Gestión de la Eficiencia Energética Mediante Técnicas de Minería de Datos

Luis Gonzaga Baca Ruiz

Tesis doctoral realizada dentro del programa:

Doctorado en Tecnologías de la Información y la Comunicación *(R.D. 99/2011)*.

Directora:

María del Carmen Pagalajar Jiménez

Grupo de Investigación:

TIC 111

Universidad de Granada
Escuela Técnica Superior de Ingeniería Informática y Telecomunicaciones
Departamento de Ciencias de la Computación e Inteligencia Artificial
Granada

Granada, noviembre de 2019.

Editor: Universidad de Granada. Tesis Doctorales
Autor: Luis Gonzaga Baca Ruiz
ISBN: 978-84-1306-407-9
URI: http://hdl.handle.net/10481/58703

Vanidad de vanidades, todo es vanidad. He reflexionado para conocer la sabiduría y el saber, la locura y la necedad, y comprendí que también esto es vanidad y caza de viento.

~ Qohelet.

Cercai l'amore dell'anima mia, lo cercai senza trovarlo. Trovai l'amore dell'anima mia, l'ho abbracciato, non lo lascerò mai.

~ Cantico dei Cantici.

AGRADECIMIENTOS

Gracias a Dios por haberme puesto en el lugar y en el momento adecuado.

A mis padres, Domingo y Encarni, por su gran paciencia y por ser el bastión donde apoyarme en los momentos más difíciles. También a mis hermanos y amigos por todo el apoyo y la ayuda que ha supuesto tenerlos a mi lado durante todo este tiempo.

Quiero agradecer también a todos aquellos compañeros que han estado a mi lado del Departamento de Ciencias de la Computación y del CITIC durante mi tesis. Especialmente a Miguel Delgado Calvo-Flores y a Manolo Pegalajar Cuéllar por estar siempre pendientes y por la confianza que han depositado en mí. Y también a Manuel I. Capel Tuñón por su apoyo. Incluyo aquí a todos los compañeros que me gustaría nombrar con los que he compartido el día a día.

A Héctor Pomares Cintas, coordinador del programa de Doctorado de Tecnologías de la Información y la Comunicación, y a Paco Blas Hernández Hoces y Beate Krug, administradores del CITIC, por su ayuda, paciencia y consejos.

Me faltarían palabras para describir mi inmensa gratitud a mi, más que tutora, María del Carmen Pegalajar Jiménez. Por su apoyo, comprensión y cariño durante todo este tiempo. Por tratarme como a un hijo y por pensar en mí antes incluso que en ti en todo momento. Gracias por cada uno de los minúsculos detalles que has tenido conmigo. Si los tuviera que nombrar uno a uno, superaría con creces el tamaño de esta memoria.

Finalmente, y en especial, a mi novia Michela que ha aparecido como un ángel de luz en medio de la oscuridad.

Gestión de la **E**ficiencia **E**nergética **M**ediante **T**écnicas de **M**inería de **D**atos

Luis Gonzaga Baca Ruiz

Tabla de Contenidos

Lista de gráficos

Pág.

<u>Artículo 3</u>

Artículo 5

Lista de tablas

Pág.

Artículo 1

Artículo 2

Artículo 3

Artículo 4

Artículo 5

Lista de abreviaturas

AE	Algoritmos Evolutivos.
AG	Algoritmo Genético.
ANN	Artificial Neural Network.
ARIMA	Autoregressive integrated moving average.
CE	Computación Evolutiva.
CHC	Cross generational elitist selection, Heterogeneous recombination, and Cataclysmic mutation.
EA	Evolutionary Algorithm.
EE	Energy Efficiency.
ENN	Elman Neural Network.
GA	Genetic Algorithm.
HVAC	Heat, ventilation and air-conditioning.
LSTM	Long Short-Term Memory.
MA	Memetic Algorithm.
MSE	Mean Squared Error.
NAR	Nonlinear autoregressive.
NARX	Nonlinear autoregressive with external input.
RNA	Red Neuronal Artificial.
SVR	Support Vector Regression.
UGR	Universidad de Granada.

Resumen

La eficiencia energética se presenta como una de las áreas con mayor interés gubernamental en los últimos tiempos. Por otro lado, el aumento de las capacidades computacionales para el procesamiento de la información y su almacenamiento, unido a la amplia disponibilidad de redes de sensores han propiciado un incremento masivo de datos en el campo de la Energía. Como consecuencia, surge el desafío de tratar y procesar la ingente cantidad de datos generados por los sistemas de gestión de los edificios con el fin de obtener indicios de comportamiento energético de los edificios, además de utilizar dicha información para la ayuda en la toma de decisiones. En este contexto, la Inteligencia Artificial emerge como herramienta para resolver dicho problema. En consecuencia, las aplicaciones específicas para la gestión eficiente de la energía están siendo recientemente explotadas. Así, este proyecto de tesis abarca la creación de modelos de predicción para estimar el consumo energético de edificios y la optimización de dichos modelos, junto con la implementación de un prototipo de software para la visualización y representación del conocimiento.

Abstract

The energy efficiency arises as one of the areas of greatest government interest in current times. On the other hand, the increase of computational capabilities for information processing and its storage, along with the wide availability of sensing nets have led to a massive increase in data production in the Energy field. Consequently, the challenge lies in treating and processing such information so as to obtain consumption profiles, in addition to use that information to support decision-making process. In this context, Artificial Intelligence emerges as an adequate tool for solving this problem. As a result, specific applications for efficient energy management are recently being exploited. Thus, this thesis comprises the creation of predictive models to estimate energy consumption of buildings and the optimization of those models, together with the implementation of a software prototype for visual energy monitoring and knowledge representation.

Motivación de la tesis

La eficiencia energética se presenta como una de las áreas con mayor interés gubernamental en los últimos tiempos. Gracias al desarrollo de nuevas tecnologías de sensores se ha facilitado la instalación de dispositivos de monitorización en multitud escenarios y, con ello, la disponibilidad de datos en tiempo real sobre aspectos de diferente índole. Además, actualmente ha habido una explosión de interés en el entorno de la minería de datos aplicada a series temporales que es, donde se encajan habitualmente este tipo de problemas. La presente tesis propone el desarrollo de herramientas y métodos para manejar y analizar la información proveniente de diversas fuentes con el objetivo final de entender cómo y cuándo se consume la energía mediante el Análisis Inteligente de Datos; combinando técnicas de Minería de Datos, Extracción del Conocimiento y Series Temporales [1].

Se entiende por minería de datos el conjunto de técnicas que permiten preparar los datos, construir el modelo y validarlo en un proceso conocido como KDD (*knowledge discovery on databases* —extracción de conocimiento en bases de datos). KDD abarca áreas de conocimiento tales como la estadística, inteligencia artificial—*Machine Learning*— y bases de datos. Dentro de la minería de datos, uno de los campos de investigación que más futuro tiene es la búsqueda de reglas o patrones ocultos en un conjunto de datos históricos. La finalidad de estos patrones es ayudar a la toma de decisiones o en la mejora de procesos productivos de una empresa [2]. El objetivo es extraer reglas que permitan adquirir conocimiento, no obvio, de bases de datos históricas.

Por otro lado, una serie temporal [3] es una sucesión ordenada en el tiempo de valores de una variable recogidas secuencialmente en el tiempo. El estudio de las series temporales puede consistir en el análisis aislado de una variable o referirse a la relación entre dos o más de ellas. En la primera aproximación se intenta entender cómo evoluciona en el tiempo una variable con el fin de realizar predicciones [4, 5]. El segundo enfoque intenta construir un modelo explicativo de la evolución temporal de una variable, con el fin de cuantificar los efectos de factores de riesgo.

Atendiendo a los puntos señalados anteriormente, la presente tesis se pretende abordar el problema de la eficiencia energética mediante el uso de técnicas de inteligencia artificial; con el fin de proporcionar modelos que, trabajando con series de datos históricos de consumo, puedan ayudar a la toma de decisiones en la gestión del mismo y su consiguiente ahorro energético anticipándose a futuros gastos energéticos.

Compendio de trabajos publicados

La presente tesis doctoral bajo el título de «Gestión de la Eficiencia Energética Mediante Técnicas de Minería de Datos» está formada por un compendio de cinco trabajos, tres de ellos publicados previamente y dos en proceso de revisión. Las referencias de dichos trabajos se recogen a continuación:

1. Ruiz, L.G.B., Cuéllar, M.P., Delgado, M. & Pegalajar MC. (2016). An application of non-linear autoregressive neural networks to predict energy consumption in public buildings. Energies, 9(9), 684. DOI: 10.3390/en9090684.

2. Ruiz, L. G. B., Rueda, R., Cuéllar, M. P., & Pegalajar, M. C. (2018). Energy consumption forecasting based on Elman neural networks with evolutive optimization. Expert Systems with Applications, 92, 380-389. DOI: 10.1016/j.eswa.2017.09.059.

3. Ruiz, L. G. B., Capel, M. I., & Pegalajar, M. C. (2019). Parallel memetic algorithm for training recurrent neural networks for the energy efficiency problem. Applied Soft Computing, 76, 356-368. DOI: 10.1016/j.asoc.2018.12.028.

4. Ruiz, L. G. B., Pegalajar, M. C., Molina-Solana, M. & Guo Y. A case study on understanding energy consumption through prediction and visualisation (VIMOEN).

5. Ruiz, L. G. B., Pegalajar, M.C., Arcucci, R. & Molina-Solana, M. A Time-Series Clustering Methodology for Knowledge Extraction in Energy Consumption Data.

INTRODUCCIÓN

1 Introducción

La Agencia Internacional de Energía indica que el sector de la edificación es el consumidor más grande que existe, representando más de un tercio del consumo final de energía a nivel mundial; así como una importante fuente de emisiones de CO_2 [6]. El mencionado consumo agrupa temas de uso cotidiano como lo son la calefacción, ventilación, aire acondicionado y sistemas de iluminación, entre otros. Habitualmente, este tipo de información está gestionada por sistemas especializados. Así, este tipo de sistemas están siendo recientemente instalados en multitud de instalaciones para mejorar la eficiencia energética de los mismos como, por ejemplo, hogares familiares, edificios industriales, comerciales, o, incluso, de dominio público como es el caso de las Universidades, e.g., la Universidad de Granada. De esta forma los edificios proveen de una serie de información potencialmente útil para conseguir ahorro energético si esta información se explota de forma adecuada.

El aumento de las capacidades computacionales para el procesamiento de la información y su almacenamiento, junto con la amplia disponibilidad de redes de sensores, han dado lugar a un crecimiento masivo de datos en diferentes dominios de aplicación. El campo de la Energía no es una excepción, de forma que la ingente cantidad de datos generados por los sistemas de gestión de los edificios puede ser procesada para obtener indicios del comportamiento energético de los edificios. En este contexto, la Inteligencia Artificial emerge como una herramienta adecuada para este objetivo. En consecuencia, las aplicaciones específicas para la gestión eficiente de la energía están siendo recientemente explotadas, como la predicción de la demanda de energía y la detección de los perfiles de consumo [7]. Sin embargo, aunque estos enfoques son apropiados para estos problemas, se basan en técnicas clásicas con alcance limitado. Este hecho dificulta la adaptación de dichas técnicas a diferentes escenarios y limita la interpretación de los resultados.

La gama de aplicaciones para el análisis de datos de eficiencia energética todavía es escasa, obstaculizada por la dificultad para recopilar, almacenar y dar sentido a una gran cantidad de datos que ha estado disponible sólo recientemente. Por tanto, se requiere de nuevas soluciones tecnológicas para impulsar la explotación de los datos relacionados con el consumo energético de los edificios. Tales soluciones tendrían el potencial de responder a las preguntas esenciales para reducir el consumo y las emisiones contaminantes, tales como la forma en que la energía se ha consumido en las instalaciones, cuál sería el consumo estimado de energía y los costes en el futuro próximo —minutos, horas, días—, los factores que tienen un mayor impacto en el uso de la energía y lo que hay que hacer para reducir el consumo al tiempo que se garantiza la comodidad. En esta tesis se propone un enfoque innovador basado en métodos avanzados de análisis inteligente de datos que se ocupa de estas cuestiones.

Centrándose en la visión europea, este proyecto está alineado con la *European Research Strategy* (Horizon 2020) dentro del *Societal Challenge Pillar*, en el reto 3: seguridad, limpieza y eficiencia energética. Además, siguiendo la política de la UE se centra en sectores donde el ahorro potencial es mayor, analizando la información para sugerir acciones para mejorar el uso de la Energía, reducir costes energéticos, incrementar la calidad de vida, etc. Finalmente, a nivel regional, la investigación a desarrollar está claramente relacionada con la estrategia *Andalusian RIS3*, particularmente con la línea de acción L74: Eficiencia energética en empresas, residencias e instituciones dedicadas a la investigación y el desarrollo tecnológico de sistemas innovadores de alta eficiencia energética en la construcción de edificios y la rehabilitación.

Como consecuencia, este proyecto de tesis abarcará la creación de modelos de predicción para estimar el consumo energético de edificios y la optimización de dichos modelos, junto con la implementación de un prototipo de software para la visualización y representación del conocimiento. Estos puntos serán las bases para el futuro planteamiento de medidas de ahorro y consecuentemente optimización del consumo.

Con esta tesis proponemos resolver tres desafíos dentro del ámbito de la eficiencia energética. Todo ello centrado en un problema real dentro de la Universidad de Granada. Los tres puntos clave se enumeran a continuación:

1. El primer punto está centrado en el estudio del problema, estado del arte de las soluciones propuestas ahora y estado del sistema actual de la Universidad de Granada. Debido a la relativamente reciente implantación de los sistemas de sensores en la UGR, hace que se disponga de un sistema todavía en fase *embrionaria*. Esto se traduce en que la calidad y cantidad de datos requiere de un tratamiento y un procesamiento exhaustivo. Además, se hace necesaria la realización de un estudio riguroso de los datos de que se disponen. Por ello, como primer desafío se propone una metodología que engloba una serie de procedimientos para tratar todo tipo de problemas relativos a la adquisición de datos para posteriormente ser utilizados en los modelos subsiguientes.

2. El segundo punto de esta tesis se centrará en estudiar y proponer métodos de Machine Learning tales como técnicas de predicción capaces de modelar el consumo energético con el fin de anticipar el gasto futuro en los edificios. Además de esto, se perseguirá un enfoque que habilite la optimización de dichos modelos con el fin de obtener las mejores predicciones del consumo. De esta forma, se obtendrán modelos optimizados con una alta precisión en el pronóstico del consumo y, por tanto, del gasto. Además, para ayudar aún más a este fin, se propondrán técnicas de agrupamiento para estudiar comportamientos y extraer conocimiento de patrones habituales en el consumo.

3. En último lugar, se pretende crear un prototipo de software que recoja toda la investigación desarrollada. Así, se completaría esta tesis con un producto software que incorporaría los modelos desarrollados en esta tesis. Además,

el software sería útil para el proceso de toma de decisiones en la eficiencia energética de los edificios, ya que permitiría al usuario de una forma clara y sencilla tener una visión general y completa del comportamiento de los mismos. De este modo, permitiría comprender cómo están los edificios consumiendo, y junto con esto, poder anticiparse a las demandas energéticas de los edificios y tomar decisiones al respecto, para favorecer el ahorro en los edificios.

Nótese que esta tesis está estructurada en dos partes bien distinguidas: en primer lugar, la tesis doctoral, y, en segundo lugar, las contribuciones científicas publicadas. La sección 1 presenta una introducción general e introduce el contexto en el que se encuentra esta tesis doctoral. La sección 1.1 reúne los objetivos principales de la tesis. La sección 2 da una visión general del estado del arte del problema que se va a abordar, así como los modelos de predicción y conceptos clave relacionados con las series temporales (sección 2.1). La metodología utilizada para desarrollar el proyecto vendrá descrita en la sección 3. La secciones 4 y 5 recogen un resumen y los resultados principales de las contribuciones científicas presentadas, respectivamente.

La segunda parte del presente documento está compuesto por los trabajos publicados en revistas de investigación reunidas en la sección 6:

1. An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings
2. Energy consumption forecasting based on Elman neural networks with evolutive optimization.
3. Parallel memetic algorithm for training recurrent neural networks for the energy efficiency problem.
4. A case study on understanding energy consumption through prediction and visualisation (VIMOEN).
5. A Time-Series Clustering Methodology for Knowledge Extraction in Energy Consumption Data.

Y finalmente, las conclusiones generales junto con futuras líneas de investigación vendrá detalladas en la sección 7.

Introduction

The International Energy Agency defines building sector as the largest energy consuming sector, representing more than one third of total energy expenditure in the world, along with a significant source of carbon dioxide emissions. The mentioned consumption results in day-to-day activities such as: heating, ventilation, air-conditioning and lighting systems, among others.

Fortunately, the increase of computational capabilities for information processing and its storage, together with the wide availability of sensing nets lead to a massive increase in data production in different application domains. The Energy field is no exception to this, and the vast amount of data generated by the building management systems may be processed to obtain behavioural patterns in buildings. In this context, Artificial Intelligence emerges as an adequate tool for this objective. As a result, specific applications for efficient energy management are recently being exploited, such as, the prediction of the energy demand and detection of consumption profiles. Nevertheless, even if those approaches are adequate for these problems, they are based on classic techniques which have limited scope. This makes it difficult to adapt them to a different scenarios and it limits the interpretability of the results.

The application range of data analysis within the energy efficiency realm is still limited, hindered by the difficulty in gathering, storing and translating that huge amount of data which has been recently available. Hence, innovative technological solutions are required so that data related to energy consumption can be exploited properly. Those solutions would have potential to answer essential questions to reduce consumption along with its polluting emissions, such as the way in which energy was used in facilities, what would be the estimated energy consumption and its cost in the near future, along with factors affecting the energy expenditure and actions to saving energy as long as it ensures comfort. In this project, an innovative approach based on advanced intelligent method of data analysis to deal with these matters.

Focusing on the European view, this thesis is aligned with the *European Research Strategy* (Horizon 2020) within the *Societal Challenge Pillar*, challenge number 3: security, cleaning and energy efficiency. In addition to this, by following the EU it is focused on particular sectors in which the potential savings is higher, analysing information in order to recommend actions to improve the energy usage, savings in energy costs, increasing standard of life, and so on. Eventually, at regional level, research to be carried out is clearly related to the *Andalusian RIS3* strategy, particularly with the L74 action line: Energy efficiency in enterprises, residences and institutions dedicated to research and technological develop of high-efficient innovative systems in building construction and building renovation.

Consequently, this thesis project will cover the creation of forecasting models to estimate energy consumption in buildings and the optimization of those models, along with the implementation of a software prototype for visual energy monitoring and knowledge representation. This will be the basis to suggest saving actions and, therefore, consumption optimization.

In this thesis we propose the solution to three challenges within the energy efficiency field. All our research is focused on a real problem in the University of Granada. The three key points are listed below:

1. The first item is focused on the study of the problem, state of the art of the proposed solutions so far, and the current state of the University of Granada's systems. Due to the relatively new implementation of sensors in the UGR, this system is still in an early stage. This means that the quality and quantity of data require an exhaustive treatment and processing of such information. Furthermore, it is necessary to perform a rigorous study of data which encompasses a series of procedures so as to address this kind of problems related to data acquisition, and then be used by the succeeding models.

2. Secondly, Machine Learning methods will be studied and proposed to solve our problem, such as, forecasting techniques capable of modelling the energy consumption to anticipate future expenditure in buildings. In addition to this, we are aiming for an approach that enables the optimization of those models so as to obtain better predictions in consumption. In this way, optimized models will be achieved with a high accuracy in energy consumption prognosis. Besides, in order to support this aim, clustering techniques will be proposed to study behaviours and to extract knowledge of usual consumption patterns.

3. Lastly, a software prototype is intended to be designed which gathers the whole research developed. In this manner, this thesis would be completed under a software product which incorporates all the models implemented during this thesis. Additionally, this software would be useful in decision-making process for improving energy efficiency in buildings, as it would allow the user to clearly have an overview and comprehensive view of the buildings in an easy and intuitive form. Hence, the software would allow us to understand how buildings are consuming, and thereby to anticipate energy demands and make decisions on this matter in order to support energy savings in buildings.

Notice that this thesis is made up of two clearly differentiated parts: First, the PhD dissertation; and second, the scientific contributions published. Section 1 makes an introduction to the problem and presents the context in which this thesis has been developed. Section 1.1 gathers the main objectives of this thesis. Section 2 gives an general state-of-art overview of the problem to address, as well as the forecasting models and key concepts associated with time-series (section 2.1). The used methodology to carry out this project is detailed in section 3. Section 4 and 5 present and gathers the main outcomes of the presented scientific contributions, respectively.

The second part of this document brings together all the studies performed and published in research journals (section 6):

1. An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings
2. Energy consumption forecasting based on Elman neural networks with evolutive optimization.
3. Parallel memetic algorithm for training recurrent neural networks for the energy efficiency problem.
4. A case study on understanding energy consumption through prediction and visualisation (VIMOEN).
5. A Time-Series Clustering Methodology for Knowledge Extraction in Energy Consumption Data.

Finally, general conclusions along with new proposals in terms of future work are detailed in section 7.

1.1 Objetivos

A modo de resumen, este trabajo se puede estructurar en una serie de objetivos que se detallan a continuación, y estableciendo como objetivo primario el estudio de la Eficiencia Energética mediante técnicas de Minería de Datos, concretamente se centrará en las Redes Neuronales Artificiales como método principal para el modelado y predicción de series de datos temporales:

1. Diseño, implementación y testeo de algoritmos de pre-procesamiento para el análisis de series temporales. Abarcando limpieza de ruido, cambio de dominio de los datos, selección de características de instancias. Junto con esto, se estudiarán técnicas de procesamiento de series temporales con el fin de obtener conocimiento e información relevante no obvia.

2. Análisis, estudio y desarrollo de algoritmos de optimización de problemas complejos. Concretamente, técnicas de modelado de series temporales para el problema de la Eficiencia Energética.

3. Desarrollo de modelos inteligentes orientados a series temporales de energía para la predicción del consumo energético, junto con la detección de patrones, irregularidades y tendencias. Aplicación de dichos algoritmos para la minería de datos predictiva destinada a la obtención de modelos de capaces de anticipar comportamientos energéticos de edificios.

4. Desarrollar y aplicar algoritmos de optimización para mejorar los modelos predictivos previamente desarrollados para estimar el consumo energético.

5. Diseñar e implementar un prototipo de software para poder incorporar los modelos creados y presentar de una forma fácil e intuitiva los resultados obtenidos en los objetivos anteriores para mejorar la eficiencia energética.

ESTADO DEL ARTE

2 Preliminares

Para trabajar este tipo de problemas es necesario reforzar las capacidades de los algoritmos de minería de datos para discernir entre información relevante e irrelevante, y que sea resistente al ruido, la incertidumbre y los valores atípicos. Surgen así las técnicas de *Soft Computing*. El término *Soft Computing* se suele emplear en informática para referirse a diversas técnicas como son la Lógica Difusa, el Razonamiento Probabilístico, las Redes Neuronales y Algoritmos Genéticos [8]. Estas técnicas, en contraste con la computación tradicional —*Hard Computing*— toleran la imprecisión, incertidumbre e información incompleta.

Se pueden identificar varias técnicas de minería de datos aplicables en este proyecto. En primer lugar, los métodos de agrupamiento se pueden emplear para clasificar objetos, expresados como matrices de valores en grupos similares. En el ámbito de la energía, los métodos de agrupación pueden ser empleados para clasificar edificios con comportamientos similar o, incluso, para identificar patrones de consumo dentro de un mismo edificio.

En segundo lugar, los métodos predictivos. Estos métodos pueden ser usados para la mejora de los sistemas de gestión de energía si son utilizados adecuadamente. En primer lugar, nos encontramos con las Redes Neuronales Artificiales (RNAs). Estos modelos permiten modelar numéricamente problemas complejos con interacciones no lineales entre variables. Las RNAs han demostrado tener éxito para crear modelos de caja negra para predecir el consumo de energía [9] y detección de fallos en entornos pequeños. Por otro lado, la Computación Evolutiva (CE) o Algoritmos Evolutivos (AE) son métodos de búsqueda y optimización basados en la evolución biológica y en el refinamiento iterativo de soluciones a través de la aplicación de diferentes «operadores genéticos». Entre las diversas aplicaciones de la CE de interés en el ámbito de la energía, se pueden encontrar la clasificación, selección de características, agrupamiento multidimensional y detección de anomalías [10].

2.1 Series temporales

Es habitual encontrar en multitud de escenarios hoy día fenómenos que son medidos periódicamente a través del tiempo, como en medicina, economía, física o geografía, entre otro. En medicina, por ejemplo, se registran regularmente los niveles de diferentes sustancias de un paciente para llevar un seguimiento y finalmente determinar si el paciente está necesitado de medicación o está todo correcto. En el campo de las finanzas se lleva un exhaustivo control de diversas variables para controlar aspectos económicos en el mercado sobre transacciones o el precio de las acciones. En el campo geográfico, podemos observar cómo, en cualquier país, se recoge anualmente la tasa de mortalidad o de crecimiento de una población. Un hecho cotidiano que forma parte del campo de la física se encarga de realizar predicciones meteorológicas lo cual lleva consigo un riguroso proceso de medición y recogida de datos para finalmente indicar si en los sucesivos días va a llover, aumentará la temperatura o estará nublado. Lo que tienen en común este tipo de escenarios es que todos ellos se resumen en medir de forma reiterada y con un cierto orden cronológico, en un intervalo concreto de tiempo, un conjunto de variables. Al conjunto de datos que se recoge siguiendo este procese se les denomina serie temporal.

Una serie temporal es un conjunto de medidas x_t tomadas periódicamente en un intervalo de tiempo [11]:

$$X = (x_{t1}, x_{t2}, \dots, x_{tn}) \mid (t_1 < t_2 < \dots < t_n) \quad \textbf{(1)}$$

Donde t_i representa el instante de tiempo en el que se tomó la muestra x_t. Asimismo, una muestra x_t puede contener más de una variable a la vez. Por ejemplo, la Figura 1 muestra un ejemplo de las tomas que se sacan para modelar la contaminación en un determinado lugar. Se toman la temperatura, humedad y dióxido de nitrógeno y de carbono.

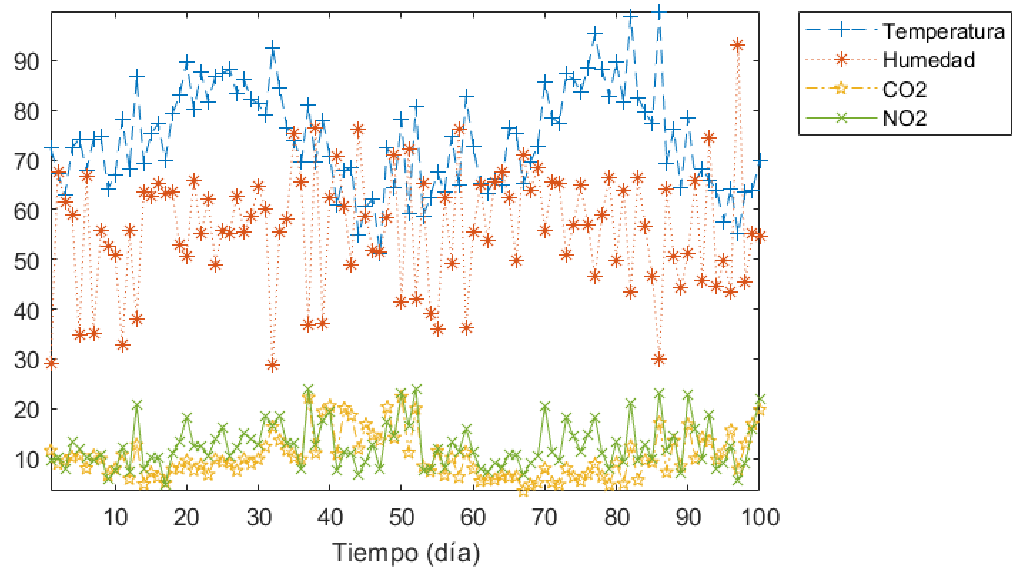


Figura 1. Ejemplo de series temporales, cuatro variables (temperatura, humedad, dióxido de carbono y dióxido de nitrógeno) para modelar la contaminación.

Habitualmente el análisis de las series temporales suele englobar procesos estadísticos y matemáticos para alcanzar fines descriptivos —analizando tendencias, influencia de periodos en el tiempo o desviaciones anómalas— o predictivos —basándose en los datos pasados, realizar una estimación futura de las series, etc.—.

En el análisis clásico de este tipo de información se suelen tener en cuenta tres componentes sobre las series [12]: 1) Tendencia. Definido como variabilidad manifestada a largo plazo. 2) Estacionalidad. Definido como la variación reflejada en los datos con cierto periodo, por ejemplo, horaria, diaria, semanal, mensual, cuatrimestral, etc., hay diversos eventos que se repiten de forma reiterada en el tiempo. 3) Aleatoriedad. Tras haber analizado las dos anteriores componentes y habiendo sido eliminados de la serie, la variación aleatoria o *residuo* es lo que queda. Y en este punto suelen entrar otros modelos probabilísticos para describir este tipo de comportamiento aleatorio *residual* que muestra la serie temporal.

Ese en este punto donde entran en juego los modelos de *Soft Computing*. En la actualidad son utilizados para modelar este tipo de comportamientos a priori arbitrarios que con otras técnicas se convierte en una tarea complicada y no se pueden tratar adecuadamente, además han demostrado dar buenos resultados en los problemas que se han aplicado [13-18]. Al igual que en el ejemplo expuesto anteriormente, en el problema que se aborda en la presente tesis se debe trabajar con una serie de variables relacionadas con energía, como, por ejemplo, la cantidad de kilo vatios consumidos cada hora, la temperatura del edificio, y todas aquellas variables disponibles que pueda proveer el sistema de gestión. De cualquier forma, es necesario estudiar modelos de predicción para estos problemas, que en definitiva se traducen en métodos de predicción de series temporales. Así el siguiente punto hará una breve introducción a los métodos actualmente utilizados en la literatura y que han sido empleados en este proyecto.

2.2 Métodos de predicción de series temporales

Los métodos de predicción de series temporales son el centro de una inmensidad de aplicaciones en la actualidad. En el área de la energía este hecho está principalmente apoyado por la instalación de tecnologías de monitorización en edificios, la cual proporciona una rica fuente de información. Sin embargo, debido a la heterogeneidad y a la diversidad de los flujos de información que existen, procedentes de diferentes sensores y orígenes, se hace complicada la tarea de manejar correctamente dicha información. Si comparamos edificios con y sin sistemas de gestión, nos encontramos que los edificios con sistemas de gestión de energía tienen costos operativos mucho más bajos que aquellos que no disponen de ellos. Por tanto, los modelos con la capacidad de predecir el consumo energético son un factor esencial en el control de los costes energéticos y reducir el impacto medioambiental. El primer objetivo de estos modelos reside en la extracción del conocimiento, traduciéndose esto normalmente en «la extracción de comportamientos habituales» y en «detección de anomalías». En este camino, se han explotado recientemente aplicaciones específicas para la construcción de sistemas eficientes de gestión energética, tales como la demanda de energía y la detección de perfiles de consumo [7].

Como se mencionaba al principio, existen infinitud de trabajos de investigación relacionadas con la predicción de series temporales, del mismo modo, hay una enorme cantidad de modelos y metodologías relacionadas a esta área aplicadas a la eficiencia energética. Como por ejemplo, para predecir la generación de energía solar y eólica algunos autores proponen el uso de árboles aditivos [19]; para predecir el consumo energético en grandes ciudades como en China hay propuestas híbridas que combinan modelos autorregresivos integrados de media móvil —ARIMA— y modelos grises [20]; la regresión lineal es también una solución, que aunque más simple que otros enfoques, han dado buenos resultados en problemas como la predicción del consumo en hogares unifamiliares [21]; y los enfoques que más van ganando popularidad son los enfoques neuro adaptativos, que también pueden ser combinados con sistemas difusos [22].

Como resultado, y debido a los excelentes resultados obtenidos en aplicaciones reales, las RNAs se han convertido en uno de los modelos más populares [23-29]. En algunos estudios se demuestra que las RNAs obtienen en general mejores resultados comparados con otras técnicas [30-33].

Una de las grandes ventajas de las redes neuronales es su capacidad de modelar relaciones no lineales en los datos. Además, una vez entrenadas, las RNAs tienen una velocidad de respuesta muy rápida. Es por ello que este proyecto de tesis irá centrado en modelos neuronales para predecir el consumo energético. Pero para que el lector pueda seguir de forma fluida el presente documento, antes han de presentarse dichos modelos brevemente. La siguiente sección abordará este aspecto.

2.3 Redes Neuronales Artificiales (RNAs)

El término de *Red Neuronal* proviene del ámbito de la Neurobiología. Si nos paramos a pensar un poco, lo primero que evoca dicho término es pensar en máquinas que guardan alguna relación con el cerebro. Y efectivamente ésta es básicamente la idea de cómo surgen estos modelos. En 1940 se introdujeron las primeras ideas acerca de cómo una neurona podía describirse de forma abstracta [34], posteriormente se propondría un procedimiento para modelar cómo podría aprender una red neuronal [35]. A partir de este momento surgen infinitud de modelos y diseños basados en cómo las neuronas del cerebro humano aprenden el conocimiento. Así, la *inteligencia* de las RNAs reside en la combinación de pequeños modelos —neuronas— que realizan operaciones muy simples con unos datos de entrada y devuelven un valor (ver Figura 2a), pero que unidas, como representa la Figura 2b, y trabajando en conjunto pueden proporcionar resultados sorprendentes en tiempos relativamente rápidos ya que pueden trabajar en paralelo [36].

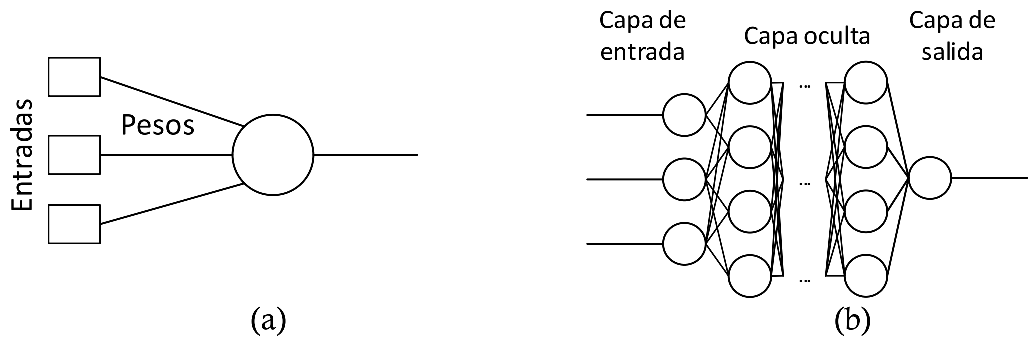


Figura 2. Ejemplo de (a) la estructura de una neurona y (b) la arquitectura de una red neuronal.

Algunos autores definen una red neuronal como «un ensamblado interconectado de unidades de procesamiento simples, llamadas unidades o nodos, cuyo funcionamiento está basado en una neurona cerebral. La capacidad de procesamiento de la red se almacena en las unidades de interconexión, o pesos, obtenidos mediante un proceso de adaptación, o aprendizaje, utilizando un conjunto de patrones de entrenamiento» [37].

La principal ventaja que presenta una RNA es su arquitectura paralela, la cual puede albergar un número considerable número de neuronas, junto con las conexiones que hay entre ellos. Este hecho permite que pequeños componentes puedan combinarse entre sí y dar como resultado una potente estructura. De forma general, una red neuronal está compuesto por neuronas y conexiones. Habitualmente, cada conexión une dos neuronas i y j, aunque puede darse el caso en el que haya más de dos neuronas interconectadas, como es el caso de las RNAs recurrentes de segundo orden. A esta conexión se le asocia un valor que sería el peso entre los dos nodos, w_{ij}, yendo la conexión desde la neurona j hasta la neurona i. Las neuronas suelen estar distribuidas por capas (ver Figura 2b), siendo las capas habituales las siguientes [36]:

1) **La capa de entrada**. En esta capa quedan recogidos los nodos que se encargan de codificar los datos que recibe la red como entrada. La

funcionalidad de esta capa es meramente distributiva. Es decir, no hacen procesamiento alguno, simplemente disponen la información de entrada al servicio del resto de los componentes en las siguientes capas.

2) **Las capas ocultas**. Las neuronas situadas entre las capas de entrada y la de salida se denominan como ocultas. En este caso pueden existir más de una capa oculta en la red. Su nombre es debido a que el procesamiento que realizan internamente se convierte en lo que se conoce como *caja negra*. Esto es, el modelo hace los cómputos en estas capas, pero su comportamiento no es evidente desde fuera. Gracias a estas capas ocultas la red es capaz de modelar aspectos no lineales en los datos que procesan.

3) **La capa de salida**. Esta es la última capa que compone una RNA. En ella se recoge toda la información procesada en las capas anteriores. Esta capa es la que dispone las salidas del sistema a través de los datos de entrada tomados por la red.

Una vez definidos los tipos de capas, se ha de definir el esquema de conexión entre las mismas. Cuando las relaciones van hacia adelante sin formar ciclos, de una capa a la siguiente, entonces recibe el nombre de *feedforward*. Es decir, las conexiones pasan de la capa de entrada a la oculta, y desde ésta a la de salida. Por otro lado, se puede considerar el diseño que incluye bucles de retroalimentación, permitiendo que haya conexiones entre nodos de capas anteriores o incluso con la misma neurona. A este tipo de enlaces se les denomina *recurrentes*.

En definitiva, una RNA está compuesta por un conjunto de pesos que definen una interconexión entre nodos de alguna forma. Estos pesos son los que se ajustan en el proceso de entrenamiento para que la red devuelva según lo que se espera de ella. Cuando el proceso de entrenamiento ha concluido esos pesos quedan fijados y es cuando se considera que los pesos asociados han sido los óptimos (producen el menor error en la salida de la red) para resolver el problema [36].

Sin embargo, los algoritmos de entrenamiento existentes hasta hoy se centran en un pequeño nicho dentro del espacio de búsqueda e intentan optimizar al máximo según los datos que están modelando. El mayor problema de esto es que fácilmente los algoritmos se quedan estancados en un mínimo local y salir de él no es una tarea que se contemple con esos métodos. Surge por tanto la necesidad de utilizar un procedimiento capaz de abordar este tipo de problema, siendo de vital importancia si se busca obtener modelos lo más óptimos posibles. Esto se traduce en modelos más precisos y de menor tamaño para realizar las predicciones y, en definitiva, poder tomar mejores decisiones para el futuro. Para este cometido es donde entran en juego los algoritmos genéticos y algoritmos multiobjetivo.

2.4 Algoritmos genéticos

Los algoritmos genéticos —AG— son unas técnicas comúnmente utilizadas para resolver problemas de optimización. De forma análoga a lo que sucede con las redes neuronales, un AG simula los mecanismos de la evolución biológica propuesta por el famoso naturalista Charles Darwin. Siguiendo esta filosofía, un AG se basa en la selección natural o la supervivencia de los más aptos. Lo que sucede en la naturaleza es que un individuo se debe adaptar al entorno que le rodea. Este proceso se le denomina evolución. Una de las características más importantes de este proceso es que los aspectos positivos de un individuo persisten en el tiempo, mientras que aquellas propiedades que hacen débil al individuo desaparecen. En otras palabras, los individuos más fuertes y más adaptados sobreviven, mientras que los más débiles se extinguen.

Un AG se puede definir como un procedimiento matemático que transforma un conjunto de representaciones de soluciones, o *individuos*. A este conjunto se le denomina *población*. A cada individuo se le asocia un valor de bondad en relación a la solución que ofrece, que es su *fitness*. De esta forma, mediante mecanismos parecidos a la reproducción en la naturaleza, los individuos se van mezclando, reproduciendo, muriendo y/o mutando según dichos operadores genéticos [38].

A modo de resumen, la Figura 3 muestra el esquema general de un algoritmo genético. Como muestra la figura, el algoritmo comienza con la generación de una población inicial. Esta población está compuesta por soluciones válidas para el problema que se pretende resolver, normalmente creadas de forma aleatoria. Tras obtener los primeros individuos se procede a comprobar la bondad de dichas soluciones, que al ser creadas de forma aleatoria comenzarán siendo, normalmente, soluciones muy mejorables. Para continuar es necesario definir alguna regla para establecer cuándo el algoritmo ha concluido, a esto se le define como *criterio de parada*. Hay diversas opciones: cuando se alcanza el óptimo (si se conoce), fijar un máximo de evaluaciones o limitar el número de iteraciones del algoritmo. En este caso, puesto que el criterio de parada no se da hasta que no se han realizado, habitualmente, un número de iteraciones del algoritmo, se procede a seleccionar los padres que se van a combinar —o *cruzar*— para producir nuevas soluciones. Una vez obtenidos los descendientes se aplica un mecanismo de mutación para otorgar al algoritmo diversidad en la exploración del espacio de soluciones. Y de nuevo se vuelve a analizar la bondad de esas soluciones calculando su *fitness*. A cada una de estas iteraciones compuestas por el conjunto de operaciones «selección, cruce, mutación y fitness» se le denomina *generación*. Esta combinación entre diversidad y convergencia provee al AG una estrategia de búsqueda que permite lidiar con el problema de los óptimos locales que otros métodos no son capaces de resolver. Y es por ese motivo que se ha optado por combinar estas técnicas para obtener los mejores modelos de predicción del consumo energético en el presente proyecto de tesis.

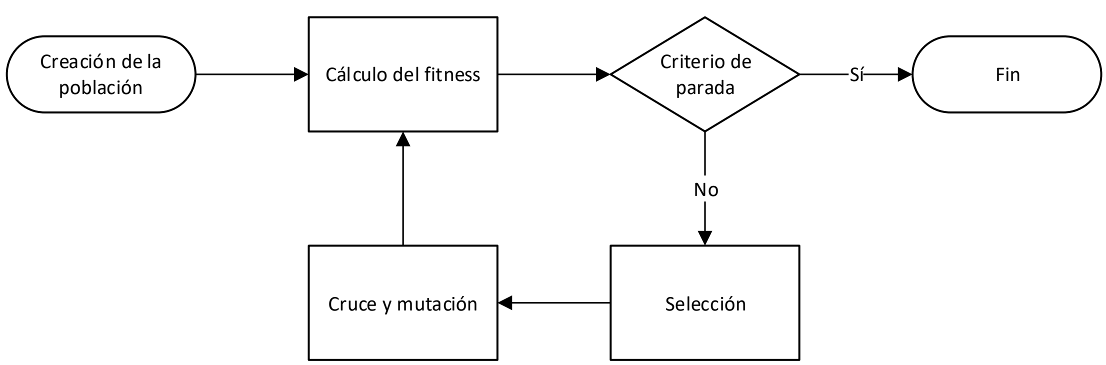


Figura 3. Esquema general de un algoritmo genético.

Téngase en cuenta que existen multitud de variaciones, modificaciones y peculiaridades dentro de los AGs. Además, según el problema a resolver existen infinitud de operadores que, dependiendo de la naturaleza del problema, tendrán mejor o peor resultado en nuestro problema. Por mencionar algunos aspectos interesantes que hacen aún más interesantes estos algoritmos, se pueden encontrar mecanismos de selección que intentan mejorar todavía más la diversidad en la población, como, por ejemplo, la *prevención de incesto* que evita que dos padres no se puedan cruzar a no ser que la diferencia entre ellos pasa un umbral. Optar por un mecanismo de reemplazo u otro puede evitar convergencias precoces del algoritmo. Los dos enfoques más conocidos son el *generacional* y el *estacionario*. El primero reemplaza completamente a la población anterior, mientras que el segundo hace un reemplazo selectivo donde padres e hijos pueden componer la siguiente población de soluciones [39]. Y el último aspecto interesante a destacar es el mecanismo de reinicialización, que es útil en varias ocasiones: cuando se llega a una convergencia prematura del algoritmo porque los individuos se parecen, o cuando se detecta que la población se ha estancado en un mínimo local. Este mecanismo hace que la población se genere de nuevo para seguir analizando el espacio de búsqueda. Este último mecanismo incluido en el conocido algoritmo de búsqueda adaptativo CHC [40] de sus siglas en inglés «*Cross generational elitist selection, Heterogeneous recombination, and Cataclysms mutation*», que como en su propio nombre se puede observar, combina a la perfección todos los mecanismos anteriormente mencionados.

Si se combina de forma adecuada los operadores de selección, cruce, mutación, reemplazamiento y reinicialización para un problema específico se puede obtener un algoritmo apropiado para explorar el espacio de soluciones de forma sistemática el cuál encontrará una buena solución a dicho problema. La tarea más compleja es por tanto diseñar la representación de las soluciones de forma conveniente y acomodar los operadores para dicho fin.

METODOLOGÍA

3 Metodología

La metodología seguida para desarrollar el presente proyecto de tesis viene recogida en las siguientes líneas. El abordaje del problema que pretendemos resolver hace necesaria conservar una metodología de trabajo. Siguiendo las líneas del método científico habitual, se proponen las siguientes formas líneas de trabajo y experimentación:

1. **Análisis y observación**. Comprobación, comprensión e interpretación del problema de la eficiencia energética y la predicción del consumo energético, junto con sus particularidad y problema concreto.

2. **Recogida de datos**. Obtención de los resultados como fruto de la aplicación de las hipótesis formuladas en los problemas abordados en distintos casos de estudio dentro de la Universidad de Granada.

3. **Formulación de hipótesis**. Diseño de nuevos métodos y modelos para mejorar la eficiencia energética. Los nuevos métodos deben cubrir las necesidades recogidas en los objetivos mencionados en la planificación de la tesis para abordar los problemas de la predicción del consumo energético.

4. **Contraste de hipótesis**. Cotejar los resultados publicados en la literatura con los obtenidos sobre la predicción y eficiencia energética.

5. **Demostración o refutación de hipótesis**. Aceptar o rehusar los modelos desarrollados a efecto de los estudios realizados. Este proceso podría repetirse en caso de necesitar alguna modificación en la hipótesis formulada.

6. **Tesis y teoría científica**. Extraer, formalizar y aceptar las conclusiones obtenidas durante todo el proceso. Todos los modelos y métodos desarrollados se concentrarán en una memoria de tesis.

RESUMEN

4 Resumen

La siguiente sección recoge las ideas principales relacionadas con los trabajos científicos de la presente tesis doctoral. En cada una de las subsecciones de la sección 5 se describen los principales logros obtenidos para cada uno de los trabajos publicados. Dichos trabajos quedan listados en las siguientes líneas:

1. Ruiz, L.G.B., Cuéllar, M.P., Delgado, M. & Pegalajar MC. (2016). An application of non-linear autoregressive neural networks to predict energy consumption in public buildings. Energies, 9(9), 684. DOI: 10.3390/en9090684.
2. Ruiz, L. G. B., Rueda, R., Cuéllar, M. P., & Pegalajar, M. C. (2018). Energy consumption forecasting based on Elman neural networks with evolutive optimization. Expert Systems with Applications, 92, 380-389. DOI: 10.1016/j.eswa.2017.09.059.
3. Ruiz, L. G. B., Capel, M. I., & Pegalajar, M. C. (2019). Parallel memetic algorithm for training recurrent neural networks for the energy efficiency problem. Applied Soft Computing, 76, 356-368. DOI: 10.1016/j.asoc.2018.12.028.
4. Ruiz, L. G. B., Pegalajar, M. C., Molina-Solana, M. & Guo Y. A case study on understanding energy consumption through prediction and visualisation (VIMOEN).
5. Ruiz, L. G. B., Pegalajar, M.C., Arcucci, R. & Molina-Solana, M. A Time-Series Clustering Methodology for Knowledge Extraction in Energy Consumption Data.

El resto de esta sección está estructurada acorde con los objetivos presentados en la sección 1.1. En primer lugar, la sección 4.1 presenta las ideas principales para resolver el problema de la predicción del consumo energético en edificios. La sección 4.2 indica los enfoques propuestos para optimizar los modelos de predicción que se han implementado. Y por último, la sección 4.3 presenta la propuesta para aprovechar el uso de los modelos implementados con el fin de obtener una forma visual de representar la información energética de los edificios.

4.1 El problema de la predicción del consumo

Existen multitud de aplicaciones y soluciones propuestas que se centran en analizar y optimizar la eficiencia energética en edificios. El principal fin que se persigue es optimizar dicho uso de energía para obtener ahorros no solo energéticos sino también ahorros económicos. En nuestro afán por resolver el problema de la Universidad de Granada para anticipar gastos en el consumo nos encontramos con la necesidad de proporcionar una metodología para dicho fin. Se hace patente la necesidad de proponer una estrategia apropiada para procesar toda la información procedente de los edificios, y con ello, crear modelos que predigan el consumo de las instalaciones.

Con esto en mente, se estudian y analizan distintos paradigmas de pre-procesamiento para el análisis de series temporales. Estas técnicas fueron estudiadas en profundidad para, finalmente optar para aquellas que fueran más apropiadas para nuestro problema. En análisis de dichas técnicas principalmente se centró en soluciones anteriores sobre problemas en el ámbito de la energía. De este modo, se consideraron aquellas técnicas que permitían tratar los datos aprovechando al máximo la información disponible sin desechar datos por presentar problemas.

Una vez hecho esto, estudiamos la viabilidad de usar los datos disponibles de los edificios de la Universidad de Granada como fuente de conocimiento para desarrollar modelos predictivos que permitiesen anticiparse de alguna forma al consumo que un edificio particular iba a tener. Sin embargo, este proceso de extracción de datos e implementación de modelos no es tan directo como a priori puede parecer. Para resolver este problema, en primer lugar, se propuso analizar el sistema proveedor de los datos de consumo energético y diseñar un procedimiento para extraer los datos de dicho sistema. Posteriormente se delineó una estrategia para tratar con los posibles problemas que podían presentar tales datos del consumo energético —datos incompletos, datos ruidosos, etc.—.

Por otro lado, se realizó un importante estudio de las técnicas más novedosas sobre predicción de datos. Tras este análisis se encontraron principalmente dos enfoques seguidos por los autores en la literatura: 1) métodos basados en técnicas clásicas de series temporales y 2) métodos basados en técnicas del ámbito de la inteligencia artificial. Los autores del primer enfoque demuestran en numerosas ocasiones la potencia y el buen rendimiento de esas técnicas. Sin embargo, cuando se comparan con las técnicas de inteligencia artificial, en raras ocasiones superan su velocidad, estabilidad y obtención de resultados.

Como fruto de todo el estudio realizado y técnicas propuestas, se propuso una metodología que fuese capaz de guiarnos desde la extracción de los datos, pasando por un procesamiento adecuado y el diseño e implementación de modelos inteligentes para realizar la predicción del consumo. En concreto, modelos de redes neuronales recurrentes fueron implementados como primera aproximación para resolver el problema, presentando buenos resultados. La publicación asociada a esta parte queda recogida en la siguiente referencia:

> *Ruiz, L.G.B., Cuéllar, M.P., Delgado, M. & Pegalajar MC. (2016). An application of non-linear autoregressive neural networks to predict energy consumption in public buildings. Energies, 9(9), 684. DOI: 10.3390/en9090684.*

4.2 Optimización de modelos de predicción

Los modelos de predicción se han convertido en una herramienta, sin duda, potente y muy útil para resolver problemas de la naturaleza que se abordan en esta tesis. Estos modelos predictivos suelen ir acompañados de algoritmos de entrenamiento para ajustar las características de los modelos que intentan ser optimizados. En particular, para las redes neuronales, como es nuestro caso, los algoritmos basados en el gradiente o similares algoritmos de entrenamiento, tienen el problema de quedar rápidamente atrapados en un mínimo local. Estos algoritmos optimizan lo que se suele decir en profundidad, pero no tienen una forma directa de solventar (salvar) dichos nichos en el espacio de soluciones.

Como solución a este problema aparecen los llamados *algoritmos genéticos* —AG—. Cientos de algoritmos de esta naturaleza han sido desarrollados en multitud de escenarios para resolver problemas de optimización. El problema que nos ocupa no es distinto, ya que los modelos con los que trabajamos son RNAs y han de ajustarse sus pesos de acuerdo a los datos que intenta modelar para obtener la mejor predicción posible con el menor error. En resumidas cuentas, nos encontramos con un problema de optimización de pesos donde el objetivo final es conseguir el mínimo error posible.

Como resultado, en nuestra investigación, proponemos utilizar ambos métodos. Es decir, aquellas técnicas que proporcionan una estrategia para optimizar las soluciones en profundidad —explorar y mejorar aún más la solución actual—, y los AG como método para realizar una exploración en anchura —explorar en el espacio de soluciones—. Sin embargo, la ventaja que otorgan estos métodos para salir de mínimos locales viene «entorpecida» de alguna forma por el costo en tiempo que requieren dichos métodos. Adicionalmente, un diseño en paralelo del método propuesto fue implementado para tratar dicha desventaja del esquema de optimización. De esta forma se obtendrían resultados óptimos disminuyendo el coste computacional de los mismos. En otras palabras, conseguir modelos capaces de predecir mejor el consumo energético en un menor tiempo. Las siguientes referencias recogen el trabajo relacionado con estos objetivos:

- *Ruiz, L. G. B., Rueda, R., Cuéllar, M. P., & Pegalajar, M. C. (2018). Energy consumption forecasting based on Elman neural networks with evolutive optimization. Expert Systems with Applications, 92, 380-389. DOI: 10.1016/j.eswa.2017.09.059.*
- *Ruiz, L. G. B., Capel, M. I., & Pegalajar, M. C. (2019). Parallel memetic algorithm for training recurrent neural networks for the energy efficiency problem. Applied Soft Computing, 76, 356-368. DOI: 10.1016/j.asoc.2018.12.028.*

4.3 Análisis y visualización del consumo energético

Una de las labores más importantes a la hora de tomar decisiones por parte del experto en energía, es la de saber interpretar y hacer uso adecuado de toda la información de que dispone sobre los edificios que está monitorizando. El experto debe de analizar todas las variables que ofrece el edificio para detectar posibles ahorros en la factura. Estas tareas normalmente suelen ser tediosas y se obvia multitud de información a causa de no disponer de una herramienta para manejar tal cantidad de datos.

De hecho, el aspecto más importante de cualquier proyecto que tiene como finalidad un producto software para un cliente, es que el software sea *intuitivo, sencillo* y *auto-explicativo*. Aunque este proyecto de tesis no está centrado en diseñar una aplicación, sino en diseñar métodos para la minería de datos energéticos, se ha querido abordar este objetivo. Téngase en cuenta que cualquier modelo inteligente desarrollado en este proyecto perdería todo su potencial si no se pone al servicio de aquellos usuarios no expertos, y es por ello que este proyecto abarca este punto, para poner su disposición dichos modelos de una forma sencilla e intuitiva.

Por tanto, considerando estos aspectos, se realizó un estudio de las soluciones expuestas ahora en la literatura y observamos que el estado del arte en el que nos encontramos carece de soluciones con este fin. Así, para llenar este hueco en la literatura se desarrollaron dos soluciones, una más analítica y otra más práctica.

La primera solución es una metodología para extraer y representar la información de los edificios de forma que se pueda extraer conocimiento oculto sobre el consumo energético de los mismos. De esta forma se propone un procedimiento para extraer automáticamente patrones en el consumo de un edificio y estudiar si dichas instalaciones presentan periodicidades en sus gastos.

La segunda propuesta está enfocada a diseñar un prototipo de software para recoger los datos que proporcionan los edificios y mostrarlos de una forma sencilla e intuitiva. Uno de los mayores objetivos a conseguir es que la aplicación fuera auto-explicativa en el mayor grado posible. De este modo, el usuario podría tener una visión general de los edificios que está monitorizando. Y no solamente esto, sino que, además de los datos propios transformados de una forma «útil» para el experto, se proporciona información extra. En otras palabras, un software inteligente que proporciona conocimiento futuro del consumo que van a realizar los edificios de forma actualizada evitando paneles de gráficos sobrecargados de información. El trabajo relacionado con esta parte de la investigación queda recogido en las siguientes referencias:

- *Ruiz, L. G. B., Pegalajar, M. C., Molina-Solana, M. & Guo Y. A case study on understanding energy consumption through prediction and visualisation (VIMOEN).*
- *Ruiz, L. G. B., Pegalajar, M.C., Arcucci, R. & Molina-Solana, M. A Time-Series Clustering Methodology for Knowledge Extraction in Energy Consumption Data.*

RESULTADOS

5 Resultados

Las secciones que siguen a continuación discuten los principales resultados obtenidos a lo largo del trabajo de investigación que ha acompañado a este proyecto de tesis. La sección 5.1 indica los resultados de los modelos implementados. Los resultados de los algoritmos de optimización para mejorar los predictores se discuten en la sección 5.2. Y finalmente se analizan brevemente las soluciones para representar el conocimiento energético en la sección 5.3.

5.1 El problema de la predicción del consumo

Suele ser muy habitual encontrarse en la literatura soluciones a problemas de difícil adaptación en otros dominios. Es decir, las aplicaciones desarrolladas para un determinado contexto, no son fácilmente aplicables cuando se intenta adaptar en un dominio completamente distinto. Esto suele deberse a que heterogeneidad y diversidad de los datos no permite una aplicación tan directa de dichas técnicas, sino que necesita de un tratamiento particular en cada caso. Cada problema es totalmente distinto de otro y requiere de un estudio exhaustivo y específico. La capacidad de generalización y flexibilidad de los modelos de Machine Learning para una adaptación adecuada hace más sencilla la resolución de estos problemas.

Es por esto que en primer lugar nuestros esfuerzos se enfocaron en analizar el problema. Una vez descubierta su naturaleza y las características de los datos, se estudiaron y examinaron los modelos actualmente utilizados en la literatura para resolver problemas de series temporales. Estudiamos los datos y los problemas que presentaban, y propusimos una serie de pasos como esquema de pre-procesado de los datos, tratando datos ruidosos e incompletos. Una vez hecho esto, llevamos a cabo la implementación de dos modelos de RNAs. Estos modelos se eligieron por su capacidad y su diseño para tratar precisamente con datos cronológicos. En la experimentación llevada a cabo se consideró el estudio de diversos parámetros para optimizar las predicciones obtenidas siguiendo el enfoque de ensayo y error. En este estudio se utilizaron los datos procedentes del sistema de monitorización de la Universidad de Granada.

Como resultado se propuso una metodología para predecir el consumo energético futuro de los edificios utilizando redes neuronales NAR —*nonlinear autoregressive*— y NARX —w*ith exogenous* inputs—. Se compararon estas dos soluciones modelando el consumo de distintos edificios de la Universidad de Granada. Los resultados obtenidos muestran como ambos modelos son adecuados para la predicción del consumo. Sin embargo, se demostró que, aunque el uso de información externa podría ser muy beneficiosa para mejorar la predicción, si no se dispone de dicha información los modelos desarrollados proporcionan también predicciones muy precisas. Finalmente, se demostró que existían similitudes en los consumos energéticos según periodos temporales en los mismos lo cual, tratando adecuadamente dicha información, podría llevarnos a una mejora notable en los modelos.

5.2 Optimización de modelos de predicción

Una vez desarrollada una metodología capaz de resolver el problema de la predicción del consumo energético en el contexto de la Universidad de Granada, se plantea la necesidad de modelos más precisos que sean capaces de modelar patrones en el consumo a más largo plazo. Es por ello que se implementaron dos nuevos modelos para este fin: las Redes Neuronales de Elman y las actualmente populares redes recurrentes con unidades de memoria a corto y largo plazo, o LSTM —del inglés *Long Short-Term Memory*—. Estos modelos, más complejos que las redes NAR y NARX, incorporan el concepto de *memoria*. De esta forma, se implementaron dos nuevos modelos de redes neuronales, en el primer caso agregamos la capacidad de memoria añadiendo una capa oculta, conocida como *capa de estado* o *capa de contexto*, donde almacenamos los valores de las neuronas ocultas con una ventana temporal. En el segundo caso, integramos la memoria con un mecanismo más complejo de *celdas* con 3 puertas, compuerta de entrada, salida y de olvido. Así regulamos el flujo de información que se va a almacenar.

Por otro lado, además de perseguir modelos más precisos modificando la estructura de los mismos, se planteó el problema de desarrollar una metodología para optimizar dichos modelos sin importar su complejidad, ya que los algoritmos clásicos de entrenamiento proporcionan una solución mejorable hasta cierto punto, quedando habitualmente estancada según el punto inicial que se lanza el método. Para este fin se desarrolló un algoritmo evolutivo híbrido cuya finalidad era optimizar los modelos de predicción en la mayor medida posible. Así, el algoritmo genético desarrollado demostró obtener mejores resultados para todas las redes neuronales implementadas y alcanzando un grado de precisión considerablemente superior a los anteriores.

Sin embargo, la ganancia en precisión y optimización de los modelos recae en la ralentización y costo temporal del algoritmo en cuestión. La principal desventaja de los algoritmos genéticos está en el tiempo de ejecución que requiere para conseguir su característica esencial: una búsqueda eficiente en el espacio de soluciones. Para tratar este problema, se desarrolló un esquema paralelo del algoritmo genético CHC, en el cual se paralelizaban todos sus operadores aprovechando la independencia existente entre cromosomas y genes. Se compararon el rendimiento y el coste computacional de las implementaciones desarrolladas en secuencial y en paralelo para obtener dichos predictores. El enfoque paralelo demostró ser computacionalmente mucho más eficiente que el secuencial, además, sin presentar impacto negativo en la calidad de las soluciones. Así, aprovechando las características intrínsecas de algoritmo y todos los recursos disponibles, este diseño paralelo consiguió mejorar en hasta en un 50% el tiempo de ejecución en el peor de los casos con un 75% en media.

5.3 Análisis y visualización del consumo energético

En esta última parte de la tesis doctoral, presentamos un modelo para analizar patrones del consumo energético y proponemos un modelo de visualización del conocimiento que sirva como herramienta de monitorización de la energía.

Por un lado, se presentó una metodología para extraer patrones de comportamiento en el consumo de los edificios por medio de técnicas de agrupamiento, o *clustering*. Nuestro método, a priori descriptivo, se centra en estudiar periodicidad en los datos que se están analizando y una vez hecho esto realizar un análisis de clustering con dicha información. Este método sirve de apoyo para mejorar en el entendimiento de cómo los edificios están consumiendo, lo cual es un aspecto clave para alcanzar ahorros energéticos.

Por otro lado, y para reunir todo el trabajo desarrollado en la presente tesis, se desarrolló un prototipo de software que incorporase estos modelos inteligentes. El propósito de esto era presentar una herramienta útil para cualquier tipo de usuario con la necesidad de trabajar con la información que proporciona el sistema para desarrollar su trabajo, con la ventaja de agregar conocimiento añadido y dotando de inteligencia al sistema. De este modo, el software desarrollado presenta el consumo energético, describiendo el comportamiento de los edificios en lenguaje natural junto con la predicción del consumo que dicho edificio va a realizar en el futuro próximo.

Los resultados mostraron que nuestra metodología es adecuada para la monitorización del consumo energético. La característica más destacable en nuestra propuesta es que el diseño llevado a cabo sigue un estilo simple y usable para cualquier usuario sin conocimientos de métodos de predicción, y evitando complejos paneles sobrecargados de información. Además, nuestro diseño permite la incorporación de nuevas instalaciones y actualización de las mismas de un modo sencillo en el sistema. El punto débil de nuestro trabajo recae sobre su propia ventaja, la simplicidad. Al querer conseguir simplicidad y usabilidad, se limita en parte la cantidad de información que puede ser visualizada y con ello limita la potencialidad del sistema. Sin embargo, nuestra propuesta modifica la forma de examinar el consumo de utilizar hojas de cálculo —la cual suele ser una tarea costosa y compleja— a utilizar un software de visualización que elimina todas las tareas intermedias y presenta la información de una forma visual e intuitiva.

COPIA DE LOS TRABAJOS PUBLICADOS

6 Trabajos publicados

6.1 An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings

Referencia
Ruiz, L., Cuéllar, M., Calvo-Flores, M., & Jiménez, M. (2016). An application of non-linear autoregressive neural networks to predict energy consumption in public buildings. Energies, 9(9), 684.
Estado
Publicado.
Factor de impacto
Factor de impacto 2.077. Índice H48.
Categoría
JCR: Posición 43/88 de la revista en el área "Energy & Fuels". Segundo cuartil. SJR: Primer cuartil en el área "Computer Science (miscellaneous)".
DOI
10.3390/en9090684
Revista ~ Editorial
Energies ~ MDPI
Logo
energies

An application of non-linear autoregressive neural networks to predict energy consumption in public buildings

L.G. Baca, M.P. Cuéllar, M. Delgado and M.C. Pegalajar *

Department of Computer Science and Artificial Intelligence, University of Granada.
c/ Periodista Daniel Saucedo Aranda, s.n., 18071 Granada, Spain.
bacaruiz@ugr.es /manupc@decsai.ugr.es / mdelgado@ugr.es /mcarmen@decsai.ugr.es*

Abstract. In the last few years, the development of new sensing technologies has led the successful installation of monitoring devices in different environments, and thereby, the wide availability of real-time data flowing from very varied elements. In this work, we use the sensor data provided by an energy consumption monitoring system in public buildings, and provide a methodology to predict future energy consumption using artificial neural networks. Our study focuses in different faculties and research centers of the University of Granada. In this context, we apply aurorregressive neural networks with and without exogenous inputs as prediction models. Results reveal that NAR and NARX neural networks are both suitable to perform energy consumption prediction, but also that exogenous data are required to improve the prediction accuracy.

Keywords: *energy efficiency; neural networks; time series prediction*

1. Introduction

According to the International Energy Agency, the building sector is the largest energy-consuming sector, accounting for one-third of final energy consumption globally, as well as an equally important source of CO_2 emissions [1]. Around 90% of buildings emissions are produced to operate the HVAC —heat, ventilation and air-conditioning— and lighting systems [2]. Thus, Energy efficiency arises as one of the more relevant objectives nowadays.

In the last few years, the development of new sensing technologies has led the successful installation of monitoring devices in different environments, and thereby, the wide availability of real-time data flowing from very varied elements. However, due to the heterogeneity and diversity of those information flows, coming from very different sensors and sources, it is difficult to appropriate manage. Buildings with energy management system have much lower operating cost than those which do not.

Therefore, forecasting models for energy consumption become an essential item for controlling energy cost and reducing the environmental impact. The first target of the models is the knowledge extraction. That means the extraction of behavioural patterns and anomaly detection. Accordingly, specific applications for building efficient energy management have been recently explored, such as prediction of the energy demand and detection of consumption profiles [3].

There are plenty of work-related with related time series prediction. As well as, there are a multitude of models and methodologies which have been applied to this aim. Gábor I. Nagy et al. [4] suggested a generalized additive tree ensemble approach to predict solar and wind power generation. Chaoqing Yuan et al. [5] used ARIMA, GM (1,1) and a hybrid model to forecast China's primary energy consumption. N. Fumo and M.A. Rafe Biswas [6] applied linear regression analysis for prediction of energy consumption in single-family homes. M. Mordjaoui et al. present a short term electric load forecasting model using an adaptive neuro-fuzzy inference system [7]. J.J. Guo et al. [8] wrote about energy consumption prediction of heat pump water heater based on grey system. F. Zhang et al. [9] applied Support Vector Regression (SVR) for building energy consumption.

The preceding paragraph lists a brief summary of articles with models and techniques used to solve time series forecasting in a specific issue. However, due to the excellent results and noteworthy success achieved in real applications [10-15], artificial neural networks (ANNs) are raised as one of the most popular models. A good few studies have illustrated that ANNs have produced better results compared with other techniques [16-19]. One of the great assets of ANNs is their potential modelling non-linear time series. Furthermore, once trained, ANNs provide swift performance of the network. Karatasou et al. [20] employed neural networks to estimate the energy consumption. Chirag Deb et al. [21] present a methodology to forecast diurnal cooling load energy consumption for institutional buildings. Gonzalez and Zamorreño [22] predicted shot term electricity by means of feedback ANN. Likewise, Camara et al. [23] applied feedforward networks to forecast energy consumption in U.S.

In the real world there are a huge amount of systems which are non-linear or whose behaviour is dynamic and depends on their current state. The dynamic recurrent neural network (RNN) and the nonlinear autoregressive (NAR), and the nonlinear autoregressive neural network with exogenous inputs (NARX) neural networks, are neural network structures that can be useful in these cases [24-25]. The first advantage of these networks is that they can accept dynamic inputs represented by time-series sets. Time series forecasting using NN is a non-parametric method. That means it is not indispensable to know any information about the process that generate the time series. On the one hand, NAR model uses the past values of the time series to predict next values. On the other hand, the RNN do not have to be organized in layers, neurons can to be linked to themselves or anyone else.

In this work, we focus on an experimentation using NAR and NARX neural networks to predict energy consumption in public buildings. The goal of the study is to provide a methodology framework to analyze the time series historical data of energy consumption, and to know if the prediction of such energy consumption for midterm can be achieved with these models. Also, we test both NAR and NARX models to study if the energy consumption depends on historical data only, or if further external variables can be used to increase the prediction accuracy. We make our study in public buildings (faculties and research centers) of University of Granada (UGR) as test case. In this University, innovative management systems have been installed in the last few years, so we can be provided with enough energy consumption data to carry out the experimentation. At least the two last years will be used in this work for each building analyzed.

The incentive of this study is to provide outcomes associated to the best model in forecasting the energy consumption in the future; and hence reaching the main purpose of this study: avoid unnecessary consumption, reducing cost over the long term, save energy and decrease the emission of harmful by products, most dangerously carbon monoxide (which is a tasteless and odorless gas with serious adverse health effects), and provide an initial framework to build more complex decision support systems.

The manuscript is organized as follows: Section 2 introduces the proposed methodology and the description of the artificial neural network methods. After that, Section 3 describes the phases of the methodology used in the study: enclosing the processing of data, noise data management, transformations and the process of training, test and validation. Section 4 shows the real data used, obtained from buildings of University of Granada. Section 5 discusses the experimentation with the proposed models and their results. Finally, Section 6 gathers the main conclusions obtained and future work.

2. Methodology

Current investigation was developed in several phases. First of them all was data collection, afterward was data selection followed by ANN modeling, and finally analysis of performance comparison between two different algorithms: NAR and NARX networks.

Previous work has shown that using ANN outstanding results are achieved [7,10-14,24,25]. Therefore, we employed two of these techniques to guarantees good results. In this way, NAR neural network models are focused on forecasting samples framed in a one dimensional time series; one input and one output are provided to this model. In this work, NAR model input is the consumption merely. Obviously, the output of NAR model will be consumption too. Considering the aim of this work is to predict energy consumption.

Likewise, NARX algorithm grants improved flexibility; it expands multidimensional time series using extra information to enhance consumption time series. Several inputs and one output are employed. We used energy consumption and certain weather information, such as temperature.

Each of which have their benefits and costs; NAR method promises to be simpler than NARX. Nevertheless, this last model promises to be more accurate. In the real world, not all buildings have the same management system or the same amount of features. Some of them only hold consumptions; others handle more information, such as external and internal temperature which we have been discussing. Selected ANN models allow working with both approaches. This section shows chosen models to solve the problem which has been described earlier.

2.1. NAR model

In the majority of cases, time series applications are characterized by high variations and fleeting transient periods. This fact makes it difficult to model time series by a liner model, therefore a nonlinear model should be suggested. A nonlinear autoregressive neural network is a recurrent network [26-27]. It describes a discrete, nonlinear, autoregressive model with and can be written as follows [28]:

$$\mathbf{y(t) = h\big(y(t-1), y(t-2), \dots, y(t-p)\big) + \epsilon(t)} \qquad (1)$$

This formula describes how a NAR network predict series $y(t)$ given p past values of $y(t)$, also illustrated in Figure **1**. The p feature is called as feedback delays. The number of hidden layers can be defined variously, and the number of neurons in each layer is completely flexible. Nevertheless, it is important to bear in mind that increasing number of neurons complicates the system.

The most common learning rule is the Levenberg-Marquardt —LMBP— [29-32]. This training function is often the fastest backpropagation algorithm. LMBP algorithm was designed to approach second-order derivative training speed without having to compute the Hessian matrix. When the performance function has the form of a sum of squares (frequently in training feedforward networks), then the Hessian matrix can be approximated as

$$H = J^{\mathbf{T}}J \qquad (2)$$

And the gradient can be computed as

$$g = J^{\mathbf{T}}e \qquad (3)$$

J is the Jacobian matrix which encloses first derivatives of the network errors with respect to the weights and biases. And e is a vector of network errors. To estimate Jacobian matrix, in [32] uses a standard backpropagation algorithm, to approximate Hessian matrix. This approach is simpler than computing the Hessian matrix.

Eventually, the LMBP algorithm uses this approach in the following Newton-like update:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}]^{-1}\mathbf{J}^T\mathbf{e} \tag{4}$$

It should be notice that this method uses the Jacobian for calculations which assumes that performance is a mean or sum of squared error. Hence, networks must use either the mean square error (MSE) or error sum of squares (SSE).

$$SSE = \sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{5}$$

$$MSE = \frac{SSE}{n-m} \tag{6}$$

Where n is the sample size and m is the number of parameters in the model.

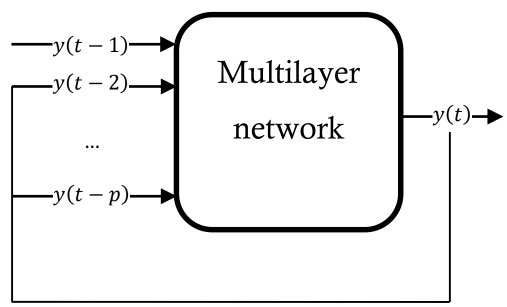


Figure 1**.** Nonlinear Autoregressive (NAR) network.

2.2. NARX model

NAR neural network only uses raw time series, without exogenous data. NAR net learns time series as is. However, many times, an important correlation exists between external data and time series. In [4], it is indicated that a usual feature of renewable energy is that the output of power plants depends largely on weather conditions. Integrate them is a basic requirement to accurate forecast, instead of a unique approach with one single value [33-34].

The model Nonlinear Autoregressive with External —Exogenous— Input (NARX) is proposed in [35], which has several advantages over simple recurrent networks. NARX predict series $y(t)$ given p past values of $y(t)$ and another —or others— series $x(t)$. It can be written as follows:

$$y(t) = h\big(x(t-1), x(t-2), \ldots, x(t-p), y(t-1), y(t-2), \ldots, y(t-p)\big) + \epsilon(t) \tag{7}$$

The NARX is a nonlinear model which estimates the future values of the time series based on its last output. Figure 2 shows its architecture, similar to NAR network. The only difference is on input, henceforth output $y(t)$ takes account of the external data as it appears on the (7). The learning rule LMBP, explained previously, is also used to train this model.

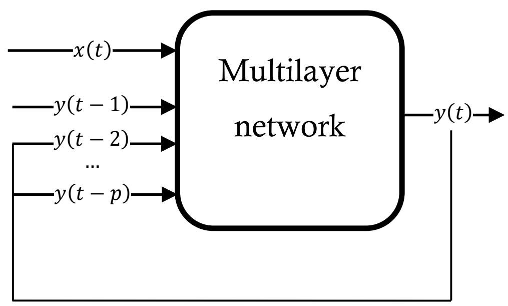


Figure 2. Nonlinear Autoregressive with External —Exogenous— Input (NARX).

3. Energy time-series modeling and forecasting

The following section describes the process to obtain an energy time-series model, along with forecasting methodology to forecasting the consumption in a building. The process overview is shown in the graph below. It would be appropriate to point out that there is no learning model indicated, so that it can support any of them indistinctively. In other words, that methodology is utilized both for NAR and NARX neural network. The chart illustrated in Figure 3 represents the methodology proposed. Each phase is detailed as follows:

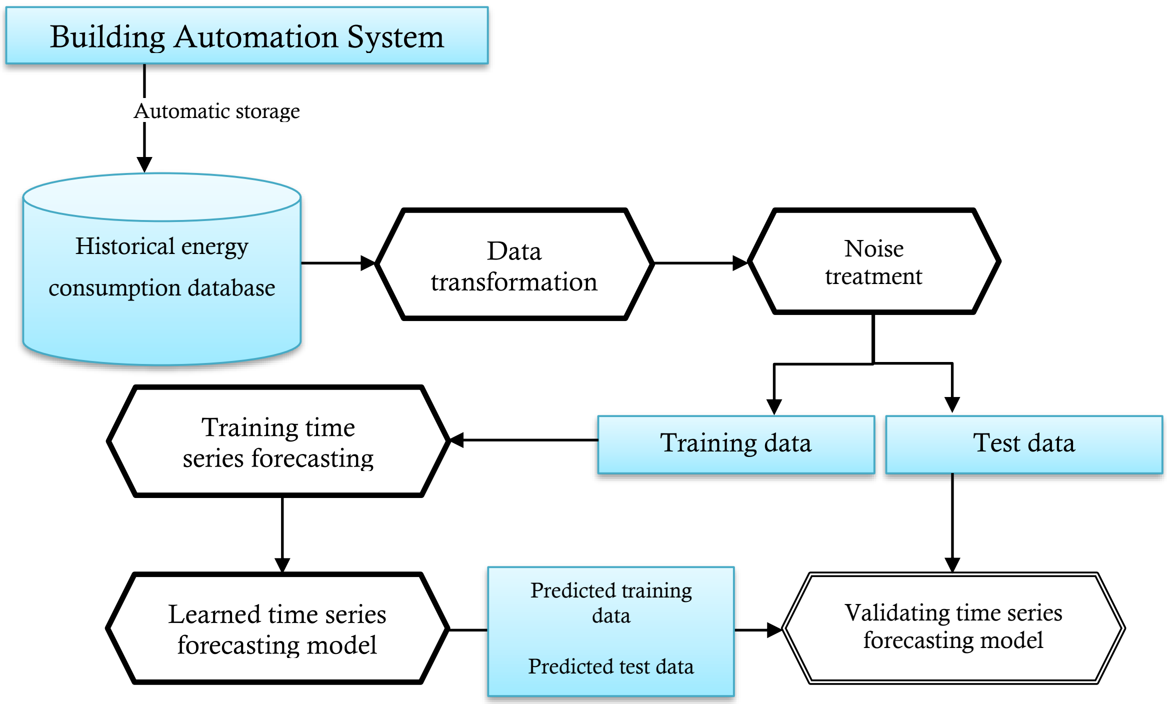


Figure 3. Flowchart of the energy time-series model.

1. **Building automation system**: it is software designed to helps building energy management efforts. It is made up of several components that provide coordinated control. It collects, summarizes and presents building data in usable way.
2. **Historical energy consumption database**: here is stored energy use of buildings. Moreover, all the consumption, energy and power recorded and monitored by diverse sensors. Apart from some external information such as temperature.
3. **Data transformation**: the building automation system is responsible for storage raw data. That is to say, collect the value in the frame counter at a certain point of time. Nevertheless, this information must be processed for the reason that we aren't always interested in what happened in a second exactly nor minutes. Maybe hours are also excessive detail in this case—always depend on the problem—. Some manuscripts propose a model for forecasting hourly [15,22,37]. However, most short-term forecasting problems involve predicting events only a few time periods: days, weeks, months [37]. Hence, this will be our approach.

4. **Noise treatment**: the data caught from physical world through sensor devices sometimes are incomplete, noisy and unreliable. This noise is mainly owing to fault detection, a broken device or a connection failure. Two main strategies are pursued: 1) Moving-average filter [38] —see **(8)**— to clean outlier. 2) Linear interpolation, based on the values at neighboring grid points to missing values.

$$y(n) = \frac{1}{windowsSize} \cdot \Big(x(n) + x(n-1) + \cdots + x\big(n - (windosSize - 1)\big)\Big) \qquad (8)$$

5. **Train and test**: The database is partitioned into a 70% training data and 30% of the test data. Splitting data randomly.
6. **Training time series forecasting model**: fitting the model with training data. NAR and NARX models have been trained using LMBP as training function. And hyperbolic tangent sigmoid transfer function [39]. The parameters will be detailed below. The model train with the 70% of the days collected, and it is tested with the remaining 30% of the days.
7. **Learned time series forecasting model**: once the model is trained and parameters have been set, it is saved. Then, prediction of training and test data are estimates; and the performance measure is calculated.
8. **Validating time series forecasting model**: the final step is to verify that appropriate results have been obtained, comparing the test data prediction with test data isolated in step 5. Such data are completely new; these have not been seen by the model.

4. Collection of data

Smart management system collects multitude of information, monitored in real-time. In this work, data are obtained from an actual Building Automation System (BAS). BAS is an automatic centralized control of building's heating, ventilation and air conditioning, lighting systems. This BAS is used by University of Granada (UGR) to monitoring his buildings consumptions, in order to analyze information coming from several sources with the final goal of better understanding how and when energy is consumed in distributed facilities. Data are generated and stored in an excessive fine-grained fashion, which hinder the processes of data analysis and decision making. The information should be condensed and simplified to extract useful forecasting.

UGR has access to energy consumption data and climatic data gathered at the buildings. In this campus, different kind of buildings and sensor setups can be found. Additionally, it has access to electricity consumption data. Teaching —units, faculties, schools and departments— are divided in five campuses — Centro, Cartuja, Fuentenueva, Aynadamar y Ciencias de la Salud— distributed in different areas of the city. In addition, UGR has offices in the autonomous cities of Ceuta and Melilla which are located in separate campus. UGR is associated with 22 faculties, 5 schools, 8 training centers and 5 culture, sport and service centers.

In this study do not apply our methodology to all buildings consumptions because of space. Otherwise, we will end up with a manuscript that is much too long. Each building consumption dataset is grouped by area. Therefore, it assumes that there is at least one relationship between those building. Eight buildings consumptions are used in this work, labelled as consumption 1, consumption 2, consumption 3… These consumptions have been selected like this: the first one is related with the second one, the thirst one is related with the fourth one, and so forth. All these consumptions are stored as a counter. Therefore, it saves the accumulated consuming up to date.

4.1. Data selection

Data selection was carried out by a first preprocessing of all data, several faculties and determined buildings. After obtaining the data that will be used to modeling ANN, the following step is to separate the data into two parts: training dataset and test dataset. The split was based on a k-fold cross-validation method to select among a set of models which have influence function representations. All of the data were normalized to have the same range of values for each of the inputs to the ANN model, and this can guarantee that there are no attributes which are more important than others, and a stable convergence of weight and biases.

4.2. Sample building

This subsection aims to summarize one sample building structure. To describe in detail its main characteristics and parameters monitored. The selected building is formed by 4 floors: basement, low level, first and second floor, see **Figure 4.**

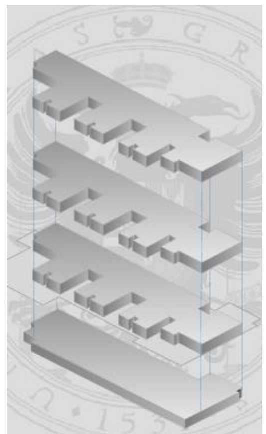


Figure 4: Sample building architecture.

Each floor monitors three different types of information: lighting, electricity, heating and air conditioning. This building has a Network Automation Engine —NAE—. It delivers comprehensive equipment monitoring and control through features like scheduling, alarm and event management, energy management, data exchange, data trending, and data storage. This tool helps lower efficiency costs for buildings. Specifically, MS-NAE5520-2 is the model used in this building (see **Figure 5**).

MS-NAE5522-2 device supports a LonWorks trunk, and two N2 Bus trunk or two BACnet MS/TP (RS-485) (or one N2 Bus trunk and one BACnet MS/TP trunk). Support a maximum of 255 devices on the LonWorks trunk and a maximum of 100 devices on each N2 Bus or BACnet MS/TP trunk. His technical characteristics are summarized as follows:

- Power requirement: dedicated nominal 24 VAC, Class 2 power supply (North America), SELV power supply (Europe), at 50/60 Hz (20 VAC minimum to 30 VAC maximum).
- Power consumption: 50 VA maximum.

- Ambient Operating Conditions: 0 to 50°C (32 to 122°F); 10 to 90% RH, 30°C (86°F) maximum dew point.
- Ambient Storage Conditions: -40 to 70°C (-40 to 158°F); 5 to 95% RH, 30°C (86°F) maximum dew point.
- Data protection battery: Supports data protection on power failure. Rechargeable gel cell battery: 12 V, 1.2 Ah, with a typical life of 3 to 5 years at 21°C (70°F); Product Code Number: MS-BAT1010-0.
- Clock battery: Maintains real-time clock through a power failure. On-board cell; typical life 10 years at 21°C (70°F).
- Processor: 1.6 GHz Intel® Atom™ processor.
- Memory: 4 GB (2 GB partitioned) flash non-volatile memory for operating system, configuration data, and operations data storage and backup 1 GB SDRAM for operations data dynamic memory.
- Operating System: Johnson Controls OEM Version of Microsoft Windows Standard 2009.
- Network and Serial Interfaces: One Ethernet port; connects at 10 Mbps, 100 Mbps, or 1 Gbps; 8-pin RJ-45 connector. Two optically isolated RS-485 ports; 9.6k, 19.2k, 38.4k, or 76.8k baud; pluggable and keyed 4 position terminal blocks (RS-485 ports available on NAE55 models only). Two RS-232-C serial ports, with standard 9-pin sub-D connectors, that support all standard baud rates. Two USB serial ports; standard USB connectors support an optional, user-supplied external modem. Options: One telephone port for internal modem; up to 56 kbps; 6-pin modular connector. One LONWORKS port; FTT10 78 kbps; pluggable, keyed 3-position terminal block (LONWORKS port available on NAE552x-x models only).
- Housing: Plastic housing with internal metal shield. Plastic material: ABS + polycarbonate UL94-5VB Protection: IP20 (IEC 60529).
- Mounting: On flat surface with screws on four mounting feet or on dual 35 mm DIN rail.

- Dimensions (Height x Width x Depth): 226 x 332 x 96.5 mm (8-7/8 x 13-1/8 x 3-13/16 in.) including mounting feet. Minimum space for mounting: 303 x 408 x 148 mm (12 x 16-1/8 x 5-13/16 in.).

Figure 5: NAE55 Network Automation Engine

This building has several meters:

- Voltage phase R, S and T-neutral (V).
- Intensity phase R, S and T (A).
- Active power phase R, S and T (kW).
- Reactive power phase R, S and T (kVAR).
- Power factor phase R, S and T.
- Three-phase active power (kW).
- Frequency (Hz).
- Voltage phases: R-S, S-T and T-R (V).
- Energy consumed (kWh)
- Neutral intensity (A).
- Maximum demand (kW).
- Total harmonic distortion (%).

Furthermore, there are other devices to monitoring lighting state. **Figure 6** show which is the state of a lamp in a room. Green state indicates a light is on, and blue state indicates a light is off. Besides, there is a pair of indicators for each room, the left one specifies the order —for example if one light is going to power off in 30 seconds it is appears blue— and the right box shows the real state. This figure shows how all the lamps are on in all the room, except in the corridor and the stairs.

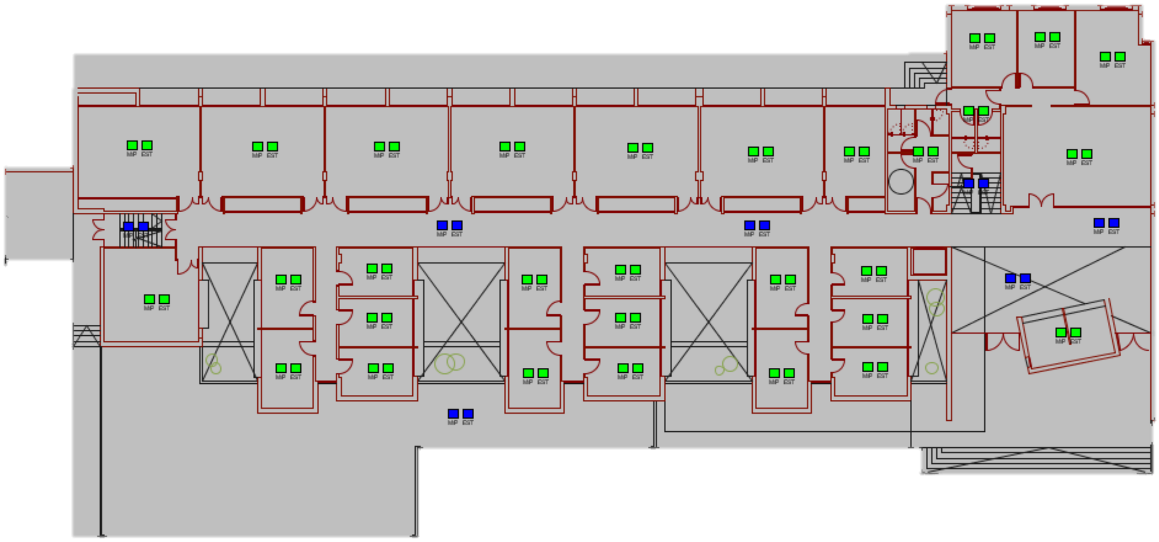


Figure 6: Lighting state of a sample floor

All this counters are gathering information 24 hours a day, every day of the year. This research is concentrated mainly in energy consumption registered and the measured temperature. **Figure 7** shows an example of a complete control ventilation system. This system manages external temperature and return temperature. It has several valves to control optimal air conditions, such as number of degrees in the environment that helps the system to close or open valves and thus introduce appropriate air. One sensor indicates if there is any problem in the system, and another sensor is used to distinguish between summer and winter. Besides, there is a handler de-icing; this one is turned on when the system detect that the temperature is too low and to prevent freezing during the winter. The most important feature in this research is the temperature. Temperature will be a useful and distinctive characteristic because of his relation with the consumption. These results will be discussed in more detail below. It is reasonable to believe that if the temperature goes too high or too low, air conditioners will be more utilized. **Figure 8** shows the consumption and the temperatures registered over time. Both time series have been normalized because each on works with different unit. Temperature is in Celsius degrees and consumption is in kW. This figure illustrates interesting trends which have been dealing before. For instance: low temperatures such as December to February have a high consumption, the air conditioners were being used more in these dates.

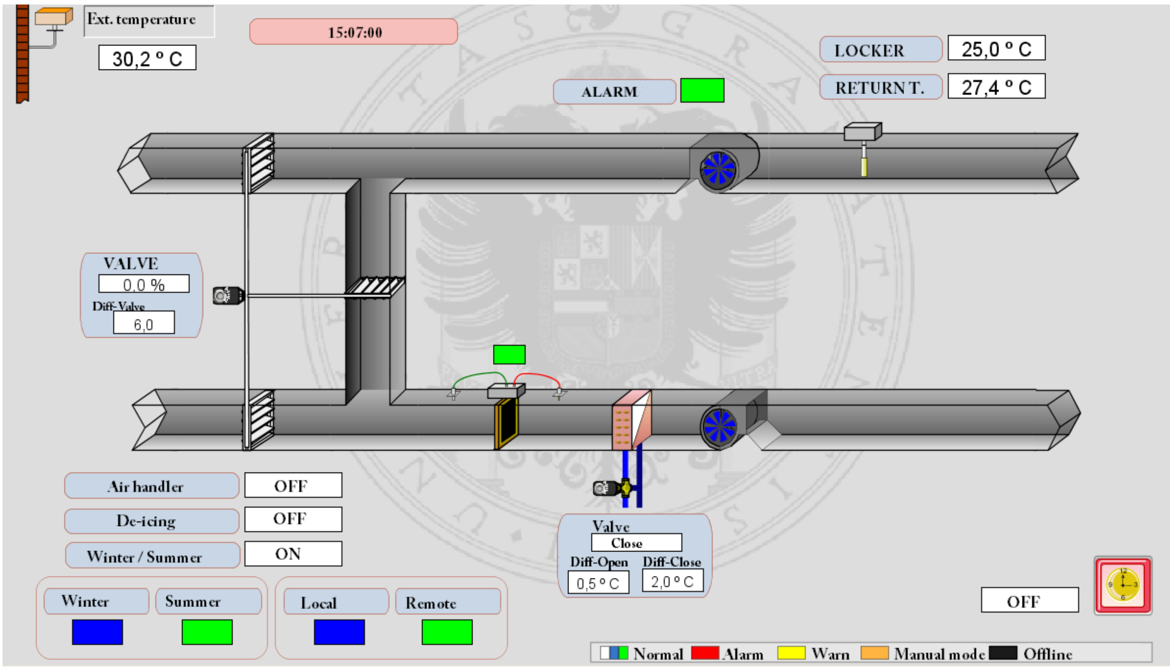


Figure 7: Sample of control ventilation system

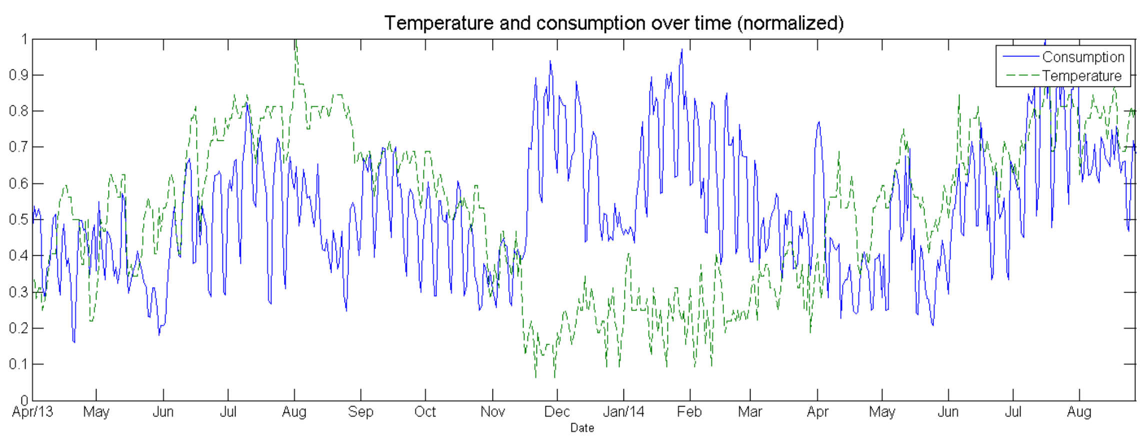


Figure 8: Temperature and consumption time series over time.

5. Results

This section describes the entire tests performed and the parameters to fit models. There are two important parameters to fit: delay and number of neurons in the hidden layer of the NN.

Figure 9 shows the behavior consumption for each day of the week. We can see in this figure that comparing closed day energy consumption is similar. For instance, if one week starts with an increase in the consumption the days around this increase have the same behavior. For this reason, appears to be important to adjust delay parameter.

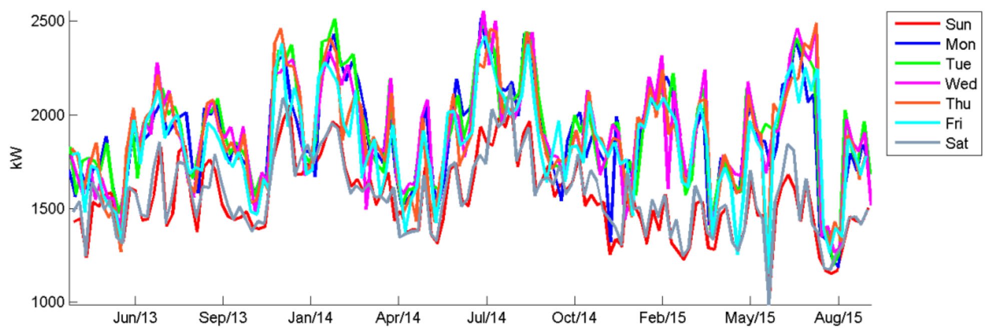


Figure 9. Example of a building consumption for every day of week

Delay parameters concern the number of days the model is going to use. In other words, the model is trained with the last p days. Tests have been executed with ten different delays. Among those are: the day before ($p = 1$), the complete last week ($p = 7$), the last two weeks ($p = 14$) and the last three weeks ($p = 21$). With the aim to find and get close best delay parameter possible, the rest of the parameters have been set. The Table 1 illustrates best delay obtained for each building. In all cases the worst error is acquired with only 1 delay. From this point, the reduction of the MSE begins. It becomes stagnant with the last 9 days; two of them extend to two weeks to get best result. From these results, conclusions can then be made relative to the past information: between past 9 and 14 days are needed to get a good forecasting NAR model.

Note that results shown in all tables derive from the average of 10 executions each cell.

Table 1. MSE of the delay parameter obtained with the NAR model

	Delay MSE									
	1	**3**	**5**	**7**	**9**	**12**	**14**	**16**	**18**	**21**
C1	~~*5.80%*~~	2.50%	2.28%	2.13%	2.06%	2.02%	**2.01%**	2.03%	2.22%	2.05%
C2	~~*3.82%*~~	2.28%	1.80%	1.53%	1.53%	**1.42%**	1.52%	1.61%	1.49%	1.57%
C3	~~*4.80%*~~	2.24%	2.09%	1.83%	1.53%	**1.30%**	1.66%	1.51%	1.69%	1.76%
C4	~~*4.87%*~~	2.68%	2.16%	1.78%	1.95%	1.85%	**1.72%**	1.75%	1.82%	1.83%
C5	~~*1.70%*~~	1.21%	1.17%	0.85%	**0.71%**	0.78%	0.74%	0.79%	0.75%	0.73%
C6	~~*2.95%*~~	1.75%	1.08%	0.80%	**0.62%**	0.81%	0.80%	0.78%	0.85%	0.88%
C7	~~*5.41%*~~	3.12%	1.87%	1.67%	1.58%	**1.35%**	1.51%	1.66%	1.62%	1.67%
C8	~~*3.29%*~~	1.97%	1.59%	1.08%	**0.94%**	1.06%	1.08%	1.09%	1.03%	1.02%
Mean	~~*4.08%*~~	2.22%	1.76%	1.46%	1.37%	**1.32%**	1.38%	1.40%	1.43%	1.44%

~~*Worst*~~
Best

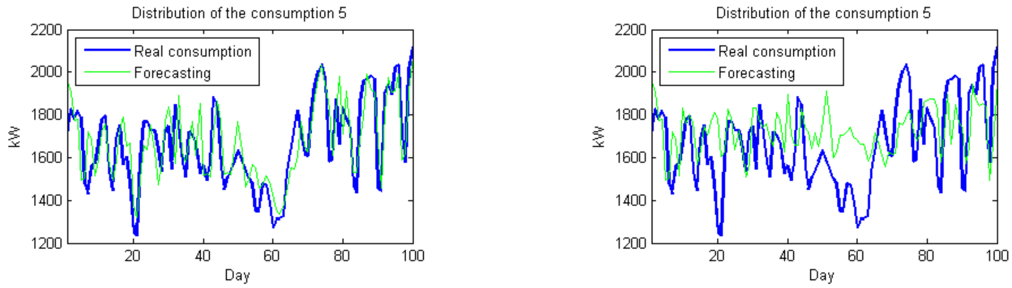
Table 2. MSE of the neurons parameter obtained with the NAR model

	Neurons MSE									
	2	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**
C1	~~*2.27%*~~	1.83%	**1.82%**	1.87%	2.10%	2.20%	2.13%	2.08%	2.14%	2.15%
C2	1.63%	1.78%	1.48%	1.54%	1.64%	**1.44%**	1.62%	1.61%	~~*1.94%*~~	1.67%
C3	1.81%	1.69%	**1.54%**	1.67%	1.63%	1.68%	1.59%	1.58%	~~*1.91%*~~	1.64%
C4	1.78%	1.90%	1.75%	1.73%	**1.70%**	1.90%	1.85%	1.77%	~~*1.93%*~~	1.87%
C5	0.72%	0.70%	0.70%	**0.67%**	0.74%	0.69%	0.75%	0.72%	0.76%	~~*0.79%*~~
C6	0.65%	0.78%	0.85%	0.77%	0.74%	**0.64%**	0.76%	0.80%	~~*0.90%*~~	0.87%
C7	1.82%	1.49%	1.70%	~~*1.89%*~~	1.62%	**1.36%**	1.73%	1.58%	1.71%	1.52%
C8	~~*1.20%*~~	1.10%	0.96%	0.96%	**0.93%**	1.00%	0.95%	1.02%	0.94%	1.03%
Mean	1.49%	1.41%	**1.35%**	1.39%	1.39%	1.36%	1.42%	1.40%	~~*1.53%*~~	1.44%

~~*Worst*~~
Best

Next point to adjust is the number of neurons needed to train the NAR network. In this case, delay parameter has been set with the best MSE attained. Table 2 collects experiments with number of hidden neurons, hereupon, it emerges that more neurons results in worse prediction. Indeed, the best average errors calculated is obtained with 6 neurons. Increase number of neurons runs the risk of over training. It is interesting to note that best delay was set between 9 and 14, and similarities outcomes get the number of hidden neurons: between 6 and 12 neurons.

We can see the result obtained for consumption 5 during 100 days in Figure 10. The selected consumption is the fifth because it has weather information associated, and we will able to compare NAR and NARX models at the end of this section. First one presents one of the best forecasting curves captured, with a MSE of 0.68 percent, and the second shows one of the worst forecasting, with a MSE of 2.23 percent.



(a) (b)

Figure 10. Best prediction for C5 (a) and worst prediction for C5 (b) using NAR model

All those results have been got by using only consumption information, without external data. Only the energy time series was used. NARX model allow extra information that could be useful. Thus, all the before experiments are going to be executed again but including temperature recording. Table 3 shows how the numbers of days required in this case to get the best model is below the week. And the worst are over two weeks in almost all cases.

Table 3. MSE of the delay parameter obtained with the NARX model

	Delay MSE									
	1	**3**	**5**	**7**	**9**	**12**	**14**	**16**	**18**	**21**
C1	1.83%	2.12%	1.63%	**1.50%**	1.52%	1.82%	~~*2.14%*~~	2.06%	1.84%	2.07%
C2	~~*1.77%*~~	1.69%	1.45%	**1.39%**	1.53%	1.48%	1.44%	1.40%	1.58%	1.59%
C3	1.59%	1.47%	1.51%	**1.35%**	1.58%	1.46%	1.55%	~~*1.90%*~~	1.66%	1.76%
C4	1.69%	**1.58%**	**1.58%**	1.71%	1.71%	1.83%	1.80%	1.99%	1.94%	~~*2.16%*~~
C5	0.91%	0.79%	**0.76%**	0.79%	0.83%	0.87%	0.90%	0.93%	0.95%	~~*1.01%*~~
C6	1.08%	0.83%	**0.77%**	1.03%	1.12%	1.16%	1.13%	1.24%	~~*1.44%*~~	1.31%
C7	1.33%	1.44%	**1.17%**	1.27%	1.36%	1.35%	~~*1.71%*~~	1.54%	1.57%	1.65%
C8	1.03%	1.09%	1.05%	**0.99%**	**0.99%**	1.06%	1.30%	1.14%	~~*1.34%*~~	~~*1.34%*~~
Mean	1.40%	1.38%	**1.24%**	1.26%	1.33%	1.38%	1.50%	1.52%	1.54%	~~*1.61%*~~

~~*Worst*~~
Best

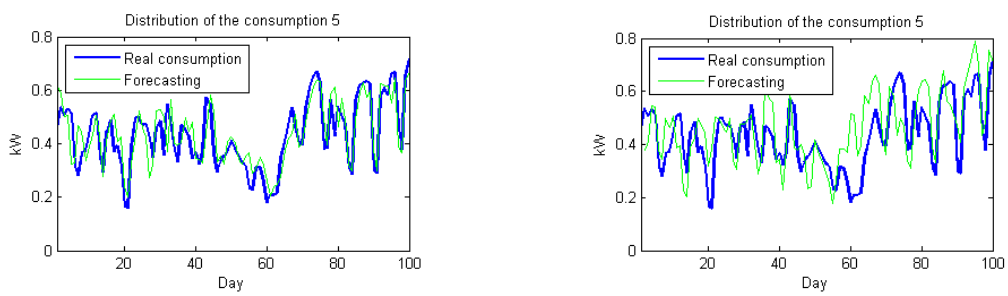
In the same way as NAR model, Table 5 shows the results of the test with the number of hidden neurons. On this occasion, however, the best results achieved have been with less number of neurons. The best average MSE involve the two neurons, despite the fact that there are several data which need neurons to improve their errors.

Table 4. MSE of the neurons parameter obtained with the NARX model

	Neurons MSE									
	2	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**
C1	1.70%	**1.33%**	1.66%	~~1.88%~~	1.70%	1.73%	1.70%	1.59%	1.69%	1.86%
C2	**1.31%**	1.47%	1.50%	1.44%	1.65%	1.74%	1.77%	1.54%	1.66%	~~1.78%~~
C3	1.24%	1.45%	**1.21%**	1.27%	1.28%	1.52%	1.34%	1.37%	1.49%	~~1.64%~~
C4	1.45%	1.74%	1.63%	1.70%	~~1.85%~~	1.77%	**1.40%**	1.62%	1.71%	1.80%
C5	**0.71%**	0.82%	0.74%	0.86%	0.85%	0.82%	0.78%	0.82%	0.82%	~~0.90%~~
C6	**0.79%**	0.87%	0.84%	0.85%	1.08%	0.94%	1.01%	1.03%	1.00%	~~1.10%~~
C7	**1.17%**	1.39%	1.27%	1.35%	1.31%	1.31%	1.24%	~~1.67%~~	1.59%	1.30%
C8	**0.88%**	1.05%	1.03%	1.01%	1.01%	1.06%	1.08%	1.11%	~~1.12%~~	1.01%
Mean	**1.16%**	1.27%	1.24%	1.30%	1.34%	1.36%	1.29%	1.34%	1.39%	~~1.42%~~

~~Worst~~
Best

Figure 11 reveals forecasting curves using NARX models. The best result obtained, with a MSE of 0.66 percent is similar to NAR model. Likewise, worst prediction curve looks like NAR network, however, this time, follows the trend of data. NARX models fit better the rapid rise and fall of consumption.



(a) (b)

Figure 11. Best prediction for C5 (a) and worst prediction for C5 (b) using NARX model

Finally, Figure 12 compares the means obtained with NAR and NARX models. Both graph shows NARX's outcomes are better than NAR's. The curve of MSE in the case of NARX network is under the other one in majority of cases. Except with the delay parameter which is extremely close when it is beyond 12 days.

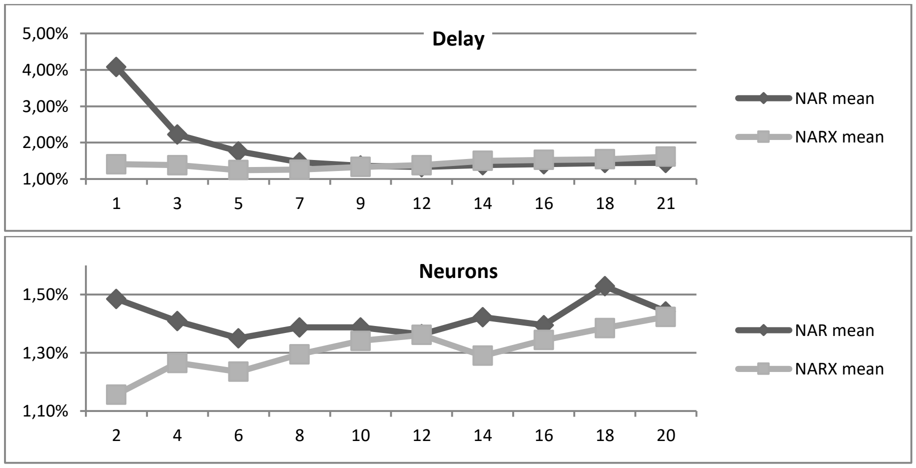


Figure 12. Comparison mean MSE of the delay parameter (up) and the neurons parameter (down).

6. Conclusions

In this work we have provided a methodology to perform energy consumption prediction using NAR and NARX models: First, the building automation system is equipped with energy consumption sensors that store the consumption data into a shared data base. From the sensor data in this data base, we applied data preprocessing techniques to transform the data and remove the data noise. After that, we have tested different prediction models coming from the neural network area: NAR and NARX. Although both approaches obtain a suitable prediction accuracy, we have concluded that there are similarities in the energy consumption of each building that, being used as exogenous inputs, can be used to improve the prediction accuracy of the remaining buildings. Future works will attempt to use these models to build high-level recommendation and decision support systems, which can aid the managers to reduce energy consumption.

Acknowledgments

This work has been developed with the support of the Department Computer Science and Artificial Intelligence of University Granada, TIC111, and the project TIN201564776-C3-1-R.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN: Artificial neural network.
ARIMA: Autoregressive integrated moving average.
BAS: Building automation system.
HVAC: Heat, ventilation and air-conditioning.
LMBP: Levenberg-Marquardt backpropagation.
ML: Machine learning.
MSE: Mean square error.
NAE: Network Automation Engine.
NAR: Nonlinear autoregressive.
NARX: Nonlinear autoregressive with external input.
NN: Neural network.
RNN: Recurrent neural network.
SVR: Support vector regression.
UGR: University of Granada.

References

[1] International Energy Agency (2013) Transition to sustainable buildings – Strategies and opportunities to 2050. [CrossRef]

[2] United Nations Environment Programme (2009) Buildings and climate change – Summary for decision-makers. [CrossRef]

[3] Data Mining for building energy management. Proceedings of Sustainable Places 2014. Nice, France. **2014.**

[4] Nagy, G.I.; et al. GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach. *International Journal of Forecasting*, **2016**. *32*(3): p. 1087-1093. [CrossRef]

[5] Yuan, C.; S. Liu; and Z. Fang, Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM (1, 1) model. *Energy*, **2016**. *100*: p. 384-390. [CrossRef]

[6] Fumo, N.; and M.R. Biswas, Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews*, **2015**. *47*: p. 332-343. [CrossRef]

[7] Haydari Z; Kavehnia F; Askari M; Ganbariyan M. Time-series load modelling and load forecasting using neuro-fuzzy techniques. In: 9th International conference on electrical power quality and utilisation; 2007. p. 1–6. [CrossRef]

[8] J.J. Guo; J.Y. Wu; R.Z. Wang. A new approach to energy consumption prediction of domestic heat pump water heater based on grey system theory. *Energy Build*, **2011,** *43*, p. 1273–1279. [CrossRef]

[9] Zhang, F., et al., Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. *Energy and Buildings*, **2016**. *126*: p. 94-103. [CrossRef]

[10] Doucoure, B.; K. Agbossou; and A. Cardenas, Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data. *Renewable Energy*, **2016**. *92*: p. 202-211. [CrossRef]

[11] Macas, M., et al., The role of data sample size and dimensionality in neural network based forecasting of building heating related variables. *Energy and Buildings*, **2016**. *111*: p. 299-310. [CrossRef]

[12] Ding, N., et al., Neural network-based model design for short-term load forecast in distribution systems. *IEEE Transactions on Power Systems*, **2016**. *31*(1): p. 72-81. [CrossRef]

[13] S.J. Kiartzis; A.G. Bakirtzis; V. Petridis, Short-term forecasting using NNs, *Electric Power Systems Research* **1995** *33*: p. 1–6

[14] Panapakidis, I.P. and A.S. Dagoumas, Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy*, **2016**. *172*: p. 132-151. [CrossRef]

[15] Chae, Y.T., et al., Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. *Energy and Buildings*, **2016**. *111*: p. 184-194. [CrossRef]

[16] Neto, A.H. and F.A.S. Fiorelli, Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. *Energy and buildings*, **2008**. *40*(12): p. 2169-2176. [CrossRef]

[17] Aydinalp-Koksal, M. and V.I. Ugursal, Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector. *Applied Energy*, **2008**. *85*(4): p. 271-296. [CrossRef]

[18] Belshaw, M., et al. Towards a single sensor passive solution for automated fall detection. in 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2011. IEEE. [CrossRef]

[19] Mehr, H.D.; H. Polat; andA. Cetin. Resident activity recognition in smart homes by using artificial neural networks. in 2016 4th International Istanbul Smart Grid Congress and Fair (ICSG). 2016. [CrossRef]

[20] Karatasou, S.; M. Santamouris; and V. Geros, Modeling and predicting building's energy use with artificial neural networks: Methods and results. *Energy and Buildings*, **2006**. *38*(8): p. 949-958. [CrossRef]

[21] Deb, C., et al., Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks. *Energy and Buildings*, **2016**. *121*: p. 284-297. [CrossRef]

[22] González, P.A. and J.M. Zamarreno, Prediction of hourly energy consumption in buildings based on a feedback artificial neural network. *Energy and Buildings*, **2005**. *37*(6): p. 595-601. [CrossRef]

[23] Camara, A.; W. Feixing; and L. Xiuqin, Energy Consumption Forecasting Using Seasonal ARIMA with Artificial Neural Networks Models. *International Journal of Business and Management*, **2016**. *11*(5): p. 231. [CrossRef]

[24] Cao, Q.; B.T. Ewing; and M.A. Thompson, Forecasting wind speed with recurrent neural networks. *European Journal of Operational Research*, **2012**. *221*(1): p. 148-154. [CrossRef]

[25] Mohanty, S.; P.K. Patra; and S.S. Sahoo. Prediction of global solar radiation using nonlinear auto regressive network with exogenous inputs (narx). in 2015 39th National Systems Conference (NSC). 2015. IEEE. [CrossRef]

[26] Nyanteh, Y.D., et al., Application of artificial intelligence to stator winding fault diagnosis in Permanent Magnet Synchronous Machines. *Electric Power Systems Research*, **2013**. *103*: p. 201-213. [CrossRef]

[27] López, M., et al., Application of SOM neural networks to short-term load forecasting: The Spanish electricity market case study. *Electric Power Systems Research*, **2012**. *91*: p. 18-27. [CrossRef]

[28] Ibrahim, M., et al., Nonlinear autoregressive neural network in an energy management strategy for battery/ultra-capacitor hybrid electrical vehicles. *Electric Power Systems Research*, **2016**. *136*: p. 262-269. [CrossRef]

[29] Alwakeel, M. and Z. Shaaban, Face Recognition based on haar wavelet transform and principal component analysis via levenberg-marquardt backpropagation neural network. *European Journal of Scientific Research*, **2010**. *42*(1): p. 25-31. [CrossRef]

[30] Marquardt, D.W., An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, **1963**. *11*(2): p. 431-441. [CrossRef]

[31] Hagan, M.T.; H.B. Demuth; and M.H. Beale, Neural network design, PWS Pub. *Co., Boston*, **1996**. *3632*.

[32] Hagan, M.T. and M.B. Menhaj, Training feedforward networks with the Marquardt algorithm. *IEEE transactions on Neural Networks*, **1994**. *5*(6): p. 989-993. [CrossRef]

[33] Hong, T., et al., Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *International Journal of Forecasting*, **2016**. *32*(3): p. 896-913. [CrossRef]

[34] Aydinalp-Koksal, M. and V.I. Ugursal, Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector. *Applied Energy*, **2008**. *85*(4): p. 271-296. [CrossRef]

[35] Lin, T., et al., Learning long-term dependencies in NARX recurrent neural networks. *IEEE Transactions on Neural Networks*, **1996**. *7*(6): p. 1329-1338. [CrossRef]

[36] Safavieh, E.; S. Andalib; and A. Andalib. Forecasting the unknown dynamics in NN3 database using a nonlinear autoregressive recurrent neural network. *International Joint Conference on Neural Networks*, **2007**. *IEEE*. [CrossRef]

[37] Alvarez, F.M., et al., Energy time series forecasting based on pattern sequence similarity. *IEEE Transactions on Knowledge and Data Engineering*, **2011**. *23*(8): p. 1230-1243. [CrossRef]

[38] S. W. Smith, Moving average filters, *The Scientist and Engineer's Guide to Digital Signal Processing,* **1999**, *2*: p. 277-284. *California Technical Publiching*, San Diego, CA, USA.

[39] Vogl, T.P., et al., Accelerating the convergence of the back-propagation method. *Biological cybernetics*, **1988**. *59*(4-5): p. 257-263. [CrossRef]

6.2 Energy Consumption Forecasting based on Elman Neural Network with Evolutionary Optimization

Referencia

Ruiz, L. G. B., Rueda, R., Cuéllar, M. P., & Pegalajar, M. C. (2018). Energy consumption forecasting based on Elman neural networks with evolutive optimization. Expert Systems with Applications, 92, 380-389.

Estado

Publicado.

Factor de impacto

Factor de impacto 3.768.

Categoría

Posición 20/132 en el área “Computer science, artificial intelligence”. Primer cuartil.

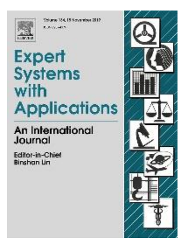
DOI

10.1016/j.eswa.2017.09.059

Revista ~ Editorial

Expert Systems With Applications ~ Elsevier.

Logo



Energy consumption forecasting based on Elman neural networks with evolutive optimization

L.G.B. Ruiz*, R. Rueda, M.P. Cuéllar and M.C. Pegalajar

Department of Computer Science and Artificial Intelligence, University of Granada.
c/ Periodista Daniel Saucedo Aranda, s.n., 18071 Granada, Spain.
bacaruiz@ugr.es* / ramonrd@ugr.es / manupc@decsai.ugr.es / mcarmen@decsai.ugr.es

Abstract. Buildings are an essential part of our social life. People spend a substantial fraction of their time and spend a high amount of energy in them. There is a grand variety of systems and services related to buildings, in order to better control and monitoring. The prompt taking of decisions may prevent costs and contamination. This paper proposes a method for energy consumption forecasting in public buildings, and thus, achieve energy savings, in order to improve the energy efficiency, without affecting the comfort and wellness. The prediction of the energy consumption is indispensable for the intelligent systems operations and planning. We propose an Elman neural network for forecasting such consumption and we use a genetic algorithm to optimize the weight of the models. This paper concludes that the proposed method optimizes the energy consumption forecasting and improves results attained in previous studies.

Keywords: *energy efficiency; neural networks; time series forecasting; evolutionary algorithm*

1. Introduction

Energy efficiency is an area of increasing importance because of the rising cost of energy and growing environmental concerns. The building sector consumes one-third of the global consumption and records close to 40% of global CO2 emissions (International Energy Agency, 2013). The operations of building systems produce approximately 90% of these pollutant gases, in order to maintain a certain comfort level, closely associated with the heating, cooling and lighting systems. The worries over energy consumption and related noise, light pollution and safety problems have motivated the sustainable building facilities by means of construction practices and context-sensitive design (Olubunmi, Xia, & Skitmore, 2016).

Due to the increased awareness of environmental issues and energy security, building regulations and polices related to new and refurbished building have been established in many countries: USA with the LEED —Leadership in Energy and

Environmental Design— and BREEAM —Building Research Establishment Environmental Assessment Methodology— and Europe with the Energy Performance of Buildings Directive (Lord, Noye, Ure, Tennant, & Fisk, 2016). Architects, planners and engineers are increasingly requiring that consider energy codes for minimizing environmental impact and resource consumption.

The most significant decisions linked to sustainable design are usually made in the early design stages, determining its environmental impact and its energy costs (Basbagill, Flager, Lepech, & Fischer, 2013). Commonly, energy analysis is habitually performed after the architectural design and related documents have been produced. This practice into the design process leads to an inefficient way of retroactively modifying the design afterwards to achieve a set of performance criteria (Jalaei & Jrade, 2014). Energy efficiency is a decisive quality in order to reach environmentally friendly buildings, and what's more, is an effective strategy for reducing energy consumption and related gas emissions, with the consequent economic savings this can represent.

In more recent years, the new sensing technologies are continually being developed and integrated in the most diverse environments (Ekwevugbe, Brown, Pakka, & Fan, 2013). These provide useful and descriptive information of the building if we know how to take advantage of the powerful data. Nevertheless, there is a marked diversity in the data flows, owing to its irregular and varied source, coming from heating, ventilation, air-conditioning and lighting systems obtaining information such as internal and external temperature, sound level, carbon-dioxide, energy consumed, intensity, maximum demand, lighting state, wind speed, wind direction, pressure, precipitation (Khosravani, Castilla, Berenguel, Ruano, & Ferreira, 2016; Ruiz, Cuellar, Delgado, & Pegalajar, 2016) or even occupancy (Balaji, Xu, Nwokafor, Gupta, & Agarwal, 2013); making its treatment difficult. The monitoring systems offers a possibility of collecting and storing a vast quantity of data. Processing all this information is not a trivial undertaking, this task frequently requires the combination of different datasets that might be not related a priori. The crucial need for analysing big amount of data has revolutionized the Machine Learning, Data Mining and Statistics using prediction, classification, regression, clustering and dimensionality reduction techniques (Balón-Canedo, Remeseiro, Sechidis, Martinez-Rego, & Alonso-Betanzos, 2017); many tools have been developed in genomics —enabling inexpensive and high-throughput measurement of the genome—, neuroscience —important diseases have been shown to be related to bran connectivity networks—, economics and finance —implementing specialized data analytics programs to identify key business insights that can be exploited to support better decision making—, social network —data analysis of data produced by Twitter, Facebook, LinkedIn and YouTube using these data to predict influenza epidemic or stock market trend— (Fan, Han, & Liu, 2014). By extension, to make an efficient use of energy, in view of achieving remarkable reductions in consumption and significant economic saving, becomes an important and challenging issue. Lately, specific applications for building efficient energy have been investigated, as is the case of the energy forecasting (Andrade & Bessa, 2017) and consumption patterns (Chou & Ngo, 2016).

Forecasting models for energy consumption furnish intelligence with in a building for improving energy use, cost saving and reducing environmental impact without the need to compromise on performance and comfort. Predictive management of a building system can reduce peak power demands which translates into energy savings (Dhillon, Rahman, Ahmad, & Hossain, 2016).

An energy prediction model represents an essential role in smart buildings. It has been proven that a small increase in forecasting accuracy would save millions of dollars in operation costs (Bunn & Farmer, 1985). The time series prediction is habitually handled as a hard paradigm because there may be diverse influencing factors, like weather conditions, social and economic conditions. There is an abundant research literature focused on time series prediction, the most popular methods are collected by Palli's book (Pallit & Popovic, 2005), some examples are Regressive Models, Artificial Neural Network —ANN—, Trees, Fuzzy methods and Support Vector Regression.

Kaur and Sachin (Kaur & Ahuja, 2017) predict the electricity consumption using autoregressive moving average model —ARIMA—. Ma and Liu use the grey system —system with partial information known, it has two part: system with completely known information and system with completely unknown information (Julong, 1989)— theory to forecast the natural gas consumption of China (Ma & Liu, 2017). Simple and multiple linear regression is applied by Fumo and Rafe Bismas to predict energy consumption in family houses (Fumo & Rafe Biswas, 2015). Davlea and Teodorescu present a neuro fuzzy model to develop a middle-term load forecaster (Davlea & Teodorescu, 2016). Dhillon et al. employ Support Vector Regression —SVR— for short term load forecasting (Dhillon, et al., 2016). The ANN is the mostly used machine learning method and present great results, such as the Adaptive Network Based Inference System model to forecast building energy consumption in a cold region of Ekici (Bektas Ekici & Aksoy, 2011), Rodger's study which uses the fuzzy logic coupled with regression, nearest neighbour and artificial neural network to create a predictive model to make predicting demand for natural gas and energy cost savings in public buildings (Rodger, 2014) and many other works (Benedetti, Cesarotti, Introna, & Serranti, 2016; Egrioglu, Bas, Aladag, & Yolcu, 2016; Kanarachos, Christopoulos, Chroneos, & Fitzpatrick, 2017; Pino-Mejías, Pérez-Fargallo, Rubio-Bellido, & Pulido-Arcas, 2017; Rodrigues, Cardeira, Calado, & Melício, 2017).

However, the main disadvantage of ANN is its slow convergence and easy local minimum stagnation. This leads to the idea of using a technique to avoid these problems, and those are the Genetic Algorithm —GA— which is a global search and an optimization method. GA is widely used for optimizing models in time series forecasting for building energy consumption (Bhandari & Gill, 2016; Zhang, Deb, Lee, Yang, & Shah, 2016).

This paper is a straight continuation of a previous work (Ruiz, et al., 2016), and it proposes a method for predicting energy consumption by using ANN and the GA to improve the accuracy of these models. The main objective of this paper is to provide a methodology to analyse historical energy consumption, and perform the daily prediction with such models. Furthermore, a comparison is made between ANNs and identifies if the energy consumption forecasting can improve with external information or it depends entirely on historical consumption. This research has used data of the faculties, centres and schoolrooms of the University of Granada —UGR. The energy management systems are relatively new in UGR, they have been implemented and introduced in the most recent years. However, sufficient data are available to carry out this work and analyse building's behaviours.

This paper deals with energy efficiency in public and distributed buildings analysing a new proposed model of ANN with previous forecasting methods for energy consumption applied.

The present paper is divided as follows: Section 2 presents the suggested methodology, the employed artificial neural networks and the description of the genetic algorithms. Section 3 describes the proposed system containing the data processing and noise treatment. Section 4 introduces experiments performed, the description of the real data used, parameters, results and discussion achieved. The paper ends with some practical implications and concluding comments.

2. Proposed system

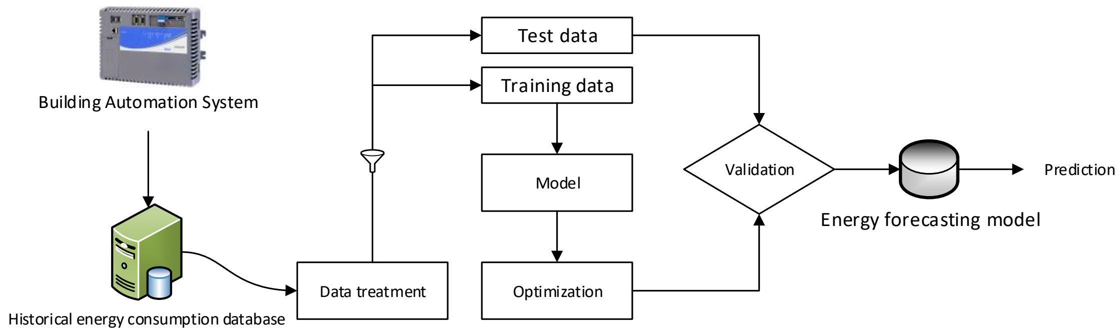


Figure 1: Proposed system flowchart.

The bulk of energy time-series modelling is represented in **Figure 1** where input is provided by building automation systems which stores all raw information in a database, and the output is forecasted consumption. The details of each component are outlined below.

A. Building automation system

Software characterized by a number of digital controllers, provides an asynchronous communication architecture for interacting with distributed building automation devices. It collects and presents building data, so that it can be interpreted.

B. Historical energy consumption database

The resources database is in charge of the information of the registered energy use. Besides, the database includes other extra knowledge from varying distributed sensors, such as power demand and temperature monitored. This constitutes a means for studying relations between energy and temperature, if exists.

C. Data treatment

Initially, database saves raw data which normally contains noise, incomplete, unreliable and missing data. It is important at this stage to transform the data to convert them to a suitable form. In this paper the time granularity energy consumption used is daily according previous works (Ruiz, et al., 2016).

Tangible and solid sensors are the link between the real world and digital world. And these devices sometimes present failures due to broken device, transmission errors or any other issues caused by the impairment. To solve this question two method have been applied: a) Energy time series consumption is filtered with a moving average filter and a sliding windows technique to eliminate breaks or other irregular patterns in the data (Smith, 1997), according to the next equation (1), b) Linear interpolation based on the immediate neighbours values at grid points to fill missing values.

$$y(n) = \frac{1}{windowsSize} \cdot \big(x(n) + x(n-1) + \cdots + x(n - (windosSize - 1))\big) \quad (1)$$

Finally, the data are normalized between $[0, 1]$ to standardize variables into same range using equation (2). And guaranteeing that there are no attributes which are more important than others, and also eases a stable convergence of network weights and biases.

$$y_{normalized} = \frac{y - y_{min}}{y_{max} - y_{min}} \quad (2)$$

D. Train and test

The dataset is split into two sets, the training set is 70% and the testing set is the remaining 30%. Both are selected randomly from the total energy consumption available.

E. Model and optimization

To predict the energy consumption at a determined time, we use three kind of ANN: NAR, NARX and ENN. These models are trained using LM algorithm and hyperbolic tangent sigmoid transfer function for the hidden layers (Vogl, Mangis, Rigler, Zink, & Alkon, 1988). The learning stage is an iterative process. At this point, GA is performed to optimize ANN, these were all depicted in section 3.3 (Genetic Algorithm). All parameters are explained in the experiments section.

F. Validation and model

The results are validated through comparison test data forecasting with the 30% of the data isolated before training. If the model's response is similar to test data, then it is assumed that the learning has been successful.

3. Methodology

The proposed methodology which follows has been developed in four stages. The first stage is data capture and preparation. Once the data have been compiled, the second stage is the ANN forecasting model. The next stage is genetic optimizing. And the final phase involves the analysis review and the use of the optimized ENN, achieving this through experimentation which appears in the next section.

The present study derives from previous researches done (Ruiz, et al., 2016) utilizing two well-known models for energy consumption forecasting: the non-linear autoregressive neural networks —NAR— for modelling the data process of one dimensional time series using past values (Ferlito, et al., 2015); and the non-linear autoregressive neural network with exogenous inputs —NARX including another external series which might provide relevant information (Cadenas, Rivera, Campos-Amezcua, & Cadenas, 2016). This paper proposed Elman Neural Network —ENN— and using genetic optimization to enhance preceding results. There are numerous studies in literature for solving time series prediction with neural networks (Benedetti, et al., 2016; Bhandari & Gill, 2016; Egrioglu, et al., 2016; Pino-Mejías, et al., 2017).

These models can be listed in order of complexity as: 1) NAR. 2) NARX. 3) ENN. The latter two networks offer the possibility of incorporating extra information to enhance the forecasting accuracy. In the real world, not all buildings are able to stock large quantities of data or register, sometimes only the energy meter is saving. On other occasions, management systems handle more information such as temperature. The suggested models are appropriate for both strategies, and the aim of this work is to determine which model is the better choice for this problem.

Subsequent sections describe employed models and analyses its pros and cons to solve the problem of energy consumption forecasting.

3.1. NAR and NARX neural network

Time series are a sequence of data, observations or numerical values usually recorded at uniform time intervals (Brockwell & Davis, 2013). Typically measured every second, minute, hour, day, week or even each year; although another time interval is valid: every 30 seconds, 12 hours, etc. A time series associated to the variable Z over the time set T is denoted by:

$$Z = \{Z_t : t \in T\} \quad (3)$$

Where Z_t is the value of Z at time t. For example, the energy consumption of a building defines a time series, indicating in each instant t the consumption spent by the building.

In many instances, the data depend not only on the total amount spent but also on other possible influence factors. In order to tackle this issue, the NAR and NARX neural networks have proved to be a very helpful tool in time series environments (Cadenas, et al., 2016; Wang, et al., 2016). These models are a kind recurrent system which are able to learn by itself, improving the approximation of the ANN by reducing the output error. NAR uses past values for the actual time series to predict next values as determined by the following equation (Ibrahim, Jemei, Wimmer, & Hissel, 2016):

$$y(t) = h\big(y(t-1), y(t-2), \ldots, y(t-p)\big) + \epsilon(t) \quad (4)$$

Where h is a nonlinear function which depends on p past values of the output y as shown in **Figure 2** excluding the exogenous input x —in grey colour— and ϵ represents random error sequence and independent distributed (Wang, et al., 2016). In much the same way as NAR, NARX equation is defined as:

$$y(t) = h\left(x(t-1), x(t-2), \ldots, x(t-p_x), y(t-1), y(t-2), \ldots, y(t-p_y)\right) + \epsilon(t) \quad (5)$$

Which includes external input x of the neural network. It is zero in case of NAR model.

The basic structure of recurrent neural networks is presented in **Figure 2**. There is an input layer with two time series. A hidden layer with p delays. And an output layer with an activation function. The Levenberg-Marquardt —LM— (Ampazis & Perantonis, 2000) method is used for optimizing the learning rate based on this gradient, which combines the local convergence properties of Gauss-Newton method near a minimum with the consistent error decrease provided by gradient descendent far away from a solution; this error is a mean squared error based on a learning sample. The LM procedure computes the Jacobian matrix of the error function which takes great use of memory. Nevertheless, this computational cost is well worth because it increases the rate of convergence of the algorithm (Hagan & Menhaj, 1994).

The NAR is used with one input —energy consumption at the previous time $\{y(t-1), \ldots, y(t-p)\}$— and one output —the predicted value $y(t)$—. Similarly, NARX network models the same energy consumption $\{y(t-1), \ldots, y(t-p)\}$ with the external input —temperature— given the p past values. The parameters of the neurons and delays are set in accordance with the best results achieved in antecedent studies (Ruiz, et al., 2016). In the next sections, all these parameters are specified.

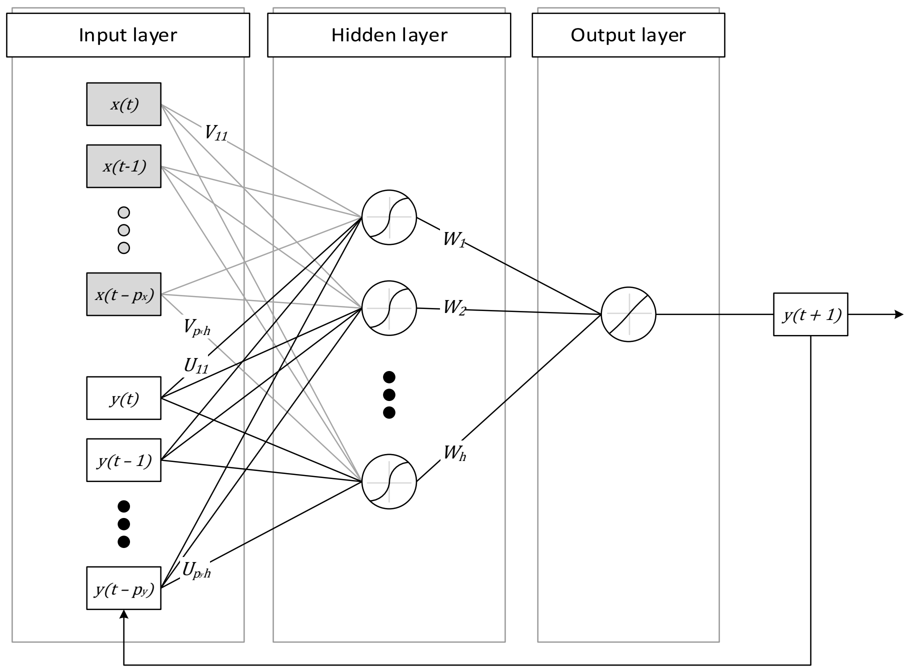


Figure 2: Representation of the structure of non-linear autoregressive neural network —NAR (without grey part)— with exogenous input —NARX (with the grey part)—.

V_{ij} represents the weight between exogenous input i and the hidden neuron j, W_i is the weight between hidden neuron i and the output neuron and U_{ij} is the weight for the connection between input i and the hidden neuron j. p_x and p_y are the past values introduced of the exogenous and input series respectively. h is the number of neurons in the hidden layer.

3.2. Elman Neural Network

Due to the present problem involves working in historical data, *memory* is an essential feature to process temporal information. The Elman Neural Network —

ENN— (Elman, 1990) introduces this concept of *memory*. Often, the past contains rich information and this is why it is stored in memory neurons. In ENN positive feedback is used to build this memory structure, and it is illustrated in **Figure 3**. These specialized units are called *context* units or *state* units store preceding outputs of hidden layer by using a positive feedback mechanism. For each unit in the hidden layer an extra context unit is fully connected with all the hidden neurons in a forward mode. State neurons are only connected to internal nodes of the network, and not with the outside world.

The following equations defines the ENN with n inputs, h hidden neurons, and o outputs:

$$S_j(t) = f\left(\sum_{r=1}^{h}\sum_{j=1}^{h} U_{jr}S_r(t-1) + \sum_{i=1}^{n}\sum_{j=1}^{h} V_{ji}X_i(t)\right) \quad \textbf{(6)}$$

$$O_k(t) = g\left(\sum_{j=1}^{h}\sum_{k=1}^{o} W_{kj}S_j(t)\right) \quad \textbf{(7)}$$

Here, $S_j(t)$ is the output of hidden neuron $j \in [1,h]$, $X_i(t)$ is the input data to neuron $i \in [1,n]$, $O_k(t)$ is the output $k \in [1,o]$ at time t. U,V,W are matrix with the network's weights. Thus, V_{ji} is the weight for the connection between input neuron i and the hidden neuron j. U_{jr} is the weight between the recurrent connection r and the hidden neuron j. And W_{kj} is the weight between hidden neuron j and output neuron k. f and g are activation functions.

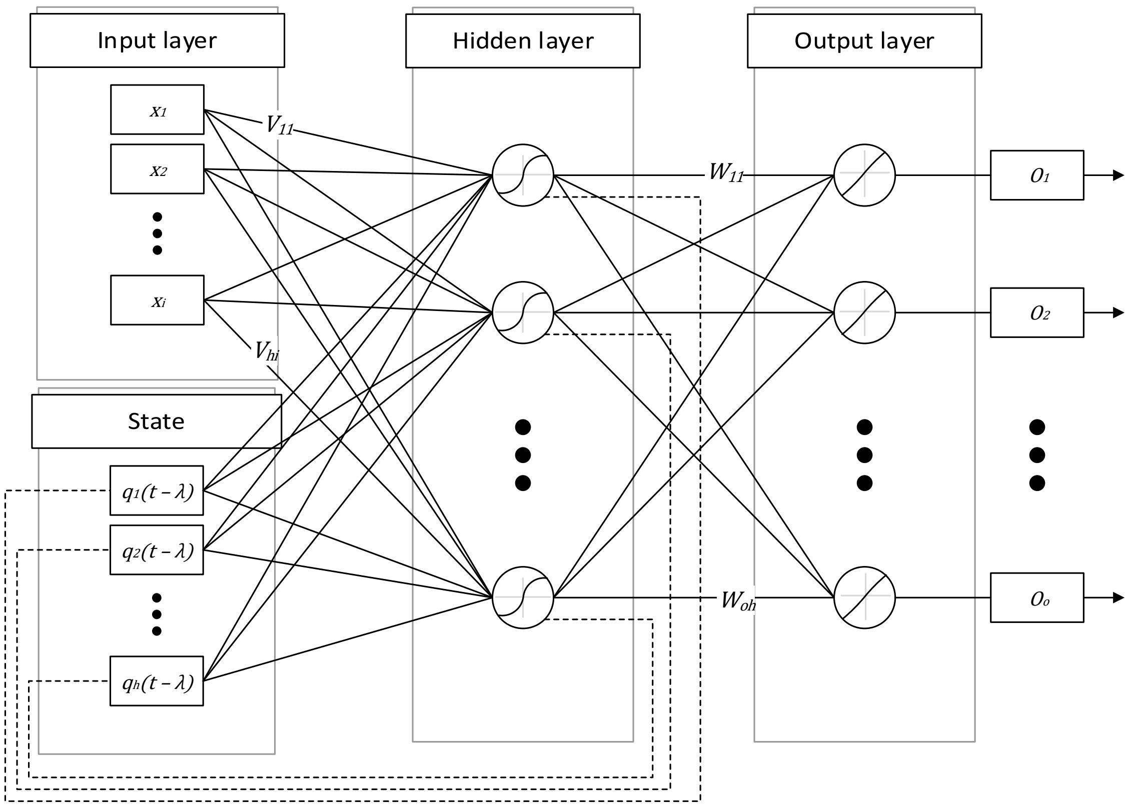


Figure 3: Architecture of the Elman Neural Network.

x_i is the input i, and the previous state in time $t-1$ of the hidden neuron j is represented by $q_j(t-\lambda)$ where $\lambda \geq 1$ is the number of previous hidden states stored.

The main dissimilarity between ENN and NAR(X) models is the recurrent link that appears from hidden layer to *Context or State Units*. In ENN, the *Context Units* store hidden neuron values at previous time step. However, there is another architecture whose *state* stores output neuron values at different past time: this is the Jordan networks (Jordan, 1997). The use of ENN in this study is supported by multiple works (Bao, Lin, Gong, & Shao, 2016; M. Cuéllar, Delgado, & Pegalajar, 2005; M. P. Cuéllar, Delgado, & Pegalajar, 2007; Miguel Delgado, M Carmen Pegalajar, & Manuel Pegalajar Cuéllar, 2006; M. Delgado, M. C. Pegalajar, & M. P. Cuéllar, 2006; Qin, Wang, Wu, & Zhao, 2016) adhering excellent results.

3.3. Genetic Algorithm

Genetic Algorithms —GA— have shown outstanding degrees of success in task related to neural network training (M. Cuéllar, et al., 2005; M. P. Cuéllar, et al., 2007; Miguel Delgado, et al., 2006; M. Delgado, et al., 2006). And for this reason, GA are used in this study in order to improve the accuracy of ANN prediction, because an improvement of a few percentages in the forecasting accuracy would bring benefits worth large amounts of money (Sadat Hosseini & Gandomi, 2012).

In essence, GAs simulate the mechanics of biological evolution. Following the philosophy of the famous naturalist Charles Darwin, the GA is based on natural selection or best adapted survival. In nature, individuals must adapt to their environment through a process name evolution. This evolving keeps positive aspects of an individual over time, and features that undermine the chromosome was ruled out. *The genetic algorithm is a highly parallel mathematical algorithm that transforms a set —population—of individual mathematical objects, each with an associated fitness value, into a new population* —i.e., the next generation— *using operations patterned after the Darwinian principle of reproduction and survival of the fittest and after naturally occurring genetic operations* —notably sexual recombination—. (Koza, 1992)

This study suggests the adaptation of the binary CHC —Cross generational elitist selection, Heterogeneous recombination, and Cataclysmic mutation— algorithm for a real-coded problem (Blanco, Delgado, & Pegalajar, 2001; Cordón, Damas, & Santamaría, 2006). The flowchart of the GA is shown in Figure 4

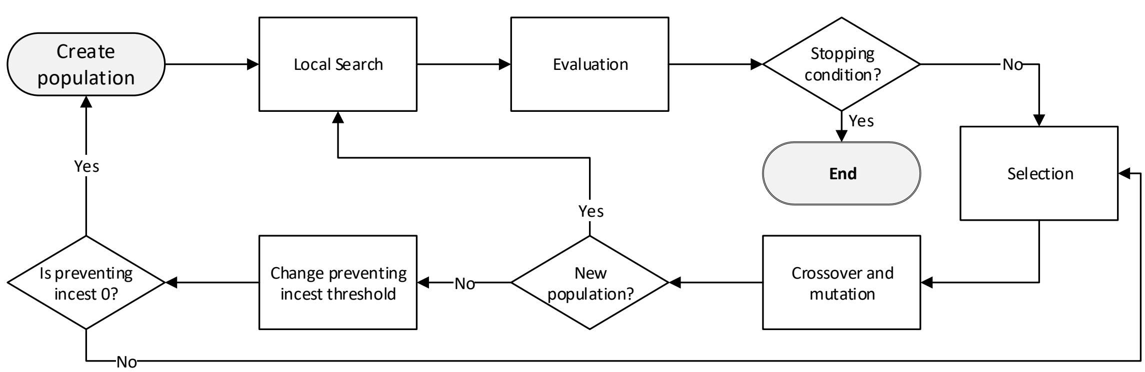
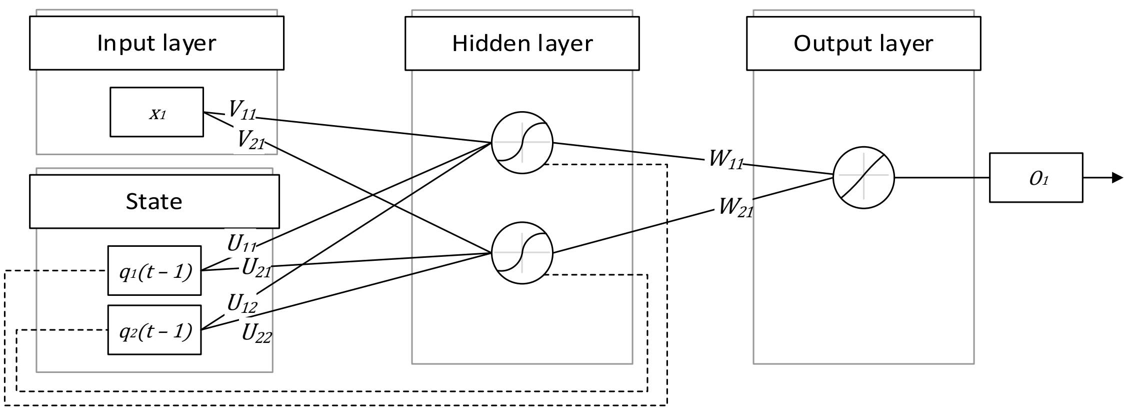


Figure 4: General Flowchart of the memetic algorithm —MA— based on binary-coded CHC schema.

This algorithm strikes a balance between diversity and convergence thanks to an elitist selection of individuals, invest prevention and initialization procedure of the population (Eshelman, 1991). For a better understanding of the algorithm components, before, there is a need to describe in detail who is an individual in this population, and then define fitness function, selection operation and local search. The hybridization of genetic algorithm and local search leads to a new kind of evolutionary algorithm commonly known as memetic algorithm.

An individual represents the structure of a NN. It would be made up of all weight which compose the network. Following the structure presented in **Figure 4**, an ENN with one input, one output and two hidden neurons storing its previous state $t-1$, **Figure 5** illustrates an example of NN encoding.



Representation

V_{11}	V_{21}	U_{11}	U_{21}	U_{12}	U_{22}	W_{11}	W_{21}

Figure 5: Encoding Elman Neural Network architecture for the Genetic Algorithm with one input, two hidden neurons, one output and one delay in the memory state.

The algorithm starts creating a first set of individuals in a totally random way. Afterward, it performs a local search to improve individual characteristics based on LM method. Thereupon, all individual in the population are evaluated by the fitness function, assigning everyone a Mean Square Error —MSE— according the network's goodness of fit. If $\hat{y}$ is a vector of m predictions, and y is the vector of observed values corresponding to the inputs, then the MSE can be estimated by equation (8):

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2 \tag{8}$$

If the stopping criteria is false, individuals who are going to cross are selected, usually known as parents. The incest prevention enables CHC to delay premature convergence: distance between the original parents must be exceed a certain limit, this limit is called as incest threshold. The Hamming distance is used in the original CHC. In this case, it makes no sense for real-coded problems, because weights of the NN are real values. Euclidean distance has been adopted instead, due to its widespread use and remarkable resolution (Han, et al., 2015).

When the crossover operator was able to build a new population, then a new local search is performed for these fresh individuals. The recombination of the ANN weights is illustrated in Figure 6. And the process repeats itself.

Otherwise, incest threshold is decreased, in favour of further crossing. At that time, parents are selected once again. A re-initialization is performed when the incest limit reaches zero, because the algorithm has fallen into local minimum. To gain an understanding of CHC algorithm, see the Eshelman's work (Eshelman, 1991).

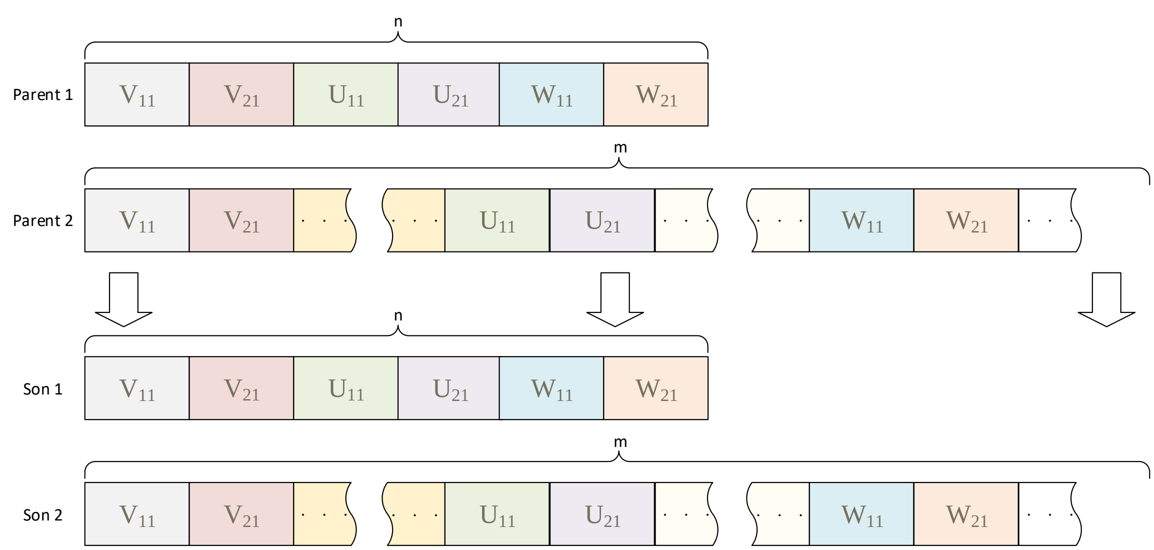


Figure 6: Example of the crossover procedure for two networks with different sizes.

The crossover procedure combines two parents and creates two children. If both parents contain that gene, the genetic recombination of this gene is carried out based on BLX-α operator, which combines two parents p^1 and p^2 to generate offspring s by sampling a new value in the range $[min_i - I \cdot \alpha, max_i + I \cdot \alpha]$ at each gene i. Where min_i and max_i are smaller and larger parent values at position i. I is $max_i - min_i$. And α value has been set to 0,2 (Picek, Jakobovic, & Golub, 2013). The remaining connections of the largest son are directly inherited from the biggest parent.

4. Experiments

4.1. Dataset

This work uses collected data from energy consumption of several University of Granada's building (UGR, Granada, Andalucía, Spain), such as classrooms, lecture rooms, laboratories and research centres. UGR owns a smart management system in its distributed facilities. It collects and monitored information of the building in real time from diverse sensors, for the purpose of being analysed and better understanding buildings behaviours.

Because of management systems implementation is relatively new, not all building dispose the same sensors. Therefore, not all buildings record the same information. Most of them collect energy consumption and climatic data. The UGR is made up of five campuses: *Centro, Cartuja, Fuentenueva, Aynadamar and Ciencias de la Salud*. These campuses are spread around different places of the city. Autonomous cities of Melilla and Ceuta also contain UGR's centres, placed in separate campus. As a whole, the UGR is assembled of 22 faculties, 5 schools, 8 training centres and 5 culture, sport and service centres. Due to Data Protection Act this study cannot reveal details over the facilities, and hence, buildings consumptions are labelled with a number. Eight edifices have been chosen, two representative building from the campuses.

Figure 7 presents the temperature and consumption pattern during 1 year. Both series have been normalized because they have distinct units. Energy consumption is recorded in kW, and temperature is in degrees Celsius. This picture gives evidence of a peculiar behaviours, for example, highest consumption is performed at lowest temperatures, securely caused by air conditioners and heating. For this reason, it is interesting to include temperature in our predictive models. But on the other hand, incorporating too many variables, makes the model much more complicated and inserts uncertainty into the system because it would depend on known variables. The temperature is a sustainable dependence thanks to its smooth and regular behaviour.

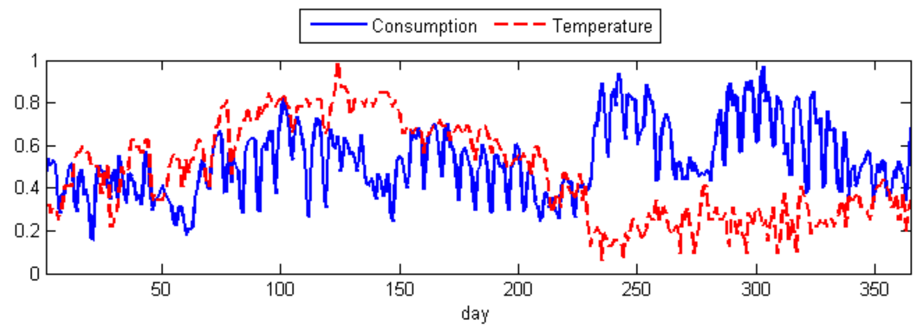


Figure 7: Example of normalized consumption and temperature recorded during one year.

We study models without temperature too, in order to have a possible alternative and how good it was compared to previous predictors and the new proposed.

4.2. Parameters

This section summarizes the parameters used for each model, and may safely be skipped by readers who are easily bored.

The number of neurons for NAR and NARX models have been gathered from the best results obtained in preceding study (Ruiz, et al., 2016). In summary: 14 neurons for the buildings 1 and 4; 12 neurons for the buildings 2, 3 and 7; 9 neurons for buildings 5, 6 and 8. These parameters have been taken to allow the comparison between outcomes with the same structure.

In the ENN case, it is not necessary establish the same number of neurons as NAR(X) because it does not follow the same architecture as NAR and NARX networks. The best number of neurons and delay have been set by testing experimentally different sizes and using cross-validation: 10 hidden neurons have been set and it has a memory of 5 past values.

The training function used in all cases is the LM backpropagation optimization. Minimum gradient is 10^{-7}, the training gain, μ, is a parameter which measures the adapting and learning rate of the model, its range is $[10^{-3}, 10^{10}]$ and the μ decrease and increase ratio are 0.1 and 10 respectively. The nets stop training at many different epochs and use validation sets to select the best model. There is an important aspect which should be considered with the ANN: a low number of hidden neurons was used to avoid overfitting to obtain a model as simple as possible, but too few neurons may lead to a negligent and ineffective learning. Furthermore, cross validation is performed in order to deal with this issue.

The population size has been set to 25 individuals. The error of each individual is illustrated in **Figure 8**, it shows the corresponding error curves for two executions, and in both cases the difference between the best and the worst individual is more than double. Therefore, to increase the population size makes the GA works with worse models and to decelerate the optimization process. The interval value for a gene is [-10, 10], stop criteria is set in 100 generations, crossover and mutation probability are 90% and 10%, respectively.

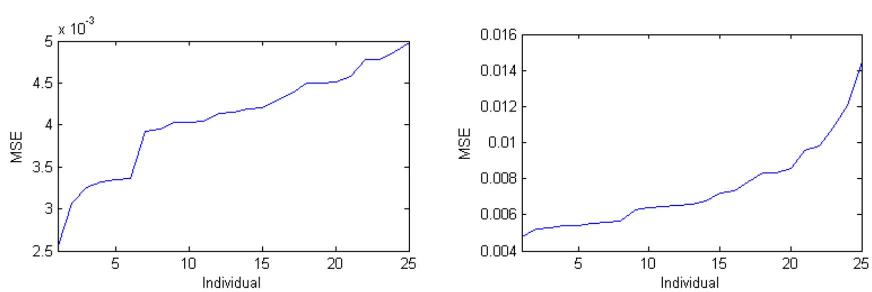


Figure 8: Two examples of Mean Square Error of each individual of a population once GA have finished.

4.3. Results and discussion

The proposed method was programmed in Matlab software run on Intel® Core™ i7-6700 CPU @ 3.40GHz. In order to confirm the robustness of the achieved results, for each experimental the simulation was run 5 times. **Table 1** gathers the outcomes achieved by using GA and previous ones. The first column lists the buildings by id, remember that each building has been selected as specified in section **4.1 Dataset**. Second column relates to the MSE obtained with NAR models optimized by GA+LM. The next column illustrates the errors of NARX models —with temperature— and improved by GA+LM. In the last two columns, the table indicate the results of the new suggested networks, without and with temperature respectively.

In all buildings, previous results are worse than new ones. Indeed, in each of these cases, they are well above the worst of the proposed model, ENN. This table illustrates the excellent performance of the GA+LM. This provides an optimization for both models, non-autoregressive and Elman. Thus, after considering the results, it may be concluded that the adapted CHC algorithm is a good method for neural network optimization.

Besides, in order to facilitate comprehension and clarity, **Figure 9**, illustrates a comparison between all models. The graphic views provide a quick overview of the improvements. The proposed method achieves a significant enhancement whichever model. In all but one case, the models are ranked as follows: 1) the ENN with temperature, 2) ENN without temperature, 3) NAR network and 4) NARX network with temperature. On the one hand, considering NAR and NARX networks only, the results are satisfactory, optimized networks are much better than previous one. These models correspond to an average of 35% improvement, with a 16% in the worst case and a 52% in the best case, acquired with NAR model and building 6. On the other and, the new proposed network, ENN, is able to obtain an even better fit. With an average of 61% improvement, 51% in the worst case and up to 82% in the best one. The propounded method with ENN provides much better results for all probe sets and considerably higher average score, compared to all other models.

Table 1. Results of 8 buildings including results of previous works and the present method. Best results in bold.

Building	Previous results	NAR	NARX temperature	Elman	Elman temperature
1	0.018200	0.012060	0.011442	0.007585	**0.005288**
2	0.014200	0.007681	0.008305	0.005203	**0.004698**
3	0.013000	0.009608	0.010558	0.005706	**0.005439**
4	0.017000	0.009736	0.010079	0.006966	**0.006604**
5	0.006700	0.005173	0.005598	0.003266	**0.002627**
6	0.006200	0.002953	0.003324	0.001376	**0.001059**
7	0.013500	0.007535	0.007998	0.006258	**0.005660**
8	0.009300	0.006830	0.007162	0.004320	**0.003929**

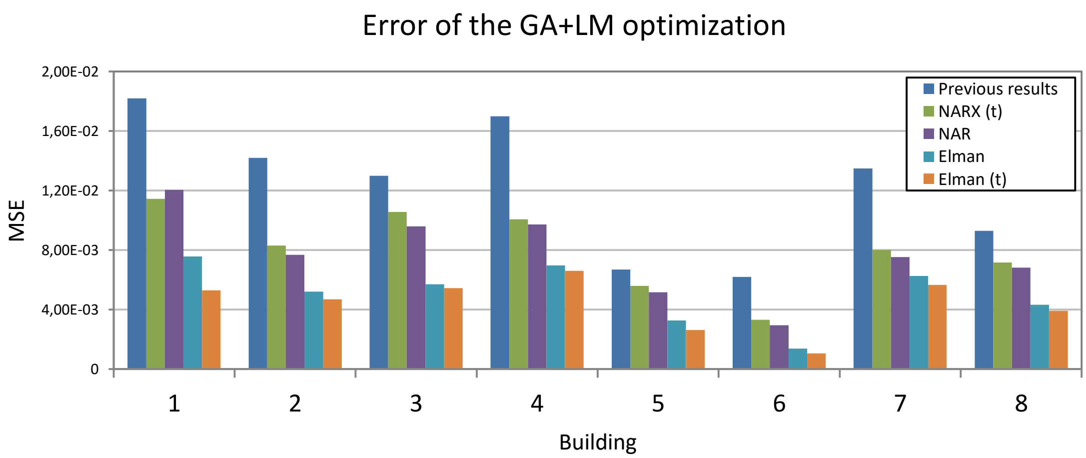


Figure 9: Comparative evaluation between preceding outcomes and new optimized models by using MA.

Figure 10 depicts an example of the forecasted consumption with all model used. The blue series is the original data normalized, so it is the desired value which models should be adjusted. The grey sequence is the response of the best models reached in previous work. Yellow and orange series concern to optimized NAR and NARX respectively thanks to GA+LM. And the green crosses are the Elman forecasting. This figure does not show two different series with and without temperature ENN because the variations are imperceptible and would make it difficult for the graph display. As shown in that picture all forecasted models follows the trend of the data quite well. However, the most faithful of the real data is Elman series which accurately predicts future values. And the worst is the non-optimized model with GA+LM of the preceding study where there are a few values which are more separated from the original data, for example in the days 25, 35, 60 and 77. The NAR and NARX models offer similar behaviour. Indeed, **Figure 9** details how optimized NAR models achieve better MSE than NARX with temperature. This is not the case of ENN where the predictor using temperature enhance a bit closer.

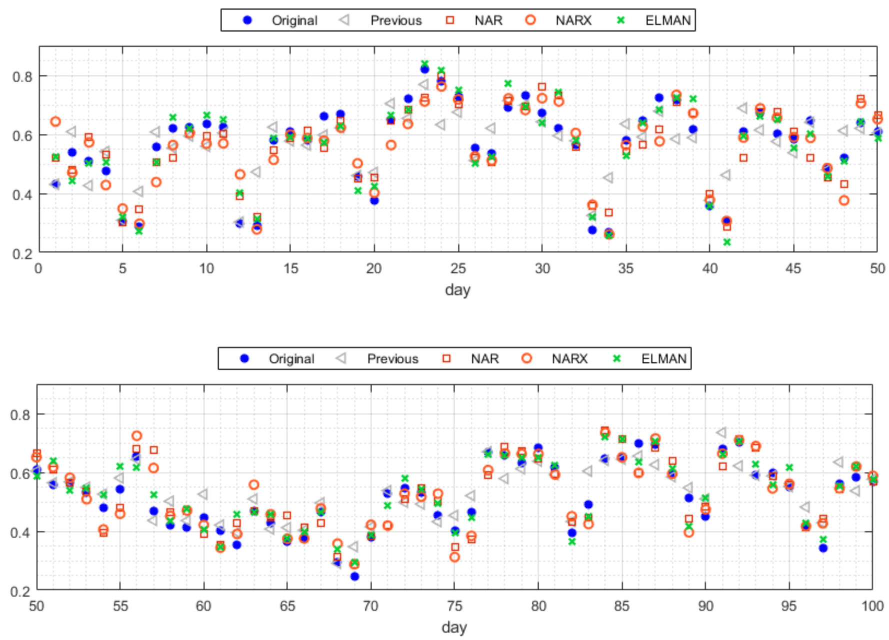


Figure 10: Forecasted and original values for energy time series consumption for 100 days. The first 50 days in the above graphic and the 50 last days in the chart below.

5. Conclusion

In this paper, we have introduced a new methodology to energy consumption forecasting and achieve optimum models. The GA has proven to be a useful and a key factor for optimizing ANN, and it helps to significantly enhance in NAR and NARX models too, used in previous works. Besides, the ENN have been very effective and it has demonstrated to be the best network in all test performed, obtaining an average improvement of 61%.

The main advantage in using NAR and NARX networks is their simplicity. However, this advantage limits its accuracy. Likewise, the major problem of the ENN lies in its complexity, because increasing components of a neural network —that is to say: including memory layer— implies increasing the number of connections, and the results is a much more complex model. This problem is known as the «curse of dimensionality», a well-known problem in statistical learning, this expression is used in phenomena that appear with high-dimensional data, and that have most often unfortunate consequences on the behaviour and performances of learning algorithms (Korn, Pagel, & Faloutsos, 2001). In this respect, it is better to seek a compromise between predictor model complexity and an acceptable level of error in the results. Although, a complex model maybe is not an important variable to consider if it can save a lot of money.

In our approach, we assume that each building has a device integrated to capture energy consumption and store it. Sometimes, external data, such as temperature, is not available, because it depends on the sensors implemented and the building's budget. Thus, this study works this two approaches which lend support to both cases, achieving a good degree of MSE: 0.005085 and 0.004413 for models without temperature and including temperature respectively. Given the importance of relationship between current and past data, further studies will focus upon developing a system to find time relations in the building consumption by using clustering methods.

Acknowledgments

This work has been developed with the support of the Department of Computer Science and Artificial Intelligence of the University of Granada, TIC111, and the project TIN201564776-C3-1-R.

Abbreviations

ANN	Artificial Neural Network.
ARIMA	Auto-Regressive Integrated Moving Average.
ENN	Elman Neural Network.
GA	Genetic Algorithm.
LM	Levenberg-Marquardt.
MA	Memetic Algorithm.

MLP	Multilayer perceptron model.
MSE	Mean Squared Error.
NAR	Nonlinear autoregressive model.
NARX	Nonlinear autoregressive model with exogenous inputs.
NN	Neural Network.
SVR	Support Vector Regression.
UGR	University of Granada.

Bibliography

Ampazis, N., & Perantonis, S. J. (2000). Levenberg-Marquardt algorithm with adaptive momentum for the efficient training of feedforward networks. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium* (Vol. 1, pp. 126-131 vol.121). [http://dx.doi.org/10.1109/IJCNN.2000.857825]

Andrade, J. R., & Bessa, R. J. (2017). Improving Renewable Energy Forecasting with a Grid of Numerical Weather Predictions. *IEEE Transactions on Sustainable Energy, PP*, 1-1. [http://dx.doi.org/10.1109/TSTE.2017.2694340]

Balaji, B., Xu, J., Nwokafor, A., Gupta, R., & Agarwal, Y. (2013). Sentinel: occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* (pp. 1-14). Roma, Italy: ACM. [http://dx.doi.org/10.1145/2517351.2517370]

Balón-Canedo, V., Remeseiro, B., Sechidis, K., Martinez-Rego, D., & Alonso-Betanzos, A. (2017). Algorithmic challenges in Big Data analytics. [http://www.cs.man.ac.uk/~sechidik/publications/2017/esann2017.pdf]

Bao, G., Lin, Q., Gong, D., & Shao, H. (2016). Hybrid Short-term Load Forecasting Using Principal Component Analysis and MEA-Elman Network. In D.-S. Huang, K. Han & A. Hussain (Eds.), *Intelligent Computing Methodologies: 12th International Conference, ICIC 2016, Lanzhou, China, August 2-5, 2016, Proceedings, Part III* (pp. 671-683). Cham: Springer International Publishing. [http://dx.doi.org/10.1007/978-3-319-42297-8_62]

Basbagill, J., Flager, F., Lepech, M., & Fischer, M. (2013). Application of life-cycle assessment to early stage building design for reduced embodied environmental impacts. *Building and Environment, 60*, 81-92. [http://dx.doi.org/10.1016/j.buildenv.2012.11.009]

Bektas Ekici, B., & Aksoy, U. T. (2011). Prediction of building energy needs in early stage of design by using ANFIS. *Expert Systems with Applications, 38*, 5352-5358. [http://dx.doi.org/10.1016/j.eswa.2010.10.021]

Benedetti, M., Cesarotti, V., Introna, V., & Serranti, J. (2016). Energy consumption control automation using Artificial Neural Networks and adaptive algorithms: Proposal of a new methodology and case study. *Applied Energy, 165*, 60-71. [http://dx.doi.org/10.1016/j.apenergy.2015.12.066]

Bhandari, R., & Gill, J. (2016). An Artificial Intelligence ATM forecasting system for Hybrid Neural Networks. *International Journal of Computer Applications, 133*, 13-16. [http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.734.7389&rep=rep1&type=pdf]

Blanco, A., Delgado, M., & Pegalajar, M. C. (2001). A real-coded genetic algorithm for training recurrent neural networks. *Neural Networks, 14*, 93-105. [http://dx.doi.org/10.1016/S0893-6080(00)00081-2]

Brockwell, P. J., & Davis, R. A. (2013). *Time series: theory and methods*: Springer Science & Business Media. [http://dx.doi.org/10.1007/978-1-4419-0320-4]

Bunn, D., & Farmer, E. D. (1985). Comparative models for electrical load forecasting.

Cadenas, E., Rivera, W., Campos-Amezcua, R., & Cadenas, R. (2016). Wind speed forecasting using the NARX model, case: La Mata, Oaxaca, México. *Neural Computing and Applications, 27*, 2417-2428. [http://dx.doi.org/10.1007/s00521-015-2012-y]

Cordón, O., Damas, S., & Santamaría, J. (2006). Feature-based image registration by means of the CHC evolutionary algorithm. *Image and Vision Computing, 24*, 525-533. [http://dx.doi.org/10.1016/j.imavis.2006.02.002]

Cuéllar, M., Delgado, M., & Pegalajar, M. (2005). Multiobjective evolutionary optimization for Elman recurrent neural networks, applied to time series prediction. *Fuzzy Economic Review, 10*, 17. [http://search.proquest.com/docview/228956321]

Cuéllar, M. P., Delgado, M., & Pegalajar, M. C. (2007). Topology Optimization and Training of Recurrent Neural Networks with Pareto-Based Multi-objective Algorithms: A Experimental Study. In F. Sandoval, A. Prieto, J. Cabestany & M. Graña (Eds.), *Computational and Ambient Intelligence: 9th International Work-Conference on Artificial Neural Networks, IWANN 2007, San Sebastián, Spain, June 20-22, 2007. Proceedings* (pp. 359-366). Berlin, Heidelberg: Springer Berlin Heidelberg. [http://dx.doi.org/10.1007/978-3-540-73007-1_44]

Chou, J.-S., & Ngo, N.-T. (2016). Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns. *Applied Energy, 177*, 751-770. [http://dx.doi.org/10.1016/j.apenergy.2016.05.074]

Davlea, L., & Teodorescu, B. (2016). A neuro-fuzzy algorithm for middle-term load forecasting. In *2016 International Conference and Exposition on Electrical and Power Engineering (EPE)* (pp. 005-009). [http://dx.doi.org/10.1109/ICEPE.2016.7781292]

Delgado, M., Pegalajar, M. C., & Cuéllar, M. P. (2006). Evolutionary training for Dynamical Recurrent Neural Networks: an application in finantial time series prediction. *Mathware and Soft Computing, 13*, 89-110. [http://eudml.org/doc/41874]

Delgado, M., Pegalajar, M. C., & Cuéllar, M. P. (2006). Memetic evolutionary training for recurrent neural networks: An application to time-series prediction. *Expert Systems, 23*, 99-114. [http://dx.doi.org/10.1111/j.1468-0394.2006.00327.x]

Dhillon, J., Rahman, S. A., Ahmad, S. U., & Hossain, M. J. (2016). Peak electricity load forecasting using online support vector regression. In *2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (pp. 1-4). [http://dx.doi.org/10.1109/CCECE.2016.7726784]

Egrioglu, E., Bas, E., Aladag, C. H., & Yolcu, U. (2016). Probabilistic Fuzzy Time Series Method Based on Artificial Neural Network. *American Journal of Intelligent Systems, 6*, 42-47. [http://dx.doi.org/10.5923/j.ajis.20160602.02]

Ekwevugbe, T., Brown, N., Pakka, V., & Fan, D. (2013). Real-time building occupancy sensing using neural-network based sensor network. In *2013 7th IEEE International Conference on Digital Ecosystems and Technologies (DEST)* (pp. 114-119). [http://dx.doi.org/10.1109/DEST.2013.6611339]

Elman, J. L. (1990). Finding Structure in Time. *Cognitive Science, 14*, 179-211. [http://dx.doi.org/10.1207/s15516709cog1402_1]

Eshelman, L. J. (1991). The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination. In R. Gregory J.E (Ed.), *Foundations of Genetic Algorithms* (Vol. Volume 1, pp. 265-283): Elsevier. [http://dx.doi.org/10.1016/B978-0-08-050684-5.50020-3]

Fan, J., Han, F., & Liu, H. (2014). Challenges of Big Data Analysis. *National science review, 1*, 293-314. [http://dx.doi.org/10.1093/nsr/nwt032]

Ferlito, S., Atrigna, M., Graditi, G., Vito, S. D., Salvato, M., Buonanno, A., & Francia, G. D. (2015). Predictive models for building's energy consumption: An Artificial Neural Network (ANN) approach. In *2015 XVIII AISEM Annual Conference* (pp. 1-4). [http://dx.doi.org/10.1109/AISEM.2015.7066836]

Fumo, N., & Rafe Biswas, M. A. (2015). Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews, 47*, 332-343. [http://dx.doi.org/10.1016/j.rser.2015.03.035]

Hagan, M. T., & Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks, 5*, 989-993. [http://dx.doi.org/10.1109/72.329697]

Han, X., Liang, Y., Li, Z., Li, G., Wu, X., Wang, B., Zhao, G., & Wu, C. (2015). An Efficient Genetic Algorithm for Optimization Problems with Time-Consuming Fitness Evaluation. *International Journal of Computational Methods, 12*, 1350106. [http://dx.doi.org/10.1142/s0219876213501065]

Ibrahim, M., Jemei, S., Wimmer, G., & Hissel, D. (2016). Nonlinear autoregressive neural network in an energy management strategy for battery/ultra-capacitor hybrid electrical vehicles. *Electric Power Systems Research, 136*, 262-269. [http://dx.doi.org/10.1016/j.epsr.2016.03.005]

International Energy Agency. (2013). Transition to Sustainable Building. Strategies and Opportunities to 2050. In. [http://www.iea.org/publications/freepublications/publication/transition-to-sustainablebuildings.html]

Jalaei, F., & Jrade, A. (2014). An Automated BIM Model to Conceptually Design, Analyze, Simulate, and Assess Sustainable Building Projects. *Journal of Construction Engineering, 2014*, 21. [http://dx.doi.org/10.1155/2014/672896]

Jordan, M. I. (1997). Serial Order: A Parallel Distributed Processing Approach. *Advances in Psychology, 121*, 471-495. [http://dx.doi.org/http://dx.doi.org/10.1016/S0166-4115(97)80111-2]

Julong, D. (1989). Introduction to grey system theory. *The Journal of grey system, 1*, 1-24. [https://pdfs.semanticscholar.org/3013/67cf1d3f0aa5248a5b8a9ffa4813ef861620.pdf]

Kanarachos, S., Christopoulos, S.-R. G., Chroneos, A., & Fitzpatrick, M. E. (2017). Detecting anomalies in time series data via a deep learning algorithm combining wavelets, neural networks and Hilbert transform. *Expert Systems with Applications, 85*, 292-304. [http://dx.doi.org/10.1016/j.eswa.2017.04.028]

Kaur, H., & Ahuja, S. (2017). Time Series Analysis and Prediction of Electricity Consumption of Health Care Institution Using ARIMA Model. In K. Deep, J. C. Bansal, K. N. Das, A. K. Lal, H. Garg, A. K. Nagar & M. Pant (Eds.), *Proceedings of Sixth International Conference on Soft Computing for Problem Solving: SocProS 2016, Volume 2* (pp. 347-358). Singapore: Springer Singapore. [http://dx.doi.org/10.1007/978-981-10-3325-4_35]

Khosravani, H., Castilla, M., Berenguel, M., Ruano, A., & Ferreira, P. (2016). A Comparison of Energy Consumption Prediction Models Based on Neural Networks of a Bioclimatic Building. *Energies, 9*, 57. [http://www.mdpi.com/1996-1073/9/1/57]

Korn, F., Pagel, B. U., & Faloutsos, C. (2001). On the “dimensionality curse” and the “self-similarity blessing&rdquo. *IEEE Transactions on Knowledge and Data Engineering, 13*, 96-111. [http://dx.doi.org/10.1109/69.908983]

Koza, J. R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection.*

[http://www.cs.bham.ac.uk/~wbl/biblio/cache/bin/cache.php?koza:book,http___mitpress.mit.edu_books_genetic-programming,http://mitpress.mit.edu/books/genetic-programming]

Lord, S.-F., Noye, S., Ure, J., Tennant, M. G., & Fisk, D. J. (2016). Comparative review of building commissioning regulation: a quality perspective. *Building Research & Information, 44*, 630-643. [http://dx.doi.org/10.1080/09613218.2016.1181955]

Ma, X., & Liu, Z. (2017). Application of a novel time-delayed polynomial grey model to predict the natural gas consumption in China. *Journal of Computational and Applied Mathematics*. [http://dx.doi.org/10.1016/j.cam.2017.04.020]

Olubunmi, O. A., Xia, P. B., & Skitmore, M. (2016). Green building incentives: A review. *Renewable and Sustainable Energy Reviews, 59*, 1611-1621. [http://dx.doi.org/http://doi.org/10.1016/j.rser.2016.01.028]

Pallit, A., & Popovic, D. (2005). *Computational Intelligence in Time Series Forecasting*: Springer, Heidelberg. [http://dx.doi.org/10.1007/1-84628-184-9]

Picek, S., Jakobovic, D., & Golub, M. (2013). On the recombination operator in the real-coded genetic algorithms. In *2013 IEEE Congress on Evolutionary Computation* (pp. 3103-3110). [http://dx.doi.org/10.1109/CEC.2013.6557948]

Pino-Mejías, R., Pérez-Fargallo, A., Rubio-Bellido, C., & Pulido-Arcas, J. A. (2017). Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO2 emissions. *Energy, 118*, 24-36. [http://dx.doi.org/10.1016/j.energy.2016.12.022]

Qin, S., Wang, J., Wu, J., & Zhao, G. (2016). A hybrid model based on smooth transition periodic autoregressive and Elman artificial neural network for wind speed forecasting of the Hebei region in China. *International Journal of Green Energy, 13*, 595-607. [http://dx.doi.org/10.1080/15435075.2014.961462]

Rodger, J. A. (2014). A fuzzy nearest neighbor neural network statistical model for predicting demand for natural gas and energy cost savings in public buildings. *Expert Systems with Applications, 41*, 1813-1829. [http://dx.doi.org/10.1016/j.eswa.2013.08.080]

Rodrigues, F., Cardeira, C., Calado, J. M. F., & Melício, R. (2017). Family Houses Energy Consumption Forecast Tools for Smart Grid Management. In P. Garrido, F. Soares & A. P. Moreira (Eds.), *CONTROLO 2016: Proceedings of the 12th Portuguese Conference on Automatic Control* (pp. 691-699). Cham: Springer International Publishing. [http://dx.doi.org/10.1007/978-3-319-43671-5_58]

Ruiz, L. G. B., Cuellar, M. P., Delgado, M., & Pegalajar, M. C. (2016). An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings. *Energies, 9*, 21. [http://dx.doi.org/10.3390/en9090684]

Sadat Hosseini, S. S., & Gandomi, A. H. (2012). Short-term load forecasting of power systems by gene expression programming. *Neural Computing and Applications, 21*, 377-389. [http://dx.doi.org/10.1007/s00521-010-0444-y]

Smith, S. W. (1997). Moving average filters. *The scientist and engineer's guide to digital signal processing*, 277-284. [http://www.dspguide.com/ch15.htm]

Vogl, T. P., Mangis, J. K., Rigler, A. K., Zink, W. T., & Alkon, D. L. (1988). Accelerating the convergence of the back-propagation method. *Biological Cybernetics, 59*, 257-263. [http://dx.doi.org/10.1007/bf00332914]

Wang, X., Song, H., Bai, S., Zeng, T., Zhao, S., Wu, W., Li, W., & Liu, Y. (2016). Study of the monitored method of atomic clock data exception based on the model of dynamic neural network time series-NAR. In *2016 IEEE International Frequency Control Symposium (IFCS)* (pp. 1-4). [http://dx.doi.org/10.1109/FCS.2016.7546764]

Zhang, F., Deb, C., Lee, S. E., Yang, J., & Shah, K. W. (2016). Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. *Energy and Buildings, 126*, 94-103. [http://dx.doi.org/10.1016/j.enbuild.2016.05.028]

6.3 Parallel memetic algorithm for training recurrent neural networks for the energy efficiency problem

Referencia

Ruiz, L. G. B., Capel, M. I., & Pegalajar, M. C. (2019). Parallel memetic algorithm for training recurrent neural networks for the energy efficiency problem. Applied Soft Computing, 76, 356-368.

Estado

Publicado.

Factor de impacto

Factor de impacto 4.873.

Categoría

Posición 20/133 en el área “Computer Science, artificial intelligence” y 11/105 en “Computer Science, interdisciplinary applications”.

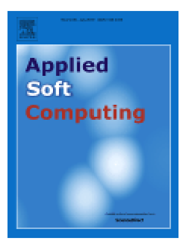
DOI

10.1016/j.asoc.2018.12.028

Revista ~ Editorial

Applied Soft Computing

Logo



Parallel Memetic Algorithm for Training Recurrent Neural Networks for the Energy Efficiency Problem

L.G.B. Ruiz [a,*], M.I. Capel [b], M.C. Pegalajar [a]

[a] Department of Computer Science and Artificial Intelligence, University of Granada.
[b] Department of Software Engineering, University of Granada.
bacaruiz@ugr.es* / manuelcapel@ugr.es / mcarmen@decsai.ugr.es

Abstract. In our state-of-the-art study, we improve neural network-based models for predicting energy consumption in buildings by parallelizing the CHC adaptive search algorithm. We compared the sequential implementation of the evolutionary algorithm with the new parallel version to obtain predictors and found that this new version of our software tool halved the execution time of the sequential version. New predictors based on various classes of neural networks have been developed and the obtained results support the validity of the proposed approaches with an average improvement of 75% of the average execution time in relation to previous sequential implementations.

Keywords: *energy efficiency; neural networks; time series prediction; evolutionary algorithms; manager-worker parallelization algorithms*

1. Introduction

Energy Efficiency (EE) is currently one of the greatest areas of interest for governments since the implementation of building energy-saving policies has a major impact on reducing pollutant emissions and has the greatest potential for delivering significant economic savings. It is also undoubtedly a matter of concern to the international community who must adhere to Directive 2012/27/EU of the European Parliament regarding energy management in terms of compliance with the overall objectives of improving energy efficiency, increasing the use of renewable energy sources, and reducing greenhouse gas emissions [1]. This paper presents a performant solution to the energy consumption prediction problem with time based on a simple parallelization of the CHC adaptive search algorithm [2].

The successful application of models to address the EE issue in buildings or in a distributed location environment typically requires the use of real-time applications [3, 4] if we need to obtain timely, dependable data to feed the mentioned models. Processing this type of information, however, requires the use of fast, accurate techniques so that decisions can be made quickly [5], and energy consumption prediction with time restrictions therefore represents a fundamental problem to be solved in order to achieve exhaustive energy management in buildings.

The problem of consumption forecasting has been explored for building energy management in recent years [6-8]. Nevertheless, these approaches are ad hoc solutions to EE problems and are based on classical techniques which require high computational costs if we want to achieve optimal solutions. Different soft computing paradigms such as artificial neural networks (ANN) [8-10] or evolutionary computation (EC) [6, 11, 12] can be used to find predictable energy management systems that reduce energy waste in buildings: ANN-based systems have proved to be successful for models for energy consumption prediction [9, 10, 13], energy inefficiency diagnosis and fault detection [14, 15]; EC-based ones support multi-objective applications that can be combined with Data Mining (DM). There are several EC applications worth mentioning and these include ones that deal with cost-optimal analysis [16], classification of new electricity customers [17], or selection of the most relevant features [18] and detection of outliers [14, 19, 20].

It is expected that the growing availability of data will soon require innovative tools to face the challenges posed by the variety, volume and velocity of data generation [21]. A sensorized environment generates high speed data streams, which involve the development of fast technology that is capable of processing large amounts of data from all the sensors of the equipment currently used for heating, ventilation, air-conditioning (HVAC) and lighting. We can deploy DM techniques that are strongly supported by parallel computing [22] to greatly reduce the data transfer overload while huge volumes of data are processed.

Furthermore, the adaptation of EE-prediction sequential algorithms, which are themselves highly parallelizable, requires the development of designs using powerful parallel techniques which are now available. The application of these methods allows us to obtain valuable benefits in computing time to solve more complex problems by means of parallel approaches [23]. Our proposal for solving such problems is therefore based on estimating the most efficient handling of energy possible by applying ANN and EC techniques and accelerating knowledge extraction and evaluation of data related to energy savings in real time through advanced parallelization techniques.

Although DM techniques require meaningful amounts of data in order to acquire relevant knowledge and reach useful conclusions, they are not generally suited to processing large amounts of data and responding within a reasonable period of time. Addressing these two issues represents a fundamental challenge, especially nowadays, when taking prompt decisions is essential to save costs in terms of energy consumption efficiency [5]. It is essential to optimize energy consumption prediction techniques for processing real-time data in scenarios where the dataflow is constant and permanent [24].

Generally speaking, the main goal of parallel techniques is to detect and exploit the available computational resources in order to make optimal use of them. DM techniques and evolutionary algorithms often present an iterative process which might be a significant opportunity for improvement in terms of time. Finding the optimal subset of code which supplies these requirements is an arduous task and may require a large number of modifications to be made to the sequential algorithm [25].

Our study proposes a modified implementation of the CHC algorithm [26] for optimizing the models used in energy consumption forecasting methods. This optimization algorithm has been widely used in recent studies. There are as many papers as task scheduling policies for providing services to numerous users in cloud environments for solving cloud computing problems [27]. These papers maximize resource utilization and minimize task processing time, or optimize the configuration of a new evolutionary fuzzy k-NN algorithm as in the proposal by Derrac *et al.* [28]. The CHC algorithm is used to establish the model parameters by self-optimization. Within this study domain, articles can be also found that follow the CHC scheme, such as the proposal shown in [29] which can forecast energy consumption from short-term to long-term time series using radial basis function neural networks.

We have used four well-known types of ANN: the non-linear autoregressive neural network (NAR) with exogenous inputs (NARX), the Elman neural network (ENN) and the Long Sort-Term Memory (LSTM) for modelling energy-consumption time series and predicting future consumption using only the historical energy-consumption record. The disadvantage of NAR models being affected by external inputs has been addressed by including NARX models, and the advantage of adding memory to the model by incorporating the ENN and LSTM are included in our study.

The main goal of this paper is to propose a methodology for energy consumption forecasting by making optimal use of existing resources. In addition, since this method provides two essential features (i.e. the good fit of the ANN for time series and the improvement of these models by GA optimization which avoids entrapment in a local minimum), our method therefore enables us to obtain optimal solutions.

The ANN deployed in our study implementation have been fed with raw data with treated missing values extracted from energy consumption meters in buildings on a daily basis. The ANN were trained with real data sets obtained from buildings at our University and the results showed a prediction mean square error of 0.013 in the worst case and 0.0003 in the best. We also showed that different types of NN such as Elman, LSTM can even improve these results.

For NN parallelization, we have deployed a simple parallelization of a map/reduce-like algorithm based on manager-workers which are connected by a crossbar switch on an Intel® Core™ i7-6700 processor (CPU 3.40GHz, 16 GB RAM), which yields an excellent enhancement of the time cost for the four NN used (NAR, NARX, Elman and LSTM) for implementing the model's algorithms.

The paper is structured as follows: Section 2 presents the methodology proposed to obtain a feasible solution to energy consumption time series prediction in buildings, this section also introduces mathematical models of NAR, NARX, Elman and LSTM and its graphical topologies; Section 3 examines the genetic algorithm used to model the time series of one of the energy consumptions for one of the buildings in this study; Section 4 discusses the dataset comprising raw consumption-data for one year from various buildings at our University; Section 5 details the results obtained in the different tests conducted in the study; and finally, Section 6 outlines our conclusions and details some practical implications.

2. Methodology

This section presents the proposed method for energy time series prediction and for minimizing cost over time, enabling full advantage to be taken of available energy resources and innovations to be developed that will provide better results when applied to the use of these resources. The first part of our method deals with data collection and pre-processing. The second part examines the forecast modelling tool. The third step explores genetic optimization with the integration of parallelization techniques. In the final step, the obtained results are validated and analyzed.

Diverse techniques have been employed to solve forecasting problems for many years with different scopes. For instance, in medicine —studies have been carried out to predict and reduce abdominal aortic aneurysm diseases using hemodynamic prediction [30] or to predict drug responses in cancer based on multiple types of genome using Regression Vector Machine [31]—, in marketing —data analysis of data produced by social networks such as Facebook, YouTube, LinkedIn and Twitter to predict influenza epidemic or stock market trends using Self Organizing Fuzzy Neural Networks and Support Vector Machine [32]— or environmental sciences —Jung et al. [33] applied a Genetic Algorithm and a Least Squares Support Vector machine to predict daily building energy consumption and in Deb et al. [34] a complete time series forecasting methods review employed in this subject is done, where other techniques are employed in recent years, such as Grey prediction models and Fuzzy Systems—. There are also recent works that combine time series techniques such as Discrete Wavelet Transform and Empirical Mode Decomposition in order to improve electric load forecasting [35] or even, Deep Learning models, i.e.: Xueheng et al. [36] propose an ensemble Deep Learning model with Empirical Mode Decomposition for load demand prediction, and demonstrate that these models show advantages when prediction horizon increases.

The scientific community has conducted a large number of studies into the problem of energy time series forecasting. Artificial Neural Networks (ANN) have proved promising because of the good/excellent results [2, 37, 38] they yield. In this paper, we use four widely known ANN models: the non-linear autoregressive neural network (NAR) and the non-linear autoregressive neural network with exogenous inputs (NARX), Elman neural network (ENN) and Long Short-Term Memory neural network (LSTM).

The NAR network allows us to model energy-consumption time series most simply. This model is capable of predicting future consumption by using only the historical energy-consumption record. Xian Zhang *et al.* report good performance forecasts for electric vehicle sales in the automobile industry with an NAR neural network [39], although, as they also point out, the main disadvantage of NAR models is that they may be affected by other external factors. As a result, it is necessary to extend ANN-based models in order to be able to integrate more information that can enrich these models [40].

The ENN is a less well known model, however, this neural network introduces a new significant term, crucial when historical information is processed. This is the concept of *memory*. Thus, the ENN's architecture adds a new temporal component to consider previous states in the network to predict the future values of the time series. The ENN has demonstrated excellent performance, especially at the time series problems where past behaviour guides future responses [41, 42] and have proved to be a strong competitor against NAR and NARX models [6].

Finally, due to the increasing interest in Big Data technologies and Deep Learning methods, the LSTM neural network has become very popular over the past few years. The LSTM is the most sophisticated model of all presented here and is also a strong competitor if sufficient data is available. Some studies compare ENN and LSTM architectures achieving very similar results [43], in that study, Mohab et al. show that the ENN is stronger than the authors expected and benefits greatly from their approach. The LSTM models have been also exploited in the energy field and have yielded remarkable outcomes [44, 45].

In this study, a new more efficient algorithm must therefore be implemented to use computing resources in the best possible way. In this context, one classic CHC [26] has been adapted here to improve both ANN accuracy and time-cost.

2.1. NAR and NARX models

Artificial neural network are very powerful, accurate techniques and are currently used for modelling and predicting in various fields. There are recent proposals in medicine which combine the NAR neural network and the autoregressive integrated moving average (ARIMA) to forecast the incidence of tuberculosis [46] or applications of this ANN for predicting incidence tendency of haemorrhagic fever with renal syndrome [47]. In the sphere of finance, these models have been used to forecast stock market returns [48], for fraud detection [49], or even smart card security for public transportation applications based on a novel neural network analysis of cardholder behaviour [50]. In the EE domain, there is a broader array of applications, e.g. simulation-based energy optimization is presented in [51] by applying a web-based parallel genetic algorithm to reduce the computation time for a series of test buildings in Spain. Petri *et al.* present a modular optimization model for reducing energy consumption in large-scale building facilities using ANN [52]. An updated review of time series-based forecasting techniques for building energy consumption can be found in [34].

We should first define the concept of energy consumption before modelling it. To this purpose, energy consumption can be described as a time-series $y(t)$ which represents the energy consumption performed at time t. In many circumstances, the data obtained belong to a fleeting, transient and ephemeral behaviour of the building energy consumption and since this decreases the effectiveness of linear methods, a non-linear approach is therefore recommended. A non-linear autoregressive neural network (NAR) can be modelled using Equation 1, where $\hat{y}$ is the current value of a data series

y at time t, modelled by the p past-values of the series. In principle, $h(\cdot)$ is an unknown non-linear function which is approximated by the optimizing process that is carried out to obtain the optimal weights and bias of the network. The error of the network's estimation of the value y at time t [9, 53] is represented by $\epsilon(t)$:

$$\hat{y}(t) = h\big(y(t-1), \dots, y(t-p)\big) + \epsilon(t) \tag{1}$$

Similarly, the non-linear autoregressive with exogenous inputs (NARX) is also used when data not only depend on the total amount spent but also on other possible factors of influence. Nagy *et al.* [54] use weather conditions as a model feature, and a more accurate predictor of energy consumption is supplied for conducting the time series modelling process and previous work has proved that this has certain advantages [9]. However, one such advantage of using extra information by dynamic feedback input comes with one main disadvantage: it provides a more complex alternative model, where the uncertainty of the additional data may limit the expected performance of the initial model. The NARX model can be described as the following mathematical function:

$$y(t) = h(x(t-1), \dots, x(t-p), y(t-1), \dots, y(t-p)) + \epsilon(t) \tag{2}$$

where $x(t-i)$ is the external time series at time $t-i,\ i \in [1, p]$ and p are the number of past values used. It should be noted that the NAR model given by x is zero. The structure of these two models can be found in Fig.1. Both models are described by the U and W matrices. U_{ij} is the weight between input i and the hidden neuron j, and h is the number of neurons in the hidden layer. W_{ij} represents the weight between the hidden neuron i and the output neuron j and o defines the number of output neurons in this layer. The last matrix, V, specifies the weights for the connections between the exogenous input and the hidden layer. The inputs $y(t-p)$ and $x(t-q)$ where $p, q \in [1, n]$ are the input time series with p past values and the q previous values employed to model the future value of the series $y(t+1)$.

In every case, the b parameter is the bias associated with its neuron.

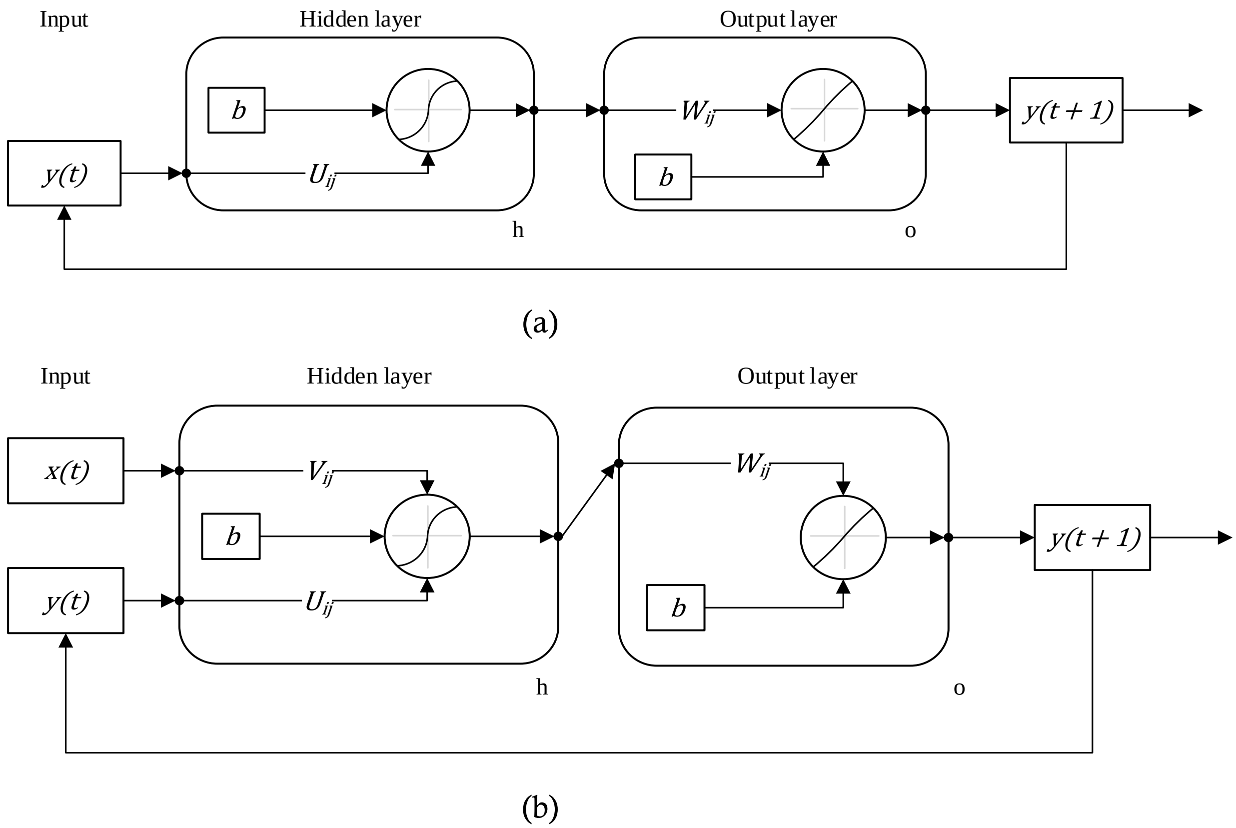


Fig. 1. Representing the structure of a non-linear autoregressive neural network (NAR) using one time series as the input and the next value (a) as the output, and with exogenous inputs NARX (b) where an extra time series is included in the input layer.

2.2. Elman Neural Network

The Elman model (ENN) is a type of network with a recurrent topology and was elaborated by Jeffrey Elman [55]. ENN are a satisfactory time series forecasting method and have proved to be a fast, accurate tool for making future predictions in a wide range of scenarios [56-58]. As with the previous ANN, these models can be found in financial time series prediction to forecast the stock market price indexes [41] and have many different applications such as that proposed by Chu *et al.* in [59] which presents an ENN to identify elderly fall signals. In the EE field, an ENN was developed by Kelo and Dudul [42] to predict electrical power load due to temperature variation. A hybrid model is proposed in [60] for short-term load forecasting and the article also includes a genetic algorithm to achieve the optimal ENN structure. A combination of wavelet and this recurrent network is developed by Sami *et al.* to identify the location of energy transmission faults [61].

The energy forecasting problem works with the evolution of data over time and results in a model that is capable of recording previous results, because the consumption normally shows a cyclic behaviour that justifies ENN deployment. This network introduces recurrence to the network through the addition of a set of units called context (or state) to introduce the concept of *memory*. State neurons acquire the input from the previous hidden layer and return the output to the next hidden layer. This recurrent connection allows the ENN to detect and learn time-varying patterns.

The first difference between the ENN and NAR(X) models is the context layer shown in Fig. 2. The *state neuron layer* has the previous values of the hidden nodes obtained previously: at time t, the output of the hidden neuron will be the input of all the hidden neurons at time $t+1$ and therefore, at time t, the context units will have the hidden neurons values at time $t-1$ [62]. Our decision to adopt this kind of neural network is supported by previous studies, where ENN have yielded significant results, thus demonstrating their usefulness and effectiveness [6, 57, 60-63].

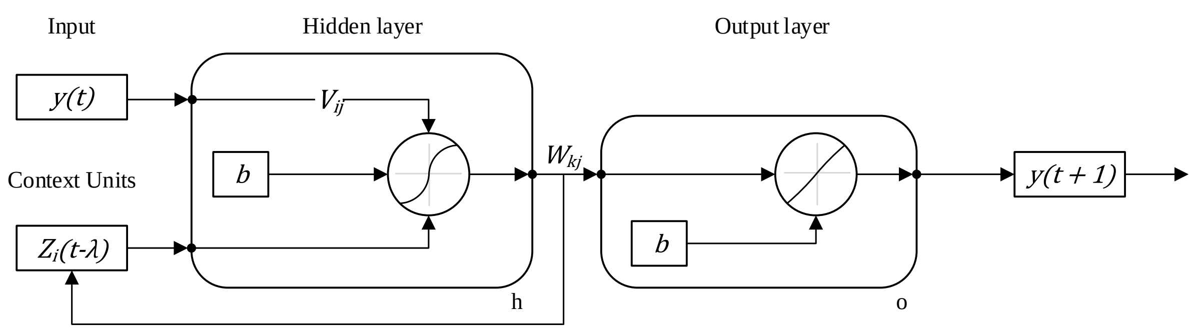


Fig. 2. Representing the topology of the Elman Neural Network.

The output of the ENN $y(t+1)$ is calculated in Equations 3 and 4. The equations detail a more general ENN, with a number of k outputs, the values of index k of the equation above can be further adjusted to fit future applications accordingly, and $y_k(t+1)$ represents the output of neuron k in the last layer. In this study, $k=1$, and G is the activation function of the output layer; W_{kj} is the weight associated with the connection of the hidden neuron j and the neuron k of the output layer; and $S_r(t-1)$ is the state value corresponding to the neuron r at time t.

$$y_k(t+1) = G\left(\sum_{j=0}^{h} W_{kj} S_j(t-1)\right) \quad \textbf{(3)}$$

Similarly, $S_j(t)$ is the output of the neuron j in the hidden layer and is calculated as follows:

$$S_j(t) = F\left(\sum_{r=1}^{h} U_{jr} S_r(t-1) + \sum_{i=1}^{I} V_{ji} y_i(t)\right) \quad \textbf{(4)}$$

where V_{ji} is the weight of the connection between the input neuron j and the hidden neuron i; $y_i(t)$ is the input i at that time and F is the activation function of the hidden neurons; I and h are the number of neurons in the input and hidden layer, respectively; and U_{jr} is the weight of the connection between the neuron j in the context layer and the neuron r in the hidden layer. In the figure, $Z_j(t-\lambda)$ represents the value of the hidden neuron j at time $t-\lambda$ where $\lambda \in [1, \mathbb{N}]$ and indicates the past values of the hidden neurons stored.

The Levenberg-Marquardt (LM) backpropagation algorithm has been used as the network training function in order to update weight and bias values, according to LM optimization. This is often the fastest algorithm that ensures the best convergence and yields a minimum error in function approximation problems [64].

2.3. Long Short-Term Memory Neural Network

The recurrent neural networks with long short-term memory (LSTM) have recently risen as a powerful and scalable model for diverse learning problems related to sequential data. LSTM, in a similar way to ENN, are effective at learning temporal dependences with the advantage that they do not experience the optimization barriers of the simple recurrent networks [65] and have been employed to solve countless problems. This covers activity recognition —i.e.: Ordóñez and Roggen [66] suggest a Deep Convolutional framework for activity recognition based on convolutional and LSTM recurrent units— handwriting recognition —i.e.: Xiaoqiang et al. propose an innovative recurrent neural method to learn discriminate binary codes, and they use

LSTM to learn feature vector by using the convolutional feature map as input for image retrieval [67]—, handwriting generation, language modeling and translation, acoustic modeling of speech, speech synthesis, analysis of audio and video among others [65].

Due to the explosive growth of data in recent years, it is common to find these kind of models with several hidden layers and with a high number of neurons in order to deal with the high complexity and the vast amount of information to be processed. As stated above, the LSTM neural network provides a more complex architecture than ENN, and it is not utilized if little data is available. In this paper, the LSTM architecture chosen for this problem is illustrated in Fig. 3. The main components of the LSTM are a sequence input layer and the LSTM layer. The first layer is the time series data, the second layer is a recurrent layer that enables support for time series and sequence data in the ANN and learns temporal dependencies between time steps of sequence data. Finally, the architecture ends with a fully connected layer which multiplies the input by a weight matrix and then adds a bias vector, and a regression output layer. An important characteristic of the LSTM is that it has been designed to learn to bridge time intervals in excess of 1000 steps even in case of incompressible, noisy input data, without loss of short-time-lag capabilities [68].

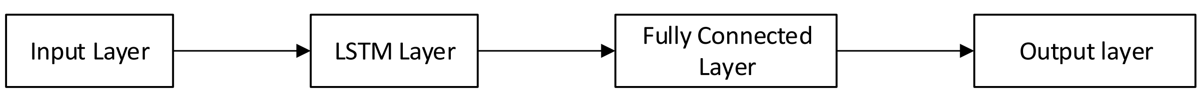


Fig. 3. Representing the topology of the Long Short-Term Memory Neural Network.

LSTM architecture distinguishes itself from the rest by the Memory Cells and the Gate Units. Fig. 4 shows the structure of the cell c_j and its gate units in and out. The self-recurrent connection indicates feedback with a delay of d time steps. The hidden units explicitly manage the flow of information as a function of both the state and input. The state stored in this structure is either deleted by a forget gate or saved indefinitely. Knowledge is thus guaranteed to be transferred over long lapses of time.

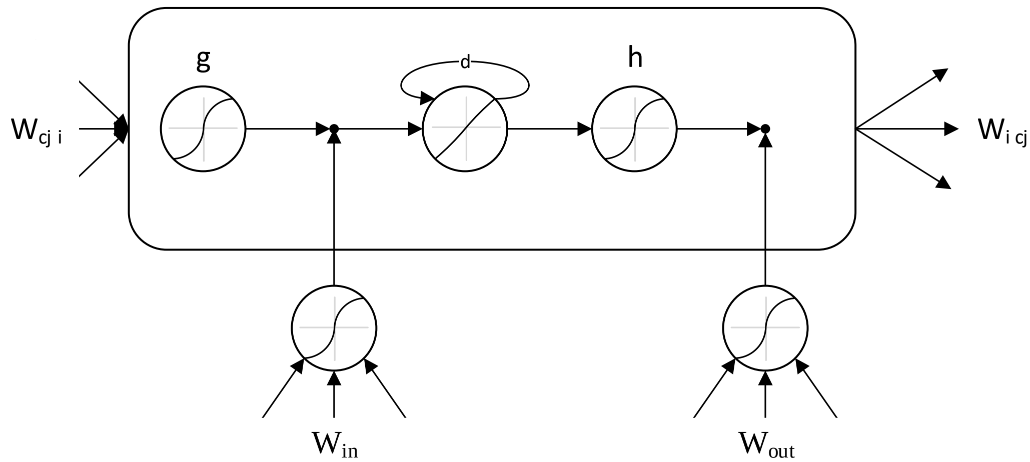


Fig. 4. Basic architecture of memory cell of the LSTM neural network.

In this case, the ADAM algorithm is used for training as the gradient based optimizer, instead of SGD method. The ADAM outperformed SGD in terms of faster convergence and lower error ratios [69]. The main challenge of the training algorithms —in our case: LM for NAR, NARX and Elman networks, and ADAM for LSTM— is that it is typical for solutions to converge to a local minimum. We therefore propose an evolutionary algorithm in order to obtain better outcomes and to optimize the ANN results. The suggested GA is discussed further in the following section. We have selected these four neural networks for testing the validity, adaptive capacity and the reliability of our proposal working with different ANN architectures.

3. Genetic Algorithm

A Genetic Algorithm (GA) is a stochastic optimization method based on the concept of natural evolution. GAs comprise a population of chromosomes (or individuals) and each represents the possible solution to the problem. Each individual has an associated objective value which designates the goodness degree of a solution. Furthermore, the GA has three essential functions: selection, crossover and mutation.

In this paper, the adaptation of the binary "Cross generational elitist selection, Heterogeneous recombination, and Cataclysmic mutation" (CHC) algorithm has been adapted for real-coded solutions [26]. The CHC algorithm finds the optimal neural network weights and biases, and has been adapted to a parallel approach in order to reduce time-cost computation.

The computation time taken to reach a good solution, and of course to improve this solution, is one of the main driving forces behind this study. Maximizing the potential of available resources is an important task and one that is often neglected in many studies and rarely explored as it is in this paper. A large number of publications focus on developing excellent models but do not mention the time cost involved, even though constrained time cost is a common requirement for industry and business.

Nevertheless, various examples of published studies can be found and He and Sun [70] presented their convolutional neural network research to fulfil the requirement of a constrained time budget. They investigate the accuracy of these models under a constrained time cost and design a very fast, accurate architecture that reaches the top-5 error. Lee *et al.* presented an advanced stochastic time-cost trade-off analysis, based on a critical path method guided by a genetic algorithm in order to reduce the computation time, reliability and usability of a previous algorithm. They use the GA for optimization and also to identify the new initial parent chromosomes [71]. A least squares support vector machine to predict building energy consumption improved with real coded GA is used in [33], the purpose of which is to obtain a faster computation speed and greater prediction accuracy. The method performed better in terms of convergence time and iteration economy.

The study presented here exploits the advantages of the CHC algorithm for searching for good solutions and the ANN's disadvantage of falling in a local minimum. Our proposal also benefits from the computational capabilities of the CPU to cope with the high GA time-cost. The proposed parallel GA improves the excessive computation time by distributing the iterative tasks to different workers. We should first explain our coding of the ANN algorithm and so by considering Fig. 1, an individual can be codified as shown in Fig. 5.

w_{11}^{x}	w_{12}^{x}	$\cdots$	w_{hn}^{x}	w_{11}^{y}	w_{12}^{y}	$\cdots$	w_{hn}^{y}	w_{11}	w_{12}	$\cdots$	w_{ho}

Fig. 5. Genetic encoding of a neural network

where n is the number of connexions between the input and the first network layer, h is the number of neurons in the hidden layer and o the number of neurons in the output layer; w_{ij}^{x} represents the weight between the input neuron j and the hidden neuron i; w_{pq}^{y} represents the weight associated with the recurrent connection between the output neuron p, and the hidden neuron q, and w_{uv} represents the weight associated with the connection between the hidden neuron u and the output neuron v. When the chromosome acts as NAR neural network architecture, then w^{x} is not part of the solution. The structure of the assumed algorithm is shown in Fig. 6. This figure includes a flowchart to explain the different steps of the procedure.

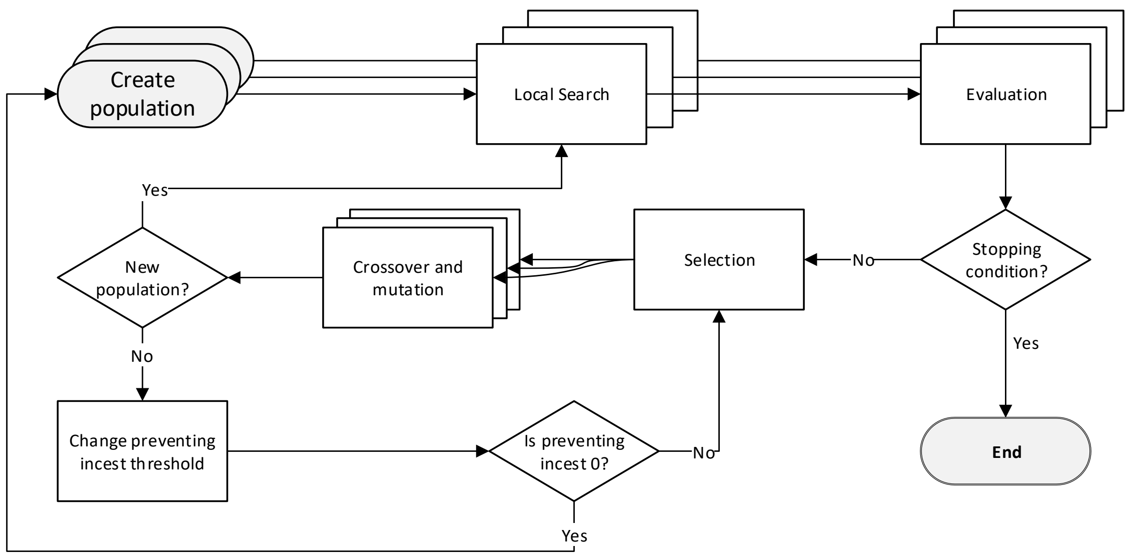


Fig. 6. Flowchart of the proposed genetic algorithm, adapted from Eshelman's CHC.

Each step is described as follows:

1. A totally random, initial population is created. Each gene is initialized in a defined range $[G_{min}, G_{max}]$ and the individuals are created in parallel. Parallelism is performed at the level of the chromosome, illustrated in Fig7a.

2. The local search is performed according to the following formula:

$$D(t) = F\big(S(t), X(t)\big) \tag{5}$$

where $D(t)$ is the expected response in the output neurons on the instant time t, $F(\cdot)$ is the activation function of the output neurons, $S(t)$ is a vector containing the hidden neuron state of the network on t, and $X(t)$ represents a vector with input values of the network on t.

The prime aim is to minimize an objective function which depends on the network weights and the expected output. This idea is illustrated by the following equation:

$$G(s) = \sum_{t=0}^{T} (\|D(t) - O(t)\|)^2 = E^T E \quad (6)$$

where $D(t)$ is the expected output on the instant time t, $O(t)$ is the output on the instant t, E refers to a matrix with $Tx1$ values with the output layer error on t, and s is the variable vector to optimize. The parallelism in this step is presented at the level of the individual. According to resources available, a number of individuals are simultaneously optimized with the local search as shown in Fig. 7b.

The hybridization of the evolutionary algorithm and local search method results in the well-known memetic algorithm (MA).

3. The individual objective value is measured using mean square error (MSE) in order to optimize an individual with the local search:

$$MSE = \frac{1}{n-m} \sum_{t=1}^{n} \left(D(t) - O(t)\right)^2 \quad (7)$$

where n is the sample size and m is the number of parameters in the model. This function is parallelized as in the two previous steps. Computer resources are divided up among the individuals so as to avoid idle times which force 100% CPU to be used. This approach is illustrated in Fig. 7a.

4. The individuals are selected using the roulette method [72]. Each individual in the population is assigned a probability of being selected. This probability is proportional to its adjustment, in other words, to its error. The best individuals receive a greater slice of roulette than the worst.

The operator for generating new offspring is the BLX-α crossover operator [73]. A new son $H = (h_1, \dots, h_n)$ is born according to a random number h_i selected in interval

$[c_{min} - I \cdot \alpha, c_{max} + I \cdot \alpha]$, where $c_{max} = \max(c_i^1, c_i^2)$, $c_{min} = \min(c_i^1, c_i^2)$, $I = c_{max} - c_{min}$, and c_i^p is the gene i of the parent chromosome p. The parameter α is introduced by the user between $[0, 1]$. This procedure has been parallelized at the gene level as illustrated in Fig. 7b. The new genes of an individual are calculated in tandem. In this phase, this approach has been followed because the crossover is performed if it overcomes a defined probability. Since not all the chromosomes reproduce, this could result in idle times.

5. The mutation operator is responsible for selecting an individual gene and setting a random value between $[G_{min}, G_{max}]$. This step has been parallelized as in the previous phase.

6. The next population shall be constructed by the m best individuals by considering individuals of the previous population and its offspring.

7. The algorithm ends when a number of generations g is achieved.

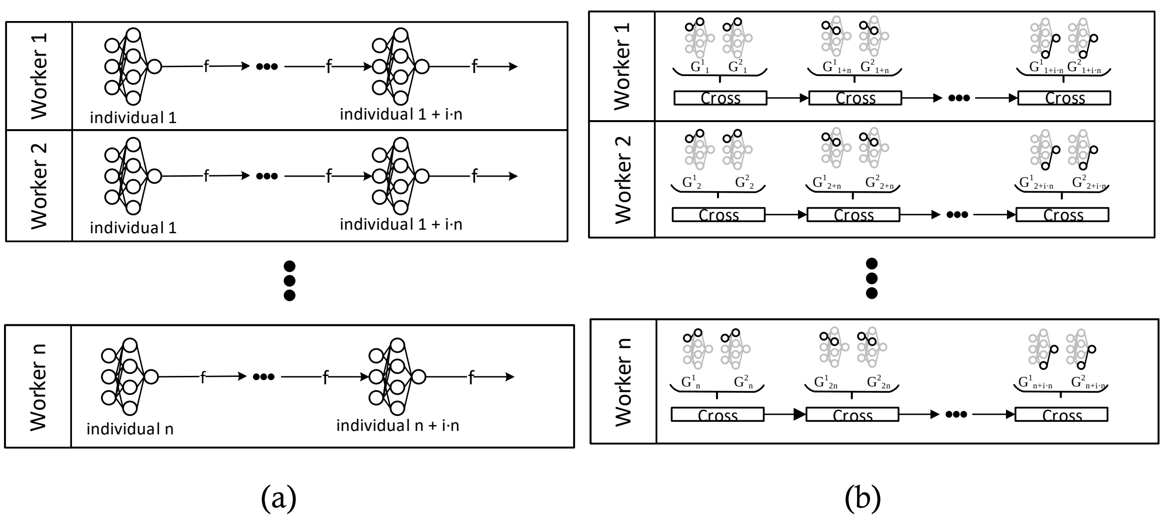


Fig. 7. Parallel distribution of the loads into n workers, f denotes procedures performed on the individual, procedures which include individual creation, local search and objective function (a). Parallel distribution of the chromosome divided into n workers. G_t^p denotes the gene t of the parent p (b)

There is also a diversification strategy to prevent local minimum stagnation so that the algorithm can continue to search the solution space and for this purpose, a re-initialization mechanism is applied. This procedure means that it is necessary to return to Step 1. The criterion established for resorting to this procedure is if there has been no improvement in the results in x generations.

The parallelism depicted in Fig. 7a splits the load into n workers defined as follows:

$$CPU = \{w_1, w_2, \dots, w_n\} \tag{8}$$

The total number of models is k:

$$individuals = \{1, 2, \dots, k\} \tag{9}$$

These workers are the control process units (CPUs) and each CPU w_j assumes the operations for the individual set C_{Wj}, according to Equation 10:

$$C_{w_j} = \{w_j + i \cdot n \,|\, i \in \{0, 1, \dots, \alpha - 1\} \text{ and } w_j + i \cdot n \leq k \} \tag{10}$$

where α is the number of assignments given to each worker and calculated as follows:

$$\alpha = \left\lceil \frac{k}{n} \right\rceil \tag{11}$$

Similarly, the cross-cutting of chromosome H (t genes in length) is carried out by means of the following equation:

$$C_{w_j} = \{w_j + i \cdot n \mid i \in \{0, 1, \dots, \beta - 1\} \text{ and } w_j + i \cdot n \leq t \} \tag{12}$$

where β is the total number of genes allocated to the available workers:

$$\beta = \left\lceil \frac{t}{n} \right\rceil \tag{13}$$

4. Dataset

In this work, a data set has been collected from a building automation system that records energy-consumption over time. These specific systems usually control energy wastage due to the heating, ventilation, air conditioning and lighting systems of a building. Our study uses data on energy-consumption and weather conditions of the University of Granada (UGR, Granada, Spain). The dataset includes data from two buildings with the same demographic characteristics.

The UGR comprises five campuses: Centro, Cartuja, Fuentenueva, Aynadamar and Ciencias de la Salud, spread over the city of Granada. In total there are 22 colleges, 5 technical engineering schools, 8 training centres and 5 additional centres for culture, sport and general services.

Since current Spanish Data Protection Laws prevent us from specifying the exact location of the buildings and facilities, we numbered them from 1 to 8. Buildings have been selected in light of two representative energy-consumption data for each campus.

Fig. 8 depicts two examples of raw consumption-data for one year. These two consumptions show a linear upward trend since the original data are recorded by the building's energy meter. The raw data shown in the figure therefore represent consumption to date, i.e. the information stored is cumulative consumption. The energy consumed c_t at time t is calculated using the energy price at that moment in time, D_t, and the previous one D_{t-1} as the following equation illustrates:

$$c_t = D_t - D_{t-1} \tag{14}$$

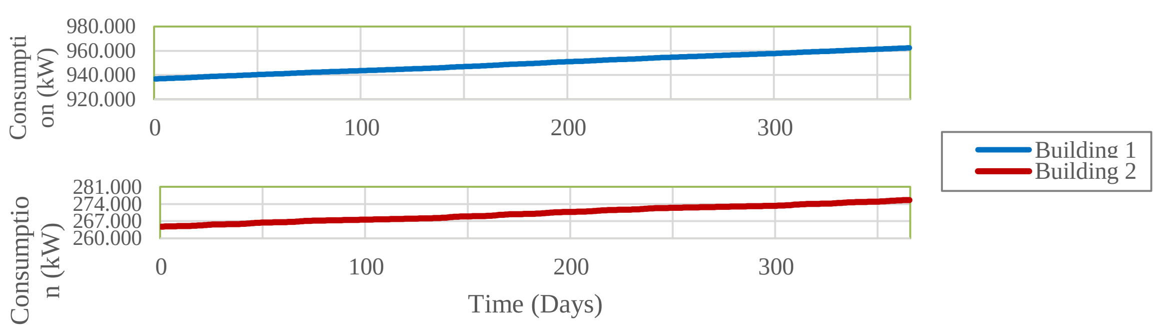


Fig. 8. Meter readings for two examples of raw data consumption over a year

5. Results

This section presents all of the results obtained in the various tests that have been conducted. Fig. 8 gathered an example of raw data from an energy meter, which periodically accrue the total consumption. It is therefore necessary to summarize and transform the data in order to present them in a usable format. Any incomplete, noisy or unreliable data are also dealt with, and incomplete data have been filled using a linear interpolation imputation method. This method fits a straight line between the endpoints of the gap and enables the missing values to be calculated in a straightforward way by employing the following line equation [74]:

$$y = y_1 + k(x + x_1) \tag{15}$$

The value k is calculated as follows:

$$\mathrm{k} = \frac{(\mathrm{y}_2 - \mathrm{y}_1)}{\mathrm{x}_2 - \mathrm{x}_1}; \mathrm{x}_1 < \mathrm{x} < \mathrm{x}_2 \text{ and } \mathrm{y}_1 < \mathrm{y} < \mathrm{y}_2 \tag{16}$$

where y is the interpolant, x symbolizes the time point of the interpolant, y_1 and x_1 are the coordinates of the starting point of the gap, y_2 and x_2 indicate the coordinates of the end point of the missing interval.

Fig. 9 depicts two examples of the results captured by the imputation method. It should be noted that linear imputation is the best choice in this problem due to its simplicity and quick computation and because it responds reasonably to the consumption behaviour. Although other imputation methods have been tested (e.g. cubic spline imputation, shape-preserving piecewise cubic interpolation, previous neighbour interpolation, next neighbour interpolation, nearest neighbour interpolation), their behaviour is undesirable and meaningless in this problem. Two such examples of this behaviour are negative consumption or a zero consumption period with a single large spike.

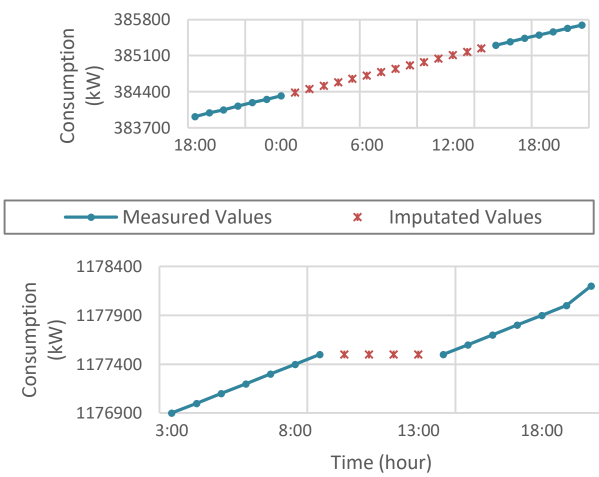


Fig. 9. Two examples of linear imputation values for the energy meters.

The raw data with missing values that have been treated are modified to assemble daily consumption. Energy consumption has also been normalized between 0 and 1 to have the same range of values for each input to the NN, thereby ensuring that the model does not give more weight to the higher range attributes. The data would then be reconstructed as shown in Fig. 10.

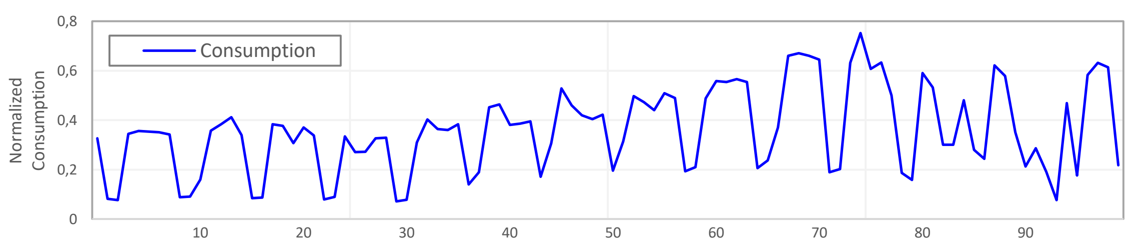


Fig. 10. Daily normalized consumption over a 100-day period.

So that such results may be obtained, the following parameters have been set for the GA: the population size has been set to 25 individuals; the maximum and minimum gene values are 10 and -10, respectively; the stop criteria were established at 100 generations; and the crossover and mutation probability has been set at 90% and 10%. In order to train all the models, the dataset has been randomly split into training and test sets with 70% of the examples allocated to training and 30% to testing. So as to obtain the best parameters for the models, ten tests have been run. Table 1 displays all the errors for each network structure and reports the best number of neurons for each case. It is evident that the NAR and NARX models obtain the best error with 10 neurons, although the optimum is achieved in two consumptions with 8 and 9 neurons, respectively. However, for most buildings, the ENN obtains the optimum solution with 9 neurons. Likewise, the best results of the LSTM network are acquired with 8, 9 and 10, although if Table 1 is looked closely, we realise that if we were to remove the top three cases (cases 2, 3 and 10), 5 and 7 become the best neuron parameters.

Table 1. Results of experimental time series consumptions: comparison of the three neural networks optimized using the memetic algorithm to display the mean square error

	Neurons									
Model	1	2	3	4	5	6	7	8	9	10
NAR										
Cons. 1	0.0215181	0.0129090	0.0118595	0.0110900	0.0107224	0.0100682	0.0099646	0.0098942	0.0093493	0.0092648
Cons. 2	0.0401715	0.0211694	0.0190867	0.0172446	0.0162044	0.0159827	0.0154314	0.0146710	0.0149999	0.0147922
Cons. 3	0.0367283	0.0205938	0.0168575	0.0155420	0.0146244	0.0142023	0.0132543	0.0128636	0.0128842	0.0122394
Cons. 4	0.0008630	0.0008324	0.0008291	0.0008266	0.0008268	0.0008216	0.0008212	0.0008205	0.0008222	0.0008217
Cons. 5	0.0134595	0.0105834	0.0100412	0.0091431	0.0089837	0.0086461	0.0083847	0.0082914	0.0082652	0.0080677
Cons. 6	0.0279272	0.0132078	0.0101502	0.0089602	0.0083126	0.0079794	0.0073542	0.0069932	0.0067491	0.0064842
Cons. 7	0.0377470	0.0191129	0.0151643	0.0140364	0.0134968	0.0128224	0.0121629	0.0117132	0.0114873	0.0114223
Cons. 8	0.0254187	0.0153050	0.0136017	0.0112555	0.0105364	0.0100957	0.0101201	0.0102848	0.0097331	0.0097186
NARX										
Cons. 1	0.0172769	0.0134489	0.0110335	0.0102405	0.0105713	0.0100784	0.0096436	0.0096312	0.0090996	0.0092894
Cons. 2	0.0351362	0.0206990	0.0192705	0.0180993	0.0180576	0.0168802	0.0158672	0.0155706	0.0144174	0.0146818
Cons. 3	0.0326071	0.0191206	0.0170564	0.0165537	0.0158502	0.0149622	0.0143571	0.0129777	0.0135710	0.0127612
Cons. 4	0.0009212	0.0008938	0.0008991	0.0009013	0.0009001	0.0008833	0.0009095	0.0008910	0.0008932	0.0008797
Cons. 5	0.0104558	0.0084836	0.0079764	0.0073544	0.0073420	0.0070150	0.0069223	0.0066998	0.0065222	0.0064522
Cons. 6	0.0200192	0.0112923	0.0096900	0.0089700	0.0085108	0.0079184	0.0078619	0.0074871	0.0074983	0.0071489
Cons. 7	0.0396858	0.0197610	0.0159657	0.0150626	0.0146919	0.0125092	0.0116735	0.0117385	0.0113365	0.0112290
Cons. 8	0.0187236	0.0130301	0.0106711	0.0099076	0.0097993	0.0093895	0.0092843	0.0092153	0.0090727	0.0088996
Elman										
Cons. 1	0.0181928	0.0080617	0.0071009	0.0065562	0.0064105	0.0057446	0.0053506	0.0039186	0.0030866	0.0036516
Cons. 2	0.0327652	0.0145094	0.0133701	0.0120689	0.0110340	0.0092417	0.0093951	0.0072378	0.0075160	0.0058396
Cons. 3	0.0326860	0.0116246	0.0099089	0.0095616	0.0092085	0.0083863	0.0079345	0.0076725	0.0065537	0.0060289
Cons. 4	0.0008494	0.0007999	0.0007916	0.0006404	0.0007576	0.0006625	0.0003674	0.0005721	0.0006332	0.0007974
Cons. 5	0.0118462	0.0056386	0.0053088	0.0052252	0.0049409	0.0044106	0.0040186	0.0036410	0.0033035	0.0033780
Cons. 6	0.0220548	0.0065676	0.0056810	0.0056224	0.0052802	0.0042862	0.0038940	0.0033211	0.0024795	0.0028352
Cons. 7	0.0331632	0.0108750	0.0094952	0.0093967	0.0088470	0.0085891	0.0076741	0.0071685	0.0062178	0.0068952
Cons. 8	0.0200358	0.0075694	0.0068414	0.0067670	0.0063285	0.0061455	0.0056810	0.0048660	0.0041578	0.0047037
LSTM										
Cons. 1	0.0092897	0.0033385	0.0020597	0.0025722	0.0021547	0.0014317	0.0010614	0.0011855	0.0006317	0.0009949
Cons. 2	0.0421114	0.0129763	0.0084729	0.0081993	0.0050532	0.0041174	0.0031338	0.0032180	0.0033550	0.0027195
Cons. 3	0.0291953	0.0085740	0.0035432	0.0044459	0.0028404	0.0025456	0.0017075	0.0012760	0.0016341	0.0014364
Cons. 4	0.0022123	0.0019221	0.0018247	0.0018727	0.0018015	0.0017007	0.0015570	0.0015211	0.0013734	0.0014758
Cons. 5	0.0191191	0.0102992	0.0082567	0.0065753	0.0049700	0.0046626	0.0048182	0.0052757	0.0042051	0.0037791
Cons. 6	0.0397496	0.0232968	0.0143120	0.0088149	0.0077170	0.0065207	0.0069664	0.0081868	0.0033985	0.0045700
Cons. 7	0.0293439	0.0106638	0.0057946	0.0041042	0.0034333	0.0012890	0.0023552	0.0009120	0.0014171	0.0010371
Cons. 8	0.0232106	0.0067959	0.0054637	0.0031309	0.0036474	0.0024250	0.0023462	0.0020388	0.0013697	0.0016317

The results are summarized in Table 2, in nearly all cases the best are obtained with the LSTM network, nevertheless, there are three cases where the Elman network achieves better outcomes: in every case, Elman and LSTM are well below half the computed MSE values for the other two models. It is also interesting to note that the NAR and NARX networks have a similar error. Although in previous studies, the

NARX models with the exogenous input performed best in every case [9], here the MA optimizes the NAR in such a way that it enables a better result to be obtained in the fit of the NARX neural networks in three cases: Consumption 3, 4 and 6. The errors of both models are quite similar.

Table 2. Mean square error performance of the best prediction NAR, NARX, Elman and LSTM networks optimized with the memetic algorithm.

Building	NAR	NARX	Elman	LSTM
Consumption 1	0.0092648	0.0090996	0.0030866	**0.0006317**
Consumption 2	0.0146710	0.0144174	0.0058396	**0.0027195**
Consumption 3	0.0122394	0.0127612	0.0060289	**0.0012760**
Consumption 4	0.0008205	0.0008797	**0.0003674**	0.0013734
Consumption 5	0.0080677	0.0064522	**0.0031310**	0.0037791
Consumption 6	0.0064842	0.0071489	**0.0024795**	0.0033985
Consumption 7	0.0114223	0.0112290	0.0062178	**0.0009120**
Consumption 8	0.0097186	0.0088996	0.0041578	**0.0013697**

One example of the application of our proposal is illustrated in Fig. 11 for case 5, which has the most similar MSE. This illustrates the prediction evolution of the different neural networks performed by the memetic algorithm at various instances of the algorithm, and more specifically, Generations 1, 25, 50, 75 and 100. This graph displays the evolution of MSE performed by the three ANN during the optimization process. It should be noted that the first error obtained has been omitted because of its high value so as not to distract attention from the other results. Fig. 11a shows ANN prediction and the real value of the series in the first generation of the algorithm. It is easily apparent that these models return an almost random prediction because their weights and bias have been randomly initialized, and it is not possible to obtain good results in a single generation. Otherwise, adopting this approach would not be justified.

Nevertheless, Fig. 11b shows how all the models are able to fit the curves more clearly with 25 generations. Fig. 11f supports this assertion because there is an important decrease in the estimated error in every case between generations 40 and 50, after which there is a gradual reduction in the MSE. During the 50th generation, there is improvement in the neural networks and various local peaks have been refined, such as the estimation of the prediction of the consumption at Day 48.

Successive improvements, however, are barely noticeable. Furthermore, from the 50th generation to the end, as the NARX model is not able to improve, the population is reinitialized since the incest threshold has been crossed and no improvement has been found. Something similar occurs with the NAR models which have a softer learning curve than other models. The NAR network was trapped in a local minimum and its population is reset, but in this case, the model achieves a better solution near the 90th generation.

In conclusion, the ENN and LSTM produces very similar results. It is interesting to see how there are considerably wider fluctuations in the ENN in Fig. 11a and these are soon well calibrated. These are not apparent in Fig. 11f because of their high MSE as we explained previously. Another interesting behaviour is illustrated in the same Fig. 11a where LSTM yields the worst prediction, however, it presents a fair view of the trend throughout the whole series. The Fig. 11b shows how the LSTM begins to adjust better and its predictions are in much the same way NAR and NARX models, but soon starts improving and its results are close to the results of the ENN network. In this instance the ENN model produces the best fit in every generation compared to LSTM and also achieves a better forecasting the more generations are performed.

Table 3. Execution time in seconds. Comparison of sequential and parallel memetic algorithm with NAR, NARX, Elman and LSTM models.

	Sequential				Parallel			
Cons.	NAR	NARX	Elman	LSTM	NAR	NARX	Elman	LSTM
1	756	1207	10351	2123	169	242	4714	681
2	835	1582	10330	1990	194	372	4652	557
3	74	1095	10402	2508	166	231	4523	698
4	675	1524	10260	2560	181	353	5084	694
5	555	1040	8908	2480	161	243	4571	718
6	386	1556	5462	2082	117	192	1957	517
7	623	1582	11073	2560	189	351	4631	720
8	574	1129	10937	2504	169	268	4630	646
Mean	644	1339	9715	2351	168	281	4345	654

Finally, Table 3 shows the time cost executions and the computational cost in seconds for every experiment. The table has nine columns: the first identifies the building, and the remaining columns summarize the average execution time breakdown for each test performed. The MA takes 644.29 seconds (≈ 11 minutes) with NAR networks with the sequential version and 168.33 seconds (≈3 minutes) with the parallel process. This represents a time cost improvement of up to 73.87%. It should be noted that the NAR model is the fastest method because its topology is simpler than the NARX and Elman networks. Similarly, the parallel and sequential approaches of the MA with the NARX networks have a time cost of 1339.43 seconds (≈22 minutes) and 281.48 (≈5 minutes), respectively, with a time cost improvement of 78.98%. An unexpected result is obtained with the LSTM and ENN networks, the ENN takes longer to provide the optimal results, and it has an average improvement of 55.28%. Nonetheless, the LSTM achieves in more than half the cases a better error than ENN. On the other hand, LSTM spends far less time to optimize the models, improving time cost by 72.18%. According to the Matlab documentation, the Elman networks are no longer recommended to use, instead they suggest NARX and NAR. This is probably happening because ENN is not optimized in the same way than NAR, NARX and LSTM neural networks.

The codes are executed in Intel® Core™ i7-6700 CPU 3.40GHz, 16 GB RAM memory and Microsoft's Windows 10 (x64) operating system with Matlab R2018a. In accordance with these features, the number of workers has been set to 8.

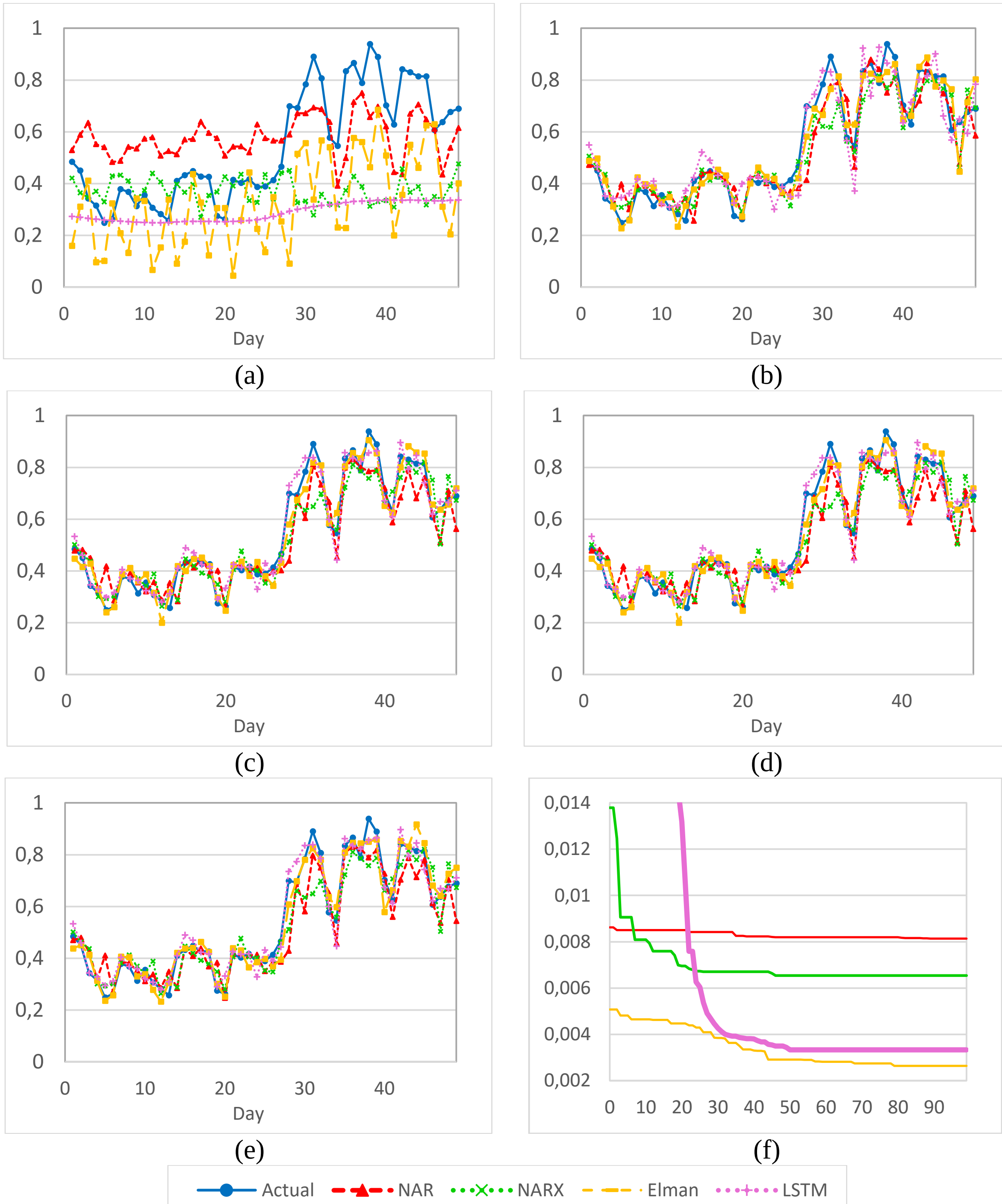


Fig. 11. Example of the optimization process in 100 generations; the model forecasting for the 50-day period in the first generation (a); optimization achieved in Generation 25 (b), in Generation 50 (c), in Generation 75 (d) and the final generation (e); and the mean square error achieved during the process (f) in each generation

6. Conclusions

This article introduces a modified implementation of the CHC adaptive search algorithm to improve and optimize energy consumption forecasting models. The novelty of our approach is to show how the CHC algorithm can be modified to produce a parallel memetic energy efficiency prediction proposal in order to satisfactorily improve the algorithm's time cost.

This paper examines the problem of developing technologies to predict future energy consumption in buildings and the ensuing temporary constraints which are crucial for appropriate building energy management. We tested the usefulness of the genetic algorithm to improve solutions, and explored a method to speed up the process by using parallel techniques which have been applied.

We have compared the performance and computational cost of parallel and sequential implementations of the MA to achieve optimal predictors. The parallel algorithm has been found to be computationally much more efficient than the sequential version, with no negative impact on the quality of the solutions. The proposed, properly optimized models are extremely valuable tools for predicting energy consumption, and parallelization of the optimization method provides a 50% reduction in time in the worst case. The predictor models NAR, NARX, ENN and LSTM are successfully developed and the results support the validity of the proposed approach, achieving an average improvement of 75%.

Our experiments highlight the fact that the ENN and LSTM are the most suitable technique for energy consumption prediction. We should also highlight that all of the models shown here provide good results in terms of time-cost, which confirms the importance of our proposal in every test conducted.

Although good results were obtained in the experiments, in the future we want to enhance the model to improve energy management and cost saving. By way of future work, we therefore plan to use feature selection and clustering methods to detect consumption profiles and abnormal consumption, and identify the relationship between supposedly independent consumption periods and peak demand. We also intend to pinpoint groups or patterns of behaviour with certain specific features. The use of a MapReduce approach will improve the scalability of the methodology and this will result in better system efficiency and greater computational capacity so that larger amounts of data can be processed.

By way of conclusion, although this study achieves good results in reasonable execution times, it would be extremely interesting to perform an additional comparative study using differential evolution (DE) approaches. Additionally, incorporating a new randomization-based method and applying decomposition to the time series will probably improve the study results in future work. Furthermore, the development of new deep learning models to deal with this problem will be an interesting alternative to explore for comparison with our method.

Acknowledgments

This work has been developed with the support of the Department of Computer Science and Artificial Intelligence of the University of Granada, TIC111, and Project TIN201564776-C3-1-R.

Abbreviations

ANN	Artificial Neural Network.
ARIMA	Autoregressive Integrated Moving Average.
CHC	Cross generational elitist selection, Heterogeneous recombination, and Cataclysmic mutation.
DM	Data Mining.
EC	Evolutionary Computation.
EE	Energy Efficiency.
ENN	Elman Neural Network.
GA	Genetic Algorithm.
LM	Levenberg-Marquardt.
LSTM	Long Short-Term Memory.
MA	Memetic Algorithm.
MSE	Mean Square Error.
NAR	Non-linear autoregressive.
NARX	Non-linear autoregressive with external input.
UGR	University of Granada.

Bibliography

[1] E.E. Directive, Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32, Official Journal, L, 315 (2012) 1-56. [http://ec.europa.eu/energy/en/topics/energy-efficiency]

[2] R. Bhandari, J. Gill, An Artificial Intelligence ATM forecasting system for Hybrid Neural Networks, International Journal of Computer Applications, 133 (2016) 13-16. [http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.734.7389&rep=rep1&type=pdf]

[3] T. Ekwevugbe, N. Brown, V. Pakka, D. Fan, Real-time building occupancy sensing using neural-network based sensor network, in: 2013 7th IEEE International Conference on Digital Ecosystems and Technologies (DEST), 2013, pp. 114-119. [http://dx.doi.org/10.1109/DEST.2013.6611339]

[4] C. Fiori, K. Ahn, H.A. Rakha, Power-based electric vehicle energy consumption model: Model development and validation, Appl Energ, 168 (2016) 257-268. [http://dx.doi.org/10.1016/j.apenergy.2016.01.097]

[5] J. Jackson, Promoting energy efficiency investments with risk management decision tools, Energy Policy, 38 (2010) 3865-3873. [http://dx.doi.org/https://doi.org/10.1016/j.enpol.2010.03.006]

[6] L.G.B. Ruiz, R. Rueda, M.P. Cuéllar, M.C. Pegalajar, Energy consumption forecasting based on Elman neural networks with evolutive optimization, Expert Systems with Applications, 92 (2018) 380-389. [http://dx.doi.org/10.1016/j.eswa.2017.09.059]

[7] D. Masa-Bote, M. Castillo-Cagigal, E. Matallanas, E. Caamaño-Martín, A. Gutiérrez, F. Monasterio-Huelín, J. Jiménez-Leube, Improving photovoltaics grid integration through short time forecasting and self-consumption, Appl Energ, 125 (2014) 103-113. [http://dx.doi.org/10.1016/j.apenergy.2014.03.045]

[8] M. Macarulla, M. Casals, N. Forcada, M. Gangolells, Implementation of predictive control in a commercial building energy management system using neural networks, Energy and Buildings, 151 (2017) 511-519. [http://dx.doi.org/http://dx.doi.org/10.1016/j.enbuild.2017.06.027]

[9] L.G.B. Ruiz, M.P. Cuellar, M. Delgado, M.C. Pegalajar, An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings, Energies, 9 (2016) 21. [http://dx.doi.org/10.3390/en9090684]

[10] H. Khosravani, M. Castilla, M. Berenguel, A. Ruano, P. Ferreira, A Comparison of Energy Consumption Prediction Models Based on Neural Networks of a Bioclimatic Building, Energies, 9 (2016) 57. [http://www.mdpi.com/1996-1073/9/1/57]

[11] E. Asadi, M.G.d. Silva, C.H. Antunes, L. Dias, L. Glicksman, Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application, Energy and Buildings, 81 (2014) 444-456. [http://dx.doi.org/10.1016/j.enbuild.2014.06.009]

[12] G. Kumar, M.K. Rai, An energy efficient and optimized load balanced localization method using CDS with one-hop neighbourhood and genetic algorithm in WSNs, Journal of Network and Computer Applications, 78 (2017) 73-82. [http://dx.doi.org/10.1016/j.jnca.2016.11.013]

[13] M. Beccali, G. Ciulla, V. Lo Brano, A. Galatioto, M. Bonomolo, Artificial neural network decision support tool for assessment of the energy performance and the refurbishment actions for the non-residential building stock in Southern Italy, Energy, (2017). [http://dx.doi.org/10.1016/j.energy.2017.05.200]

[14] F. Magoulès, H.-x. Zhao, D. Elizondo, Development of an RDP neural network for building energy consumption fault detection and diagnosis, Energy and Buildings, 62 (2013) 133-138. [http://dx.doi.org/10.1016/j.enbuild.2013.02.050]

[15] T. Muhammed, R.A. Shaikh, An analysis of fault detection strategies in wireless sensor networks, Journal of Network and Computer Applications, 78 (2017) 267-287. [http://dx.doi.org/10.1016/j.jnca.2016.10.019]

[16] F. Ascione, N. Bianco, C. De Stasio, G.M. Mauro, G.P. Vanoli, A new methodology for cost-optimal analysis by means of the multi-objective optimization of building energy performance, Energy and Buildings, 88 (2015) 78-90. [http://dx.doi.org/10.1016/j.enbuild.2014.11.058]

[17] J.L. Viegas, S.M. Vieira, R. Melício, V.M.F. Mendes, J.M.C. Sousa, Classification of new electricity customers based on surveys and smart metering data, Energy, 107 (2016) 804-817. [http://dx.doi.org/10.1016/j.energy.2016.04.065]

[18] J.-P. Burochin, B. Vallet, M. Brédif, C. Mallet, T. Brosset, N. Paparoditis, Detecting blind building façades from highly overlapping wide angle aerial imagery, ISPRS Journal of Photogrammetry and Remote Sensing, 96 (2014) 193-209. [http://dx.doi.org/https://doi.org/10.1016/j.isprsjprs.2014.07.011]

[19] F. Duque-Pintor, M. Fernández-Gómez, A. Troncoso, F. Martínez-Álvarez, A New Methodology Based on Imbalanced Classification for Predicting Outliers in

Electricity Demand Time Series, Energies, 9 (2016) 752-752. [http://dx.doi.org/10.3390/en9090752]

[20] A. Capozzoli, F. Lauro, I. Khan, Fault detection analysis using data mining techniques for a cluster of smart office buildings, Expert Systems with Applications, 42 (2015) 4324-4338. [http://dx.doi.org/10.1016/j.eswa.2015.01.010]

[21] A. Ahmad, M. Khan, A. Paul, S. Din, M.M. Rathore, G. Jeon, G.S. Chio, Towards modeling and optimization of features selection in Big Data based social Internet of Things, Future Generation Computer Systems, (2017). [http://dx.doi.org/https://doi.org/10.1016/j.future.2017.09.028]

[22] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, Y. Agarwal, Sentinel: occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings, in: Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, ACM, Roma, Italy, 2013, pp. 1-14. [http://dx.doi.org/10.1145/2517351.2517370]

[23] J.H. Reif, Synthesis of parallel algorithms, Morgan Kaufmann Publishers Inc., 1993. [http://dl.acm.org/citation.cfm?id=562546]

[24] A. Ali, G.A. Shah, J. Arshad, Energy efficient techniques for M2M communication: A survey, Journal of Network and Computer Applications, 68 (2016) 42-55. [http://dx.doi.org/10.1016/j.jnca.2016.04.002]

[25] J.-M. Adamo, Data mining for association rules and sequential patterns: sequential and parallel algorithms, Springer Science & Business Media, 2012. [http://books.google.es/books?hl=es&lr=&id=CXPjBwAAQBAJ&oi=fnd&pg=PA49&dq=parallel+and+sequential+implementations+data+mining&ots=X9oC3aiqq0&sig=YqLrFUfswlReoT4p0mucYz-bORU]

[26] L.J. Eshelman, The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination, in: R. Gregory J.E (Ed.) Foundations of Genetic Algorithms, Elsevier, 1991, pp. 265-283. [http://dx.doi.org/10.1016/B978-0-08-050684-5.50020-3]

[27] D. Kanani, K. Shah, R.G. Vaishnav, Cloud Computing–Task Scheduling based on Modified CHC Algorithm, environment, 1 (2016) 2. [http://www.irjet.net/archives/V3/i7/IRJET-V3I791.pdf]

[28] J. Derrac, F. Chiclana, S. García, F. Herrera, Evolutionary fuzzy k-nearest neighbors algorithm using interval-valued fuzzy sets, Information Sciences, 329 (2016) 144-163. [http://dx.doi.org/10.1016/j.ins.2015.09.007]

[29] E. Parras-Gutierrez, V.M. Rivas, J.J. Merelo, A Radial Basis Function Neural Network-Based Coevolutionary Algorithm for Short-Term to Long-Term Time

Series Forecasting, in: K. Madani, A. Dourado, A. Rosa, J. Filipe, J. Kacprzyk (Eds.) Computational Intelligence: Revised and Selected Papers of the International Joint Conference, IJCCI 2013, Vilamoura, Portugal, September 20-22, 2013, Springer International Publishing, Cham, 2016, pp. 121-136. [http://dx.doi.org/10.1007/978-3-319-23392-5_7]

[30] V. Paramasivam, T.S. Yee, S.K. Dhillon, A.S. Sidhu, A methodological review of data mining techniques in predictive medicine: An application in hemodynamic prediction for abdominal aortic aneurysm disease, Biocybernetics and Biomedical Engineering, 34 (2014) 139-145. [http://dx.doi.org/10.1016/j.bbe.2014.03.003]

[31] F. Azuaje, Computational models for predicting drug responses in cancer research, Briefings in Bioinformatics, 18 (2017) 820-829. [http://dx.doi.org/10.1093/bib/bbw065]

[32] J. Fan, F. Han, H. Liu, Challenges of Big Data Analysis, National science review, 1 (2014) 293-314. [http://dx.doi.org/10.1093/nsr/nwt032]

[33] H.C. Jung, J.S. Kim, H. Heo, Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach, Energy and Buildings, 90 (2015) 76-84. [http://dx.doi.org/10.1016/j.enbuild.2014.12.029]

[34] C. Deb, F. Zhang, J. Yang, S.E. Lee, K.W. Shah, A review on time series forecasting techniques for building energy consumption, Renewable and Sustainable Energy Reviews, 74 (2017) 902-924. [http://dx.doi.org/10.1016/j.rser.2017.02.085]

[35] X. Qiu, P.N. Suganthan, G.A.J. Amaratunga, Ensemble incremental learning Random Vector Functional Link network for short-term electric load forecasting, Knowledge-Based Systems, 145 (2018) 182-196. [http://dx.doi.org/10.1016/j.knosys.2018.01.015]

[36] X. Qiu, Y. Ren, P.N. Suganthan, G.A.J. Amaratunga, Empirical Mode Decomposition based ensemble deep learning for load demand time series forecasting, Applied Soft Computing, 54 (2017) 246-255. [http://dx.doi.org/10.1016/j.asoc.2017.01.015]

[37] M. Benedetti, V. Cesarotti, V. Introna, J. Serranti, Energy consumption control automation using Artificial Neural Networks and adaptive algorithms: Proposal of a new methodology and case study, Appl Energ, 165 (2016) 60-71. [http://dx.doi.org/10.1016/j.apenergy.2015.12.066]

[38] E. Cadenas, W. Rivera, R. Campos-Amezcua, R. Cadenas, Wind speed forecasting using the NARX model, case: La Mata, Oaxaca, México, Neural Computing and

Applications, 27 (2016) 2417-2428. [http://dx.doi.org/10.1007/s00521-015-2012-y]

[39] Z. Xian, C. Ka Wing, Y. Xuesen, Z. Yangyang, Y. Kexin, W. Guibin, A comparison study on electric vehicle growth forecasting based on grey system theory and NAR neural network, in: 2016 IEEE International Conference on Smart Grid Communications (SmartGridComm), 2016, pp. 711-715. [http://dx.doi.org/10.1109/SmartGridComm.2016.7778845]

[40] A. Afram, F. Janabi-Sharifi, A.S. Fung, K. Raahemifar, Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: A state of the art review and case study of a residential HVAC system, Energy and Buildings, 141 (2017) 96-113. [http://dx.doi.org/10.1016/j.enbuild.2017.02.012]

[41] J. Wang, J. Wang, W. Fang, H. Niu, Financial time series prediction using elman recurrent random neural networks, Computational intelligence and neuroscience, 2016 (2016). [http://dx.doi.org/10.1155/2016/4742515]

[42] S. Kelo, S. Dudul, A wavelet Elman neural network for short-term electrical load prediction under the influence of temperature, International Journal of Electrical Power & Energy Systems, 43 (2012) 1063-1071. [http://dx.doi.org/10.1016/j.ijepes.2012.06.009]

[43] M. Elkaref, B. Bohnet, A Simple LSTM model for Transition-based Dependency Parsing, arXiv preprint (2017). [http://arxiv.org/pdf/1708.08959]

[44] C. Liang, W. Zhen, W. Gang, Application of LSTM Networks in Short-Term Power Load Forecasting Under the Deep Learning Framework, Electric Power Information and Communication Technology, 15 (2017) 8-11. [http://www.dlxxtx.com/CN/article/downloadArticleFile.do?attachType=PDF&id=1380]

[45] T. Linzen, E. Dupoux, Y. Goldberg, Assessing the ability of LSTMs to learn syntax-sensitive dependencies, arXiv preprint arXiv:1611.01368, (2016). [http://arxiv.org/pdf/1611.01368]

[46] K. Wang, C. Deng, J. Li, Y. Zhang, X. Li, M. Wu, Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network, Epidemiology & Infection, 145 (2017) 1118-1129. [http://dx.doi.org/10.1017/S0950268816003216]

[47] W. Wu, S. An, J. Guo, P. Guan, Y. Ren, L. Xia, B. Zhou, Application of nonlinear autoregressive neural network in predicting incidence tendency of hemorrhagic fever with renal syndrome, Zhonghua liu xing bing xue za zhi= Zhonghua

liuxingbingxue zazhi, 36 (2015) 1394-1396. [http://europepmc.org/abstract/med/26850398]

[48] A. Zavadskaya, Artificial Intelligence in Finance: Forecasting Stock Market Returns Using Artificial Neural Networks (Available on Internet), (2017). [http://helda.helsinki.fi/dhanken/bitstream/handle/123456789/170154/zavadskaya.pdf?sequence=1]

[49] A. Verma, I. Kaur, A. Kaur, Algorithmic approach to data mining and classification techniques, Indian Journal of Science and Technology, 9 (2016). [http://www.indjst.org/index.php/indjst/article/view/88874/71913]

[50] G. Düzenli, RFID card security for public transportation applications based on a novel neural network analysis of cardholder behavior characteristics, Turkish Journal of Electrical Engineering & Computer Sciences, 23 (2015) 1098-1110. [http://journals.tubitak.gov.tr/elektrik/issues/elk-15-23-4/elk-23-4-13-1306-96.pdf]

[51] C. Yang, H. Li, Y. Rezgui, I. Petri, B. Yuce, B. Chen, B. Jayan, High throughput computing based distributed genetic algorithm for building energy consumption optimization, Energy and Buildings, 76 (2014) 92-101. [http://dx.doi.org/10.1016/j.enbuild.2014.02.053]

[52] I. Petri, H. Li, Y. Rezgui, Y. Chunfeng, B. Yuce, B. Jayan, A modular optimisation model for reducing energy consumption in large scale building facilities, Renewable and Sustainable Energy Reviews, 38 (2014) 990-1002. [http://dx.doi.org/10.1016/j.rser.2014.07.044]

[53] M. Ibrahim, S. Jemei, G. Wimmer, D. Hissel, Nonlinear autoregressive neural network in an energy management strategy for battery/ultra-capacitor hybrid electrical vehicles, Electric Power Systems Research, 136 (2016) 262-269. [http://dx.doi.org/10.1016/j.epsr.2016.03.005]

[54] G.I. Nagy, G. Barta, S. Kazi, G. Borbély, G. Simon, GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach, International Journal of Forecasting, 32 (2016) 1087-1093. [http://dx.doi.org/10.1016/j.ijforecast.2015.11.013]

[55] J.L. Elman, Finding Structure in Time, Cognitive Science, 14 (1990) 179-211. [http://dx.doi.org/10.1207/s15516709cog1402_1]

[56] A. Kose, E. Petlenkov, System identification models and using neural networks for Ground Source Heat Pump with Ground Temperature Modeling, in: 2016 International Joint Conference on Neural Networks (IJCNN), 2016, pp. 2850-2855. [http://dx.doi.org/10.1109/IJCNN.2016.7727559]

[57] G. Bao, Q. Lin, D. Gong, H. Shao, Hybrid Short-term Load Forecasting Using Principal Component Analysis and MEA-Elman Network, in: D.-S. Huang, K. Han, A. Hussain (Eds.) Intelligent Computing Methodologies: 12th International Conference, ICIC 2016, Lanzhou, China, August 2-5, 2016, Proceedings, Part III, Springer International Publishing, Cham, 2016, pp. 671-683. [http://dx.doi.org/10.1007/978-3-319-42297-8_62]

[58] S. Qin, J. Wang, J. Wu, G. Zhao, A hybrid model based on smooth transition periodic autoregressive and Elman artificial neural network for wind speed forecasting of the Hebei region in China, International Journal of Green Energy, 13 (2016) 595-607. [http://dx.doi.org/10.1080/15435075.2014.961462]

[59] C.-T. Chu, C.-H. Chang, T.-J. Chang, J.-X. Liao, Elman neural network identify elders fall signal base on second-order train method, in: Next Generation Electronics (ISNE), 2017 6th International Symposium on, IEEE, 2017, pp. 1-4. [http://dx.doi.org/10.1109/ISNE.2017.7968722]

[60] P. Li, Y. Li, Q. Xiong, Y. Chai, Y. Zhang, Application of a hybrid quantized Elman neural network in short-term load forecasting, International Journal of Electrical Power & Energy Systems, 55 (2014) 749-759. [http://dx.doi.org/10.1016/j.ijepes.2013.10.020]

[61] S. Ekici, S. Yildirim, M. Poyraz, A transmission line fault locator based on Elman recurrent networks, Applied Soft Computing, 9 (2009) 341-347. [http://dx.doi.org/10.1016/j.asoc.2008.04.011]

[62] M. Cuéllar, M. Delgado, M. Pegalajar, Multiobjective evolutionary optimization for Elman recurrent neural networks, applied to time series prediction, Fuzzy Economic Review, 10 (2005) 17-33. [http://search.proquest.com/docview/228956321]

[63] M. Delgado, M.C. Pegalajar, M.P. Cuéllar, Memetic evolutionary training for recurrent neural networks: An application to time-series prediction, Expert Systems, 23 (2006) 99-114. [http://dx.doi.org/10.1111/j.1468-0394.2006.00327.x]

[64] V. Singh, K. Tiwari, Prediction of GreenHouse Micro-Climate using Artificial Neural Network, Applied Ecology and Environmental Research, 15 (2017) 767-778. [http://www.aloki.hu/pdf/1501_767778.pdf]

[65] K. Greff, R.K. Srivastava, J. Koutník, B.R. Steunebrink, J. Schmidhuber, LSTM: A Search Space Odyssey, IEEE Transactions on Neural Networks and Learning Systems, 28 (2017) 2222-2232. [http://dx.doi.org/10.1109/TNNLS.2016.2582924]

[66] F.J. Ordóñez, D. Roggen, Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition, Sensors, 16 (2016) 115. [http://dx.doi.org/10.3390/s16010115]

[67] X. Lu, Y. Chen, X. Li, Hierarchical Recurrent Neural Hashing for Image Retrieval With Hierarchical Convolutional Features, IEEE Transactions on Image Processing, 27 (2018) 106-120. [http://dx.doi.org/10.1109/TIP.2017.2755766]

[68] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation, 9 (1997) 1735-1780. [http://www.mitpressjournals.org/doi/pdfplus/10.1162/neco.1997.9.8.1735]

[69] D.L. Marino, K. Amarasinghe, M. Manic, Building energy load forecasting using Deep Neural Networks, in: IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, 2016, pp. 7046-7051. [http://dx.doi.org/10.1109/IECON.2016.7793413]

[70] K. He, J. Sun, Convolutional neural networks at constrained time cost, in: Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on, IEEE, 2015, pp. 5353-5360. [http://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/He_Convolutional_Neural_Networks_2015_CVPR_paper.pdf]

[71] H.G. Lee, C.Y. Yi, D.E. Lee, D. Arditi, An Advanced Stochastic Time-Cost Tradeoff Analysis Based on a CPM-Guided Genetic Algorithm, Computer-Aided Civil and Infrastructure Engineering, 30 (2015) 824-842. [http://dx.doi.org/10.1111/mice.12148]

[72] T. Blickle, L. Thiele, A comparison of selection schemes used in genetic algorithms, (1995). [http://pdfs.semanticscholar.org/fef8/1135f587851f19fe515cb8eb3812e3706b27.pdf]

[73] F. Herrera, M. Lozano, J.L. Verdegay, Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis, Artificial Intelligence Review, 12 (1998) 265-319. [http://dx.doi.org/10.1023/a:1006504901164]

[74] H. Junninen, H. Niska, K. Tuppurainen, J. Ruuskanen, M. Kolehmainen, Methods for imputation of missing values in air quality data sets, Atmospheric Environment, 38 (2004) 2895-2907. [http://dx.doi.org/10.1016/j.atmosenv.2004.02.026]

6.4 A case study on understanding energy consumption through prediction and visualisation (VIMOEN)

Referencia
Ruiz, L. G. B., Pegalajar, M. C., Molina-Solana, M., Guo Y. (2019). A case study on understanding energy consumption through prediction and visualisation (VIMOEN).
Estado
En revisión.

A case study on understanding energy consumption through prediction and visualisation (VIMOEN)

L.G.B. Ruiz [a,*], M.C. Pegalajar [a], M. Molina-Solana [b], Yi-Ke Guo [b]

[a] Department of Computer Science and Artificial Intelligence, University of Granada.

[b] Data Science Institute, Imperial College London, UK.

bacaruiz@ugr.es* / mcarmen@decsai.ugr.es / m.molina-solana@imperial.ac.uk / y.guo@imperial.ac.uk

Abstract. Energy efficiency has emerged as an overarching concern due to the high pollution and cost associated with operating heating ventilation and air-conditioning systems in buildings. Thus, energy monitoring has become one of the most important research topics with global impacts. This, along with energy forecasting represent a very decisive task for energy efficiency. The goal of this study is divided into two parts. First, to provide a methodology to predict the energy usage every hour with the goal of deciding which Machine Learning is the best approach: Trees, Support Vector Machine or Neural Networks. Since the UGR lacks a tool to properly monitoring those data, then, the second aim is to propose an intelligent system to visualize and to use those models in order to predict energy consumption in real time. For this end, we will design VIMOEN (VIsual MOnitoring of ENergy), a web-based application (using Mapbox) whose objective is to provide not only visual information about the energy consumption of a set of geographically-distributed buildings but also expected expenditures in the near future. The system has been designed to be easy-to-use and intuitive for any non-expert users. Thus, VIMOEN is intended to make easier the monitoring of a large amount of buildings and to anticipate changes in consumption. The system has been validated on data coming from buildings of University of Granada.

Keywords: *energy efficiency; energy forecasting; visualization; Mapbox; energy monitoring*

1. Introduction

Buildings are known to be among the largest consumers of energy at global level (International Energy Agency, 2013). In the USA (Pérez-Lombard, Ortiz & Pout, 2008) HVAC systems account for 20% of the global consumption, and the US Department of Energy calculates that buildings are responsible of 70% of the electricity in the country; the amount of energy consumed in the European Union reaches more than 50%, and other countries such as China or Iran have increased their consumption in more than 10% in the past 20 years (Chen, Shi, Shen, Huang & Wu, 2019). A considerable portion of this energy consumption in buildings is linked to heating, ventilation and air conditioning systems —HVAC— which oversee maintaining comfort for the building's occupants. Typically, these HVAC systems run on rigid schedules and preloaded rules and do not use any external and dynamic information in order to optimize energy efficiency (Gelazanskas & Gamage, 2014). Besides, the growth rate of population and the necessity of energy conservation in this sector have become a significant concern for many governments in the world. Increasing levels of energy consumption are correlated not only with the price of the energy of each country, but also with environmental pollution and its negative impact on health. Recent studies (Santamouris & Kolokotsa, 2015) maintain that global climate change has seriously increased the frequency of extreme weather conditions, bringing with it severe energy impact, high temperature, extreme climatic conditions and pollution affecting the environmental comfort conditions and, what is more important, health. Hence, the need for prompt action to prevent imminent impact of climatic change and aggravating overheating on the energy consumption in buildings discloses the importance and seriousness of the problem.

In order to achieve more efficient buildings, it is necessary to understand how buildings consume energy (Lopes, Antunes & Martins, 2012) and have suitable monitoring systems to reduce energy waste and generate energy savings. Granderson and Lin (Granderson & Lin, 2016) indeed pointed out that building energy information systems are a powerful technology to monitor and analyse consumption which leads to considerable energy and monetary savings. Nevertheless, appropriately handling the vast amount of available data is not a trivial undertaking because of heterogeneity in data provided by the sensors and smart devices.

In fact, several building energy management systems have been developed as analysis software for data acquisition and monitoring in diverse scenes. For instance, Sarma et al. (Sarma, Singh & Bezboruah, 2018) presented a design of a low-cost data acquisition system for data monitoring, data storing and plotting of live streaming data; Tae-Keun et al. (Oh, Lee, Park, Cha & Park, 2018) proposed a three-dimensional visualization solution for building energy management targeting recommendations to enhance the energy efficiency obtained from the energy consumption; and Pahl et al. (Pahl, Goodhew, Boomsma & Sheppard, 2016) aimed at energy visualization, in which they investigate how to reduce energy consumption in buildings by combining psychology principles and intelligent techniques.

Many authors elaborate designs based on graphs, tables and figures whose purpose is to depict data in a more organized manner, but rare are the works that propose interactive and visual tools for this purpose. A notable exception is (Murshed, Al-Hyari, Wendel & Ansart, 2018) where a design of a web application for smart city visualization was implemented. Furthermore, in almost all scientific studies the solution is focused on just one building, not on a whole set of facilities distributed in diverse areas.

In addition to those, there are many solutions in literature for improving energy efficiency and many intelligent algorithms for processing and managing the energy consumption data. In nearly all instances, data are employed without the need for being visualized, using consumption data only for this purpose. The main problem here arises when non-expert users attempt to use those technologies. The task becomes an arduous and laborious issue, due to the learning difficulties to understand panels with too much information, and the relevant information seems to have less importance; making it very difficult to control and manage the energy. Thus, the basic requirement is the need to delivering an intuitive and easy-to-use software to the users and operators in order to ensure they understand what happens to the energy of their buildings. It should be noted that, using a crowded dashboard of information, or using algorithms and intelligent systems only are non-perfect solutions because its limited interpretability. The proposal we describe in this work precisely pursues this objective to improve current solutions by simplifying and clarifying the knowledge to assist users in making faster diagnoses, which can improve decision making and efficiency.

Hitherto, very little research has been done in energy efficiency and visualization incorporating new intelligent technologies. Most recent studies are focused on applying data mining techniques (Deb, Zhang, Yang, Lee & Shah, 2017; Ruiz, Cuellar, Delgado & Pegalajar, 2016; Ruiz, Rueda, Cuéllar & Pegalajar, 2018) on the one hand, and data visualization (Huacón & Pelegrin, 2017; Rodrigues, et al., 2017) on the other, without combining both approaches. Thus, to achieve an adequate monitoring of the consumption in a building complex and adding intelligent techniques for effectively realizing energy savings, this work presents a web-based application for monitoring the energy consumption of distributed buildings following a more visual strategy and thus more user-oriented.

Information visualisation has been extensively demonstrated as a viable tool for improved data understanding, as it increases concentration and makes information more attractive, hence supporting more informed decision-making (Oh et al., 2018). For this reason, we have developed an innovative application using recent web technologies and energy consumption forecasting methods to predict the waste of energy and improve the energy efficiency in a set of buildings located in different geographic zones. We adopted the Mapbox GL JS API (Mapbox & LLC, 2018) to render an interactive map with buildings' locations and we used the best Machine Learning model to predict consumption.

Due to the high difficulty of accurately forecasting energy consumption, diverse approaches can be found in literature to solve this problem. Some of those models have recently gained considerable popularity: linear regression (Fumo & Rafe Biswas, 2015) regression trees (Nagy, Barta, Kazi, Borbély & Simon, 2016), autoregressive moving average and grey models (Yuan, Liu & Fang, 2016), fuzzy-based models (Davlea & Teodorescu, 2016), support vector machine (Jung, Kim & Heo, 2015), neural network (Singh & Tiwari, 2017) among many other hybridizations. Historical data of the energy consumption of the University of Granada will be used in the comparative study. Most of the buildings present more than five years of data. This data was acquired from sensors installed in buildings.

Our first goal is to compare those techniques so as to obtain the most accurate approach for bringing it into the visualization system.

The field of energy visualization has a huge potential which have recently begun to be exploited and can make a significant contribution to the challenges surrounding energy efficiency, global warming and exploitation of energy resources. In addition to this, it is important to mention that the UGR does not currently have a support tool for monitoring or access adequately to energy consumption data. Accordingly, the second aim of this paper is to fill this gap and solve this problem by providing a unique software which combines visualization and soft computing techniques to provide an intelligent system to ensure proper energy monitoring. In this paper, we focus on developing an application for centralized monitoring of distributed buildings. Our second goal in this paper is to procure a visual framework to manage the current data of energy consumption, and to know at a glance changes in the building behaviours and to identify possible causes of low performance by using forecasting models. The tool has been tested with data and buildings from University of Granada (UGR). 25 buildings have been considered, in 7 campus located in south Europe and North Africa. VIMOEN's visualization platform will enable an accurate energy consumption management and will allow anticipating future scenarios in the near future as well as providing updated reports about buildings.

Therefore, the overall goal of this work is to design and implement an energy consumption visualization tool that provides energy information on the past, present and future state of a distributed group of buildings spread geographically throughout diverse areas. The system has been designed to ease the use and to improve the interpretability of the energy information collected from buildings in order to aid any user who needs to know in real-time how their buildings are consuming, avoiding overstuffed dashboards which often result complex and difficult to interpret. Additionally, adding to it intelligent forecasting models to anticipate and manage consumption changes and to make effective planning and decision making about energy expenditure.

The rest of the paper is structured as follows: Section 2 introduces the forecasting techniques for energy consumption. Section 3 describes data used in this study. Section 4 presents the methodology carried out, data pre-processing and VIMOEN, along with its different parts. Section 5 gathers all outcomes related to energy forecasting and demonstrates the usage of VIMOEN. The last section concludes with some comments and proposals for future works.

2. Forecasting techniques

This section contains a short presentation of the three different types of machine learning techniques used to predict energy consumption: trees, support vector machine models and artificial neural networks.

2.1. Regression Tree

Regression Tree (RT) is one of the simplest model among all the machine learning techniques. Tree-based models are commonly used in both classification and regression problems. Although, in the first case, it is called Decision Tree. Its tree-structured form is based on binary recursive partitioning of a dataset. This process divides observations into different groups and are utilized to decide the structure of the tree (Xu, Watanachaturaporn, Varshney & Arora, 2005). The main advantage of RT over classical statistical methods is that RT does not make a priori assumptions concerning the dependencies between input and output variables. Furthermore, thanks to its configuration RT allows the modelling of nonlinearities in data. CART is the standard nonparametric procedure to perform the predictor selection. In this study three varieties of RT were used: the simple RT, Boosted Regression Tree (BoRT) and Bagged Regression Tree (BaRT) (Prasad, Iverson & Liaw, 2006). BoRT and BaRT are an ensemble approach which combines the results of several RT to predict the target output. The main difference between those two approaches is that, while BaRT creates an aggregated model with less variance, BoRT focus new predictors on observations that other get wrong.

2.2. Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning method which was created to classify a dataset into two separable classes. To do so, SVM finds the best hyperplane by maximizing the distance between the two classes. However, SVM has recently gained popularity to solve regression problems and have been shown to be superior to traditional empirical risk minimization approaches. SVM turns the input space into a high-dimension space and models data according to a kernel function (Jung et al., 2015). In this work two versions of the SVM regression model were implemented. The first is the Linear SVM (L-SVM) model which maps data into a linear function. This model presents a high interpretability and a low flexibility. In contrast, the second model is the Coarse Gaussian SVM (CG-SVM) whose kernel is the radial basis function and kernel scale is set to $4 \cdot \sqrt{P}$, being P the number of predictors. This model has a hard interpretability and a rigid response function. The main advantage of SVM is finding the best and unique solution to minimize the objective function.

2.3. Artificial Neural Networks

Lastly, the Artificial Neural Networks (ANN) are a particular machine learning inspired by brain connections. ANNs are non-mathematical parametric models capable of modelling non-linear dependencies between input and output variables. In addition to the previous models, ANN are widely used in both classification (Duque-Pintor, Fernández-Gómez, Troncoso & Martínez-Álvarez, 2016) and regression (Chae, Horesh, Hwang & Lee, 2016) problems. Due to its positive results in numerous problems, ANN has become as the most popular method. Many types ANN can be found in literature. The simplest one is the Multi-Layer Perceptron (MLP) which has been utilized in this study to predict energy consumption. Also, since the need of modelling time dependencies of data to predict future consumption usage, the Elman Neural Network (ENN) has been utilized as well, as this introduces the concept of memory. To this end, ENN uses positive feedback to fit this memory in the model. This feature provides a better modelling of time-series with historical dependencies.

3. Dataset

As a case study to test the feasibility of the forecasting models and VIMOEN, we will work with data collected from University of Granada (UGR). Due to Data Protection Laws and security issues, we are unable to display or share the actual data of energy consumption; this fact however does not affect the development and achievement of this research. UGR is located and distributed in three different cities: Granada (south of Spain, Europe), Ceuta and Melilla (north of Africa). UGR is divided into 7 campuses

(*Centro, Fuentenueva, Cartuja, Ciencias de la Salud, Aynadamar, Ceuta* and *Melilla*) each one of them comprising different types of buildings, and consequently, different sensors technologies and energy consumptions need to be considered and dealt with; e.g., teaching units, faculties, schools, departments and training centres. An additional difficulty arises from the disparity of reporting software and different information registers, as each building was commissioned independently. Luckily, most of them register energy consumption and temperature.

As noted above, UGR has recently installed technologies to support building monitoring; and those pieces of software are designed to collect and transform energy data from all sensors and meters in a usable form. The data is stored in a database which gathers monitored information such as consumption, energy and power, as well as other external data such as temperature and humidity, which do not come from the installed sensors but from third party's services. This smart management system captures all this information in real-time and is used by the University for monitoring and analysing buildings' consumption. However, the building automation system is in charge of storing raw data as it comes from sensors. This makes it difficult to work with such information. For this reason, the treatment of those data is essential and must be preprocessed. Fig. 1a illustrates an example of the data from meters. As one may see, it is very hard to obtain any information as the consumption is stored considering how much energy has been used up to now. This is why, in Fig. 1a, a rising trend is depicted, until it reaches breaking point where the meter resets its counter. However, once raw consumption data are processed, they present the form shown in Fig. 1b which details a more usable way of the same information, and a particular behavior may be appreciated. Besides, this treatment is vital due to possible problems in data, such as, problems during the data transfer, fault detection, broken devices or system failure.

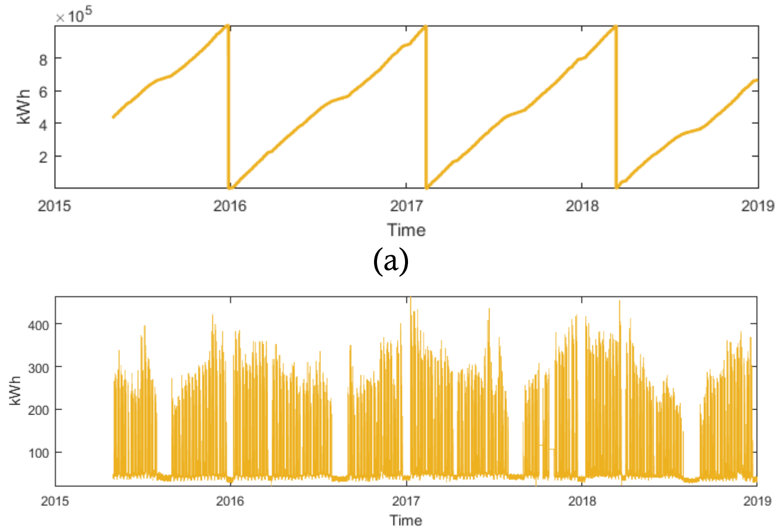


Fig. 1. Example of (a) the raw data extracted from a building's meter (b) and the transformed and processed energy consumption in an hourly time scale.

4. Methodology

This study has two main goals or stages. The first one is to find the best forecasting model in order to predict the energy consumption in buildings. The second one is to utilize past, present and future building's information so as to help improve their consumption monitoring.

4.1. Data-related treatment

In order to take up the first problem, a retrieving data procedure must be designed to extract energy consumption from buildings. As mentioned above, data present several problems and its form when stored by meters is not usable or readable directly (see Fig. 1a). Data are preprocessed through a noise treatment and data transformation. The first procedure is carried out with two solutions: 1) Moving-average filter to remove outliers (Kayacan, Ulutas & Kaynak, 2010); 2) linear interpolation using the neighboring grid values of the consumption to fill missing data (Andiojaya & Demirhan, 2019). The data extracted from the electric meters are stored as a historical time-series which represents the accumulated amount of energy used to date. In other words, one record of the raw data denotes the aggregate sum of all preceding consumption until at that time. Hence, the second operation involves changing the raw data into an hourly time scale. In this way, one sample of the data is now the consumption spent that hour.

Once that was done (see Fig. 1b), we address forecasting aim by testing several machine learning algorithms: RT, SVM and ANN, along with some variants of these models. We will work with historical data, as a result, a time-series y corresponding to a particular building will be modelled following the next equation:

$$y(t+1) = m\big(y(t), y(t-1), y(t-2), \dots, y(t-p)\big) + \epsilon(t+1) \quad \textbf{(1)}$$

In this way, the future value $t+1$ depends on the p past values, and the specific model m will adjust those data. The error of the forecasting model at time $t+1$ is represented by $\epsilon(t+1)$.

Data are also normalized in range $[0, 1]$ in order to guarantee that the models do not give more weight to those attributes with higher range in consumption by using the next equation:

$$y = \frac{y - y_{\min}}{y_{max} - y_{min}} \quad (2)$$

Besides, we applied Root Mean Squared Error (RMSE) to check the accuracy of the models:

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \big(D(t) - O(t)\big)^2} \quad (3)$$

Where $D(t)$ is the desired response from the forecasting model, $O(t)$ is the output at instant t and N is the number of instances.

Fig. 2 illustrates the process flow to train the predictive models. As a summary, this figure gathers all the operations that are carried out. Firstly, the raw data are collected from diverse sensors installed in the buildings. This information is pre-processed in order to allow us to work with the provided data in a useful way. In this stage, data granularity is set according to the established time scale, in this case, hourly. Once the data are transformed in the desired form, they should pass the missing and noising values procedures previously mentioned. Thus, wrong data and gaps in the consumption are treated. As soon as the data are considered clean, they are split into two sets: training and test. The first one is employed to train the model, and the second one is used to validate and confirm that the forecasting model has learnt satisfactorily to predict consumption. Once the model is considered validated, it is stored to later predict the energy consumption when required.

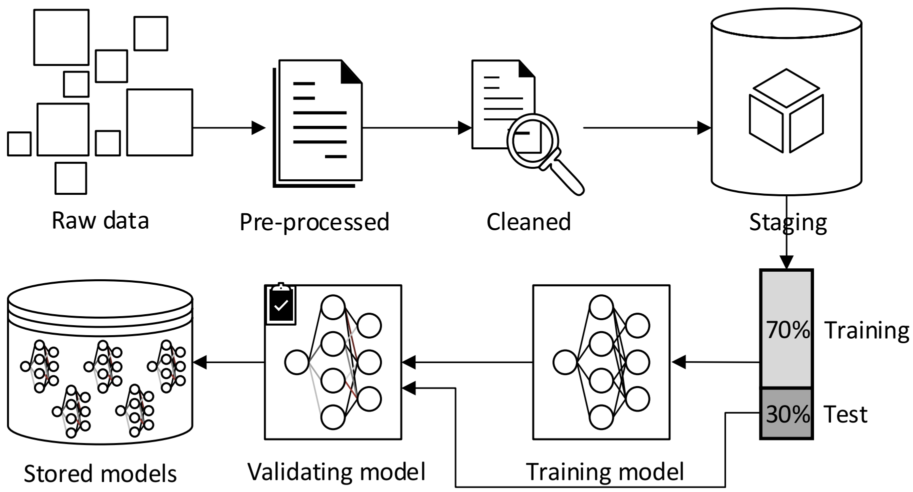


Fig. 2. Flowchart of the training process of forecasting models. This follows the usual procedures and best-practices in Machine Learning.

4.2. Proposed visualization system

VIMOEN is composed by three main components: database, forecasting models and visualization module. The system has been developed using the Mapbox GL JS library to render interactive maps, ANN to predict the consumption of the buildings and data from the facilities of the UGR to create the design of the system as illustrated in Fig. 3. The first part represents the origin of the data. All information is extracted from the buildings through the automation system which collects the data from the physical world by using the sensing technologies. Each sensor provides different information about the buildings. All the possible context information of the buildings must be pre-processed and transformed before it can be used to provide useful knowledge.

Consequently, the next step is to join and process such information in order to properly handle and manipulate it. Thus, the database gathers the remodelled data of the hourly energy consumption or the temperature, in addition to other interesting details, e.g., social networks, corporate website respective building and naturally the geographic information hereof. All that knowledge will be used to build the final system, but firstly data on energy consumption is employed by the predictive models to forecast forthcoming consumption. Note that other knowledge apart from the energy consumption, such as the geolocation information, web portal or social networks of the buildings was taken from UGR's public website University to complete the information.

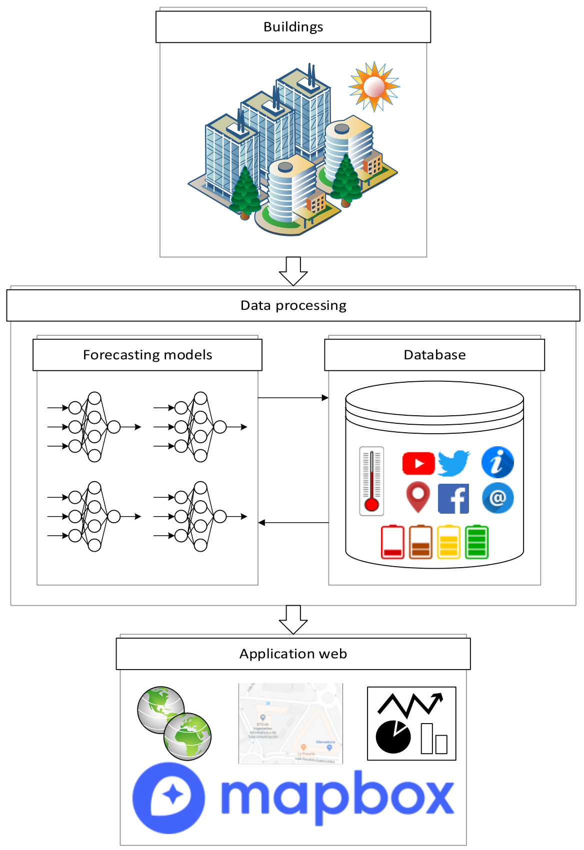


Fig. 3. Overall scheme of VIMOEN.

VIMOEN has been designed to manipulate the information of the buildings and to depict it in the form of knowledge. And thus, to clearly visualize the energy consumption of the whole distributed institution in a centralized system. This component collects the buildings' data and links them for visualization in client-side by using web browser. It also incorporates the forecasting models already trained to include the predicted consumption of the buildings. For visualization purpose, the Mapbox's JavaScript library is used here to render the maps. Geographical data about the location of UGR's buildings was available for our test.

An illustration of the proposal layout for the energy visualization can be found in Fig. 4. The first layer is provided by the Mapbox API with the street layer and all the system is based on it. The map layer includes methods and properties which allow change the appearance of the map and enable users to interact with it. The second layer includes the location of all buildings. The buildings' geometry is determined with a two-value coordinates, the longitude and the latitude, both measured in degrees. Each point is endowed with a popup component in which the real-time consumption is depicted. Furthermore, each point offers useful information at a glance and provides easy access to historical consumption (as described in the next section).

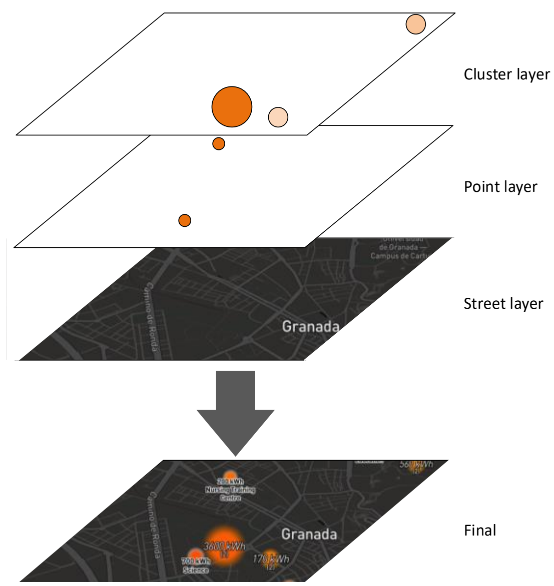
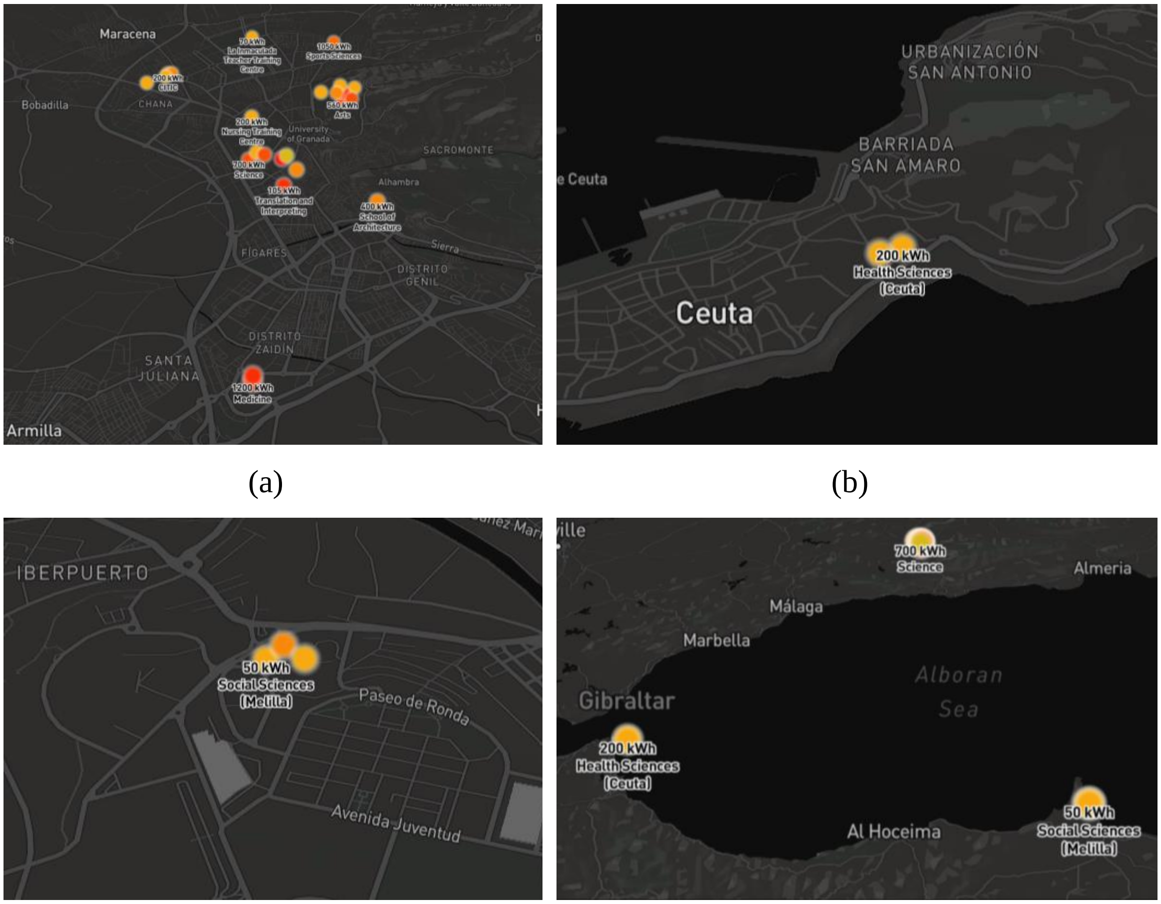


Fig. 4. Layer structure of the map visualization using Mapbox.

The last layer is designed to give an overview of the consumption in an aggregated way. Fig. 5a shows how it would be like if points were not clustered. In view of the low number of facilities in Ceuta and Melilla, buildings are relatively discernible in Fig. 5b and Fig. 5c. Nonetheless, given the amount of buildings located together, a fair number of them are overlapped in Fig. 5d. For all these reasons, the last layer has been configured to address these visualization problems. The next section will exemplify the final representation in detail.



(a) (b)

(c) (d)

Fig. 5. Overview of the buildings at different zoom levels. A) if no clustered is performed in the city of Granada with 24 buildings, in (b) the city of Ceuta with 2 facilities and in (c) Melilla with 3 Universities. In (d) the whole institution is depicted with all its buildings

5. Results

The original set of data is made up of many buildings. As a consequence, a set of representative buildings were chosen to carry out all the experiments. Besides, to prevent an excessive amount of results, intermediate results will be skipped. Cross-validation were used to validate models. Each experiment was repeated 10 times and the average of those executions was taken in order to prevent biased results.

The first model, RT, was set with the next parameters: Standard CART as the algorithm used to select the best split predictor, the minimum number of branch node observations will be 10, 36 will be the minimum leaf size, the prune procedure enabled to estimate optimal sequence of pruned subtrees. Bagged and Boosted RT will be trained with 30 learners whose minimum leaf size will be 8 observations. Secondly, support vector machine methods were set using two different kernel functions: Linear and Gaussian. Hyper-parameters were optimized by following the Bull's method (Bull, 2011). Thirdly, the number of hidden layers of the ANN were considered as one layer with 10 neurons in both cases. ENN also will use 10 feedback delays in the *context layer.* The learning algorithm will be Levenberg-Marquardt.

In order to evaluate the performance of the seven techniques, datasets are split into training set and testing set. Each model will be trained with 70% of the energy consumption data and the remaining 30% will be used to measure the goodness of the fit.

According to all these directions, Table 1 was built. This table shows the Root Mean Squared Error for all our models tested: Regression Tree (RT), Bagged Regression Tree (BaRT), Boosted Regression Tree (BoRT), Linear Support Vector Machine (L-SVM), Coarse Support Vector Machine (C-SVM), Multi-layer Perceptron (MLP) and Elman Neural Network (ENN). As mention above, eight buildings were chosen as representative examples of the whole dataset. The average RMSE of 10 executions is presented in Table 1. In addition to this, ten different models where designed so as to predict the following 10 hours. In this way, each column illustrates the RMSE obtained to predict the energy consumption within f hours, i.e., the first column shows the RMSE to predict the next hour, the second column shows the RMSE to predict the consumption within two hours, and so on.

At first glance, we can observe that all the models present a similar behavior in terms of performance except the four method, L-SVM. In the latter case, L-SVM seems to be the worst option to model this kind of data as it is, by far, the model with higher error in almost all cases. Only with the fifth building and predicting the $t + 1$ and $t + 2$ values it reaches a slightly lower error than its other version, C-SVM. On the basis of these results there appears to be a bad option to choose SVM as a forecasting method on this problem. Thus, the last place in the ranking is for L-SVM (7^{th}), followed by C-SVM (6^{th}). The fifth position in the ranking is among RG, BoRT and MLP. Interestingly, MLP better predicts in the near future, as it achieves better performance in case $f = 1$ and $f = 2$. However, from $f = 3$, MLP gets the 3^{rd} worst predictions in more than half of the cases (5^{th}). Besides, RT which is much simpler model than MLP accomplishes the 4^{th} position in the ranking in all instances except for the two first hours. The third position is reached by MLP only when predicting the next hour. In the rest of the cases this position is mainly held by BoRT (3^{rd}), followed by BaRT (2^{nd}) which attains better results on the order of a few hundredths. Note that MLP gets even the second best alternative for predicting in a few instances, such as B8 in all its prediction windows. Finally, when looking more in detail this table, one may see that the undisputed best model is ENN. Only 1 out of the 80 trials ENN gets the second best error as shown in building 6, $t + 1$. In this case, BaRT achieves an error of 0.0004 better than ENN. In spite of this, ENN predicts, by far, all the buildings and time windows. In fact, ENN has an average RMSE of 25% better than the rest of the models, and in the best case it gets up to 45%.

The best results obtained are summarized in Table 2. As has just been commented, ENN is the best approach to model the energy consumption in this case study. When looking in detail this table, one may observe that errors increase as the forecasting time step dilates. To better observe this particular behavior, Fig. 6 shows the fact of the RMSE increase while the values to be predicted are further in time. In the case of ENN this increase is slightly softer if compared to all other models.

Table 1. Root Mean Squared Error for Regression Tree, Bagged Tree, Boosted Tree, Linear Support Vector Machine, Coarse Support Vector Machine, Multi-layer Perceptron and Elman Neural Network; with a prediction window from 1 to 10 hours ahead.

Test	Forecasting $(t+f)$									
	1	2	3	4	5	6	7	8	9	10
Building 1										
RT	0.042423	0.054863	0.063985	0.070398	0.075528	0.078756	0.078588	0.079099	0.078691	0.079696
BaRT	0.037520	0.049720	0.057712	0.064939	0.070922	0.075486	0.075698	0.075596	0.075504	0.075297
BoRT	0.040604	0.052857	0.061667	0.068541	0.072794	0.075675	0.075981	0.076402	0.075726	0.075921
L-SVM	0.127427	0.178221	0.240546	0.217467	0.309545	0.443688	0.494352	0.465635	0.427360	0.528536
C-SVM	0.094915	0.105127	0.111173	0.118738	0.124602	0.124382	0.124944	0.125287	0.124864	0.124389
MLP	0.039017	0.058152	0.069397	0.078527	0.085269	0.088017	0.089041	0.089225	0.088262	0.088203
ENN	**0.034165**	**0.045475**	**0.051615**	**0.055554**	**0.056859**	**0.057094**	**0.059120**	**0.058332**	**0.059158**	**0.058138**
Building 2										
RT	0.042859	0.062125	0.076290	0.087660	0.098379	0.106283	0.110490	0.112531	0.114845	0.118492
BaRT	0.038226	0.056269	0.068648	0.079789	0.090390	0.097477	0.102186	0.105571	0.107050	0.109130
BoRT	0.044885	0.065514	0.078284	0.089700	0.099833	0.106933	0.110984	0.113611	0.116065	0.118391
L-SVM	0.102332	0.167063	0.212275	0.221770	0.266588	0.278696	0.310133	0.362435	0.328193	0.266761
C-SVM	0.054078	0.101398	0.125119	0.157197	0.178748	0.204213	0.218686	0.212897	0.222557	0.211742
MLP	0.035340	0.053955	0.067195	0.079754	0.089908	0.097028	0.102150	0.105771	0.106879	0.110309
ENN	**0.031797**	**0.044207**	**0.051942**	**0.057654**	**0.062532**	**0.063908**	**0.066410**	**0.066248**	**0.067870**	**0.067452**
Building 3										
RT	0.031547	0.049398	0.061997	0.073760	0.081238	0.089481	0.093897	0.097338	0.100127	0.102717
BaRT	0.026183	0.040338	0.051118	0.060697	0.067733	0.075033	0.079476	0.083319	0.085435	0.088874
BoRT	0.031496	0.046650	0.058416	0.066898	0.075262	0.082133	0.087292	0.090801	0.093283	0.096594
L-SVM	0.084617	0.128474	0.198676	0.259834	0.441548	0.393710	0.435010	0.510033	0.483196	0.481687
C-SVM	0.040866	0.075134	0.105658	0.148788	0.177875	0.232809	0.243370	0.253713	0.266257	0.278848
MLP	0.028075	0.045483	0.061151	0.074584	0.084634	0.094050	0.101026	0.105667	0.108644	0.111685
ENN	**0.023489**	**0.034806**	**0.042175**	**0.048729**	**0.051864**	**0.055996**	**0.057125**	**0.061281**	**0.059965**	**0.060444**
Building 4										
RT	0.032956	0.051007	0.063292	0.072987	0.080316	0.086200	0.091087	0.093816	0.095955	0.099187
BaRT	0.028625	0.041943	0.052018	0.061827	0.069075	0.075140	0.079385	0.081742	0.084069	0.087188
BoRT	0.035509	0.050631	0.061390	0.069367	0.076878	0.083751	0.087000	0.089930	0.092027	0.095168
L-SVM	0.103108	0.140093	0.252552	0.234440	0.327257	0.347307	0.364249	0.421843	0.351986	0.443102
C-SVM	0.053955	0.094886	0.134043	0.164638	0.173981	0.185513	0.198803	0.217501	0.218112	0.224203
MLP	0.032587	0.050652	0.065593	0.076973	0.086439	0.094264	0.098363	0.101248	0.102968	0.106977
ENN	**0.028453**	**0.041269**	**0.049153**	**0.053318**	**0.056964**	**0.059172**	**0.061813**	**0.062787**	**0.062372**	**0.063499**
Building 5										
RT	0.031764	0.042542	0.050314	0.055847	0.059425	0.062154	0.064093	0.065677	0.067083	0.068476
BaRT	0.028353	0.037966	0.044599	0.049644	0.053220	0.055963	0.057586	0.059398	0.061064	0.061911
BoRT	0.033847	0.042640	0.048179	0.052707	0.055439	0.057752	0.059338	0.060393	0.062129	0.062896
L-SVM	0.056076	0.072446	0.100790	0.110914	0.115419	0.131269	0.137492	0.122878	0.129351	0.146875
C-SVM	0.059594	0.079173	0.088712	0.099861	0.104707	0.106620	0.104460	0.107770	0.111743	0.098357
MLP	0.028927	0.041099	0.048657	0.055133	0.059068	0.062300	0.064611	0.066456	0.067604	0.068328
ENN	**0.026887**	**0.033301**	**0.035983**	**0.038896**	**0.040655**	**0.041526**	**0.041912**	**0.042513**	**0.043044**	**0.043061**
Building 6										
RT	0.027749	0.040425	0.048876	0.055236	0.060473	0.067483	0.074032	0.076152	0.078286	0.078564
BaRT	**0.022552**	0.032892	0.040235	0.045656	0.051210	0.056968	0.062204	0.065652	0.066764	0.068371
BoRT	0.026224	0.037051	0.043347	0.049336	0.055228	0.062078	0.067480	0.069971	0.071582	0.071924
L-SVM	0.125184	0.140588	0.149961	0.204499	0.289259	0.333380	0.358899	0.357593	0.366183	0.331850
C-SVM	0.072624	0.111474	0.133219	0.120124	0.120588	0.140640	0.142243	0.146175	0.145976	0.147139
MLP	0.031053	0.042480	0.050606	0.057242	0.063189	0.071036	0.078087	0.081710	0.083372	0.085115
ENN	0.022911	**0.029707**	**0.033855**	**0.034691**	**0.037531**	**0.038591**	**0.039204**	**0.039534**	**0.039872**	**0.040486**
Building 7										
RT	0.030488	0.049363	0.064206	0.079134	0.090885	0.099889	0.106828	0.110788	0.114986	0.118440
BaRT	0.025758	0.039776	0.052577	0.064917	0.074822	0.083428	0.089966	0.094206	0.098149	0.101023
BoRT	0.033699	0.050512	0.062146	0.074838	0.084957	0.092793	0.098470	0.102427	0.106219	0.109829
L-SVM	0.111623	0.154890	0.261622	0.327596	0.482860	0.416295	0.359808	0.420238	0.419129	0.484455
C-SVM	0.050886	0.096670	0.131943	0.192059	0.192439	0.211648	0.232536	0.270985	0.268914	0.281900
MLP	0.027762	0.047868	0.066254	0.081760	0.095626	0.105703	0.113146	0.116764	0.120329	0.125132
ENN	**0.022944**	**0.033725**	**0.042692**	**0.047951**	**0.052137**	**0.055154**	**0.057240**	**0.056857**	**0.057524**	**0.058780**
Building 8										
RT	0.045387	0.066067	0.082481	0.096021	0.108163	0.117628	0.121427	0.122936	0.124140	0.126143
BaRT	0.039959	0.059244	0.074879	0.087520	0.099768	0.109324	0.113431	0.114853	0.116204	0.118459
BoRT	0.049992	0.070498	0.085710	0.099173	0.111072	0.118904	0.121412	0.123020	0.124214	0.125712
L-SVM	0.124275	0.210009	0.202029	0.234078	0.415058	0.387901	0.443007	0.386421	0.403097	0.479505
C-SVM	0.067852	0.107659	0.138375	0.170347	0.192552	0.217577	0.225939	0.225918	0.225650	0.221182
MLP	0.038648	0.057840	0.072436	0.084582	0.096046	0.104135	0.107953	0.109543	0.110910	0.112472
ENN	**0.031425**	**0.043970**	**0.050122**	**0.056304**	**0.060288**	**0.062305**	**0.063147**	**0.062334**	**0.062709**	**0.062508**

Table 2. Root Mean Squared Error of the best prediction model with different forecasting time steps.

Building	Forecasting $(t + f)$									
	1	2	3	4	5	6	7	8	9	10
B1	0.034165	0.045475	0.051615	0.055554	0.056859	0.057094	0.059120	0.058332	0.059158	0.058138
B2	0.031797	0.044207	0.051942	0.057654	0.062532	0.063908	0.066410	0.066248	0.067870	0.067452
B3	0.023489	0.034806	0.042175	0.048729	0.051864	0.055996	0.057125	0.061281	0.059965	0.060444
B4	0.028453	0.041269	0.049153	0.053318	0.056964	0.059172	0.061813	0.062787	0.062372	0.063499
B5	0.026887	0.033301	0.035983	0.038896	0.040655	0.041526	0.041912	0.042513	0.043044	0.043061
B6	0.022552	0.029707	0.033855	0.034691	0.037531	0.038591	0.039204	0.039534	0.039872	0.040486
B7	0.022944	0.033725	0.042692	0.047951	0.052137	0.055154	0.057240	0.056857	0.057524	0.058780
B8	0.031425	0.043970	0.050122	0.056304	0.060288	0.062305	0.063147	0.062334	0.062709	0.062508

On the other hand, if we continue observing Fig. 6, we can distinguish that there are buildings whose error is moderately lower than the rest of the buildings. This fact is on account of the particularities of the building as there are buildings whose behavior are less changing and fluctuating than others, and therefore those have a more foreseeable energy consumption. In any case, as shown in Fig. 6 and Table 2, ENN's errors are relatively close to one another. Besides, when analyzing in detail these results and taking $t + 1$ as a reference point, the next predictions are worsening by 38% for $t + 2$, on average. In the next case, it reaches an increase of 62%, 78% and 90% for $t = 3,4,5$ respectively. After that, the error passes with a difference of 100% and indicates a smooth growth from that point.

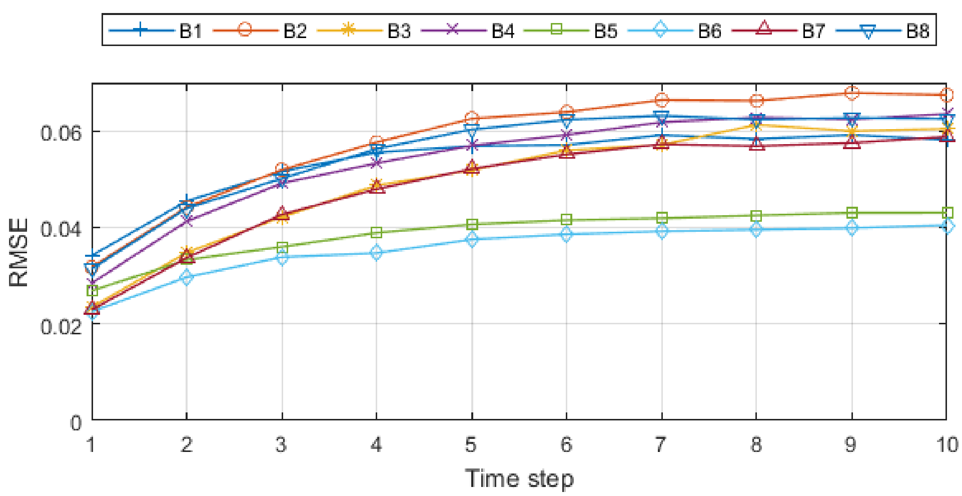


Fig. 6. Evolution of the RMSE as the forecasting time step increases.

Once forecasting models have been properly analyzed and the best option is selected. The need for going one step further is essential. As a result, the rest of this section will describe how VIOMEN was deployed at UGR to visualize the energy consumption at its campuses, along with the related predictions provide by the forecasting models.

An initial appearance of the application is depicted in Fig. 7. As seen in the figure, two thirds of the screen are dedicated to the map visualization where buildings are displayed, which is where most of the relevant knowledge will be presented. This area also contains two buttons (located in the upper left corner of the view) used to modify clustering and colour properties, and a textbox which allows us to look for places and streets on the map by typing global search terms. The remaining one third of the application (right part) lists all the buildings, serving as a "calling card" of each building.

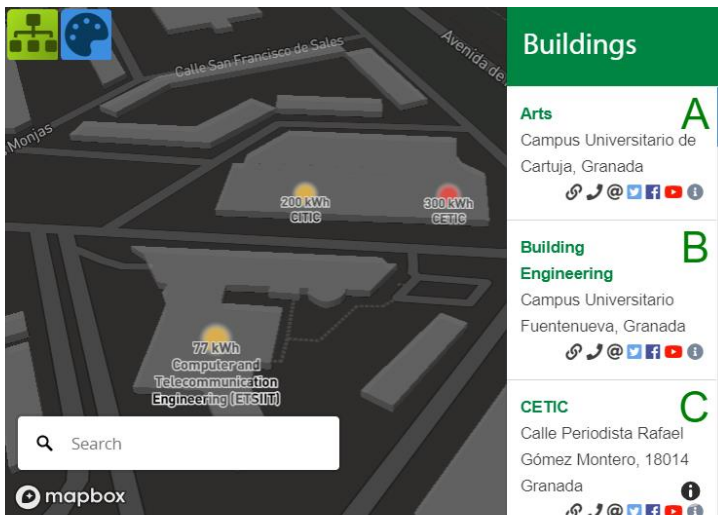


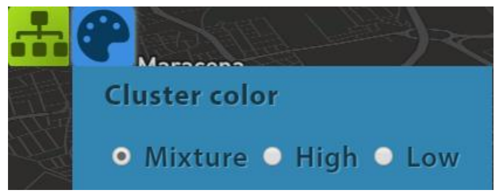
Fig. 7. Example of a general view of the proposed system.

Due to the relatively high number of co-located elements when zooming out, it is necessary to perform a building grouping. Thus, VIMOEN reduces the number of drawn points on the grid when points collide. It is possible to achieve various form of aggregation, and we consider four of them (as illustrated on Fig. 8a).

A cluster's energy consumption may be represented by the maximum, the minimum, the average or the total amount of the consumption in that group. The purpose of these adjustments is to adapt the system to user needs. For instance, the maximum value allows an easy detection of a potential building problem in the consumption; alternatively, the measure of the average gives an idea of the consumption for each of the areas concerned; similarly, the other two options furnish valuable information depending on the context from which the subject is approached.

In keeping with this philosophy, the colour of the clusters is also considered flexible to adjustments. For this end, the colour of the group admits three strategies as the Fig. 8b illustrates. The mixture option performs the averaging of the red, green and blue components and the alpha channel. The high and low options express the group according to the biggest and smallest consumption respectively.

(a)



(b)

Fig. 8. Upper left buttons: available (a) clustering properties and (b) colour behaviour of the group.

VIMOEN thus far has related to the color properties and diverse available changes in the visualization features. As the consumption profiles of each building are dissimilar, each point should be coloured differently according to the range of its historical energy use. In other words, there is a colour palette where green, orange and red represent that the building's consumption is close to its recorded minimum, medium and maximum energy respectively. Using these three colours, a set range of 10 colours have been created to distinguish the consumption considering the energy use of each building. Thereby, two buildings with the same consumption at a certain time could have different colour, because their historical consumptions are different. There is a similar situation in Fig. 7 where two buildings with different consumption are coloured with the same colour due to the different range of consumption.

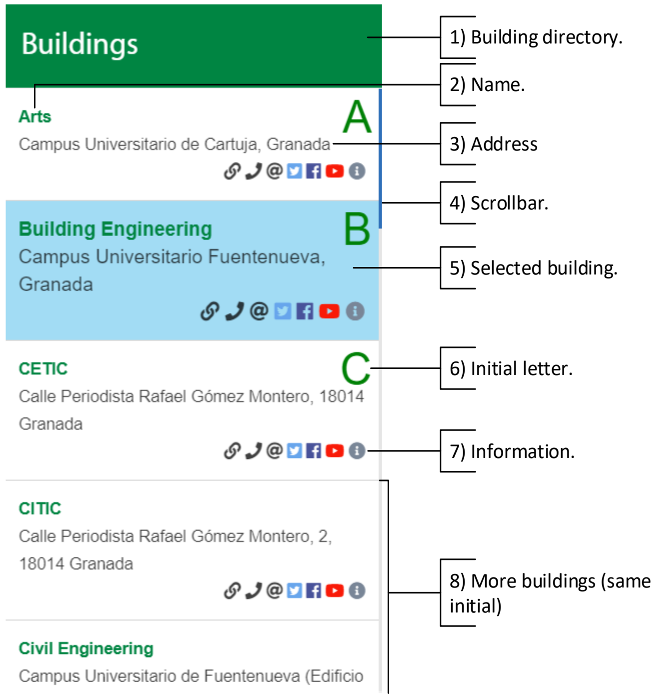


Fig. 9. Partial view of the list of UGR's buildings in VIMOEN.

Fig. 9 shows the area listing of the existing buildings. The building directory is sorted alphabetically by name for convenience. Buildings' names are distinguished by its initial letter. Each description card summarizes information about one building, e.g., the name of the building, its specific address (street, number, city), the URL of the building website, the contact phone number, an e-mail address, its social networks (twitter, Facebook and YouTube) and other information that describes the building. For our use case of a university, those parameters make sense and have actual values as they correspond to faculties. The design of the menu is a simple interface for showing and locating the buildings in the map in a quick and straightforward way.

Fig. 10. Example of map pin when building is select in the list.

Thus, a white map pin appears when the mouse pointer is placed into the building card from the list and its cell is displayed in light blue. This is reflected in Fig. 10. What's more, the application incorporates the functionality of adjusting the current location to the new selected building. This is carried out when click on one of the rows of the list, providing for easily location of the buildings.

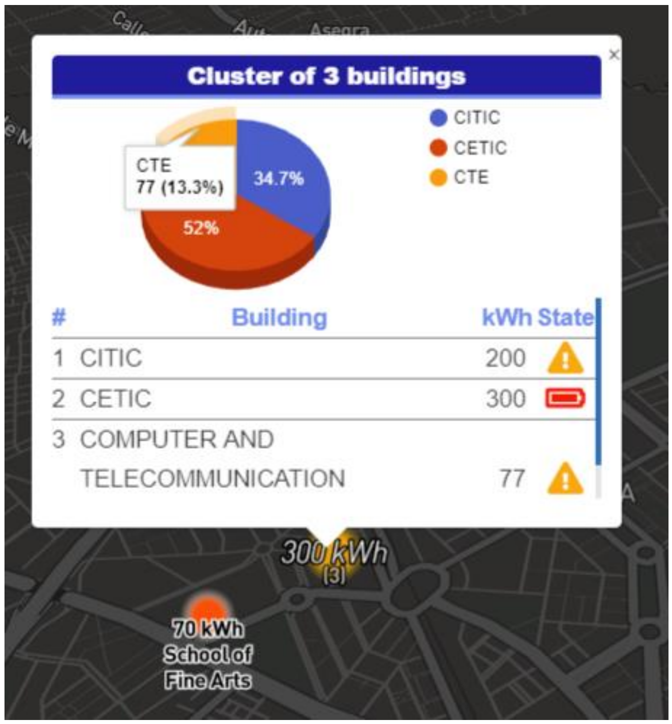


Fig. 11. Example of a popup for a cluster.

Once a building (or cluster of them) is selected, a popup appears with useful energy consumption information. Fig. 11 details the information for a building group of three edifices. The proportional amount of energy from each building is represented with an SVG-rendered pie chart. When pieces of the chart are selected, this graph also provides the detailed individual figures. In order to optimize the space of the chart, the name of the buildings is labelled by taking its initials, e.g., the CTE building is the Computer and Telecommunications Engineering School. Tooltips are also displayed on hovering.

In addition to this, a description of the consumption appears below the graph. All buildings are listed using its complete name and its current energy consumption.

The last column of the scrollable list gives visual information about the building's behaviour. As illustrated before in Fig. 7, two of the three buildings are consuming above their average, and the other one is consuming close to the maximum registered. Fig. 12a lists all available icons, together with their meaning, for visually expressing the energy behaviour of the building.

	Consumption below its average.
	It is consuming beyond the average.
	Consumption near the max recorded.
	Energy consumption unavailable.

(a)

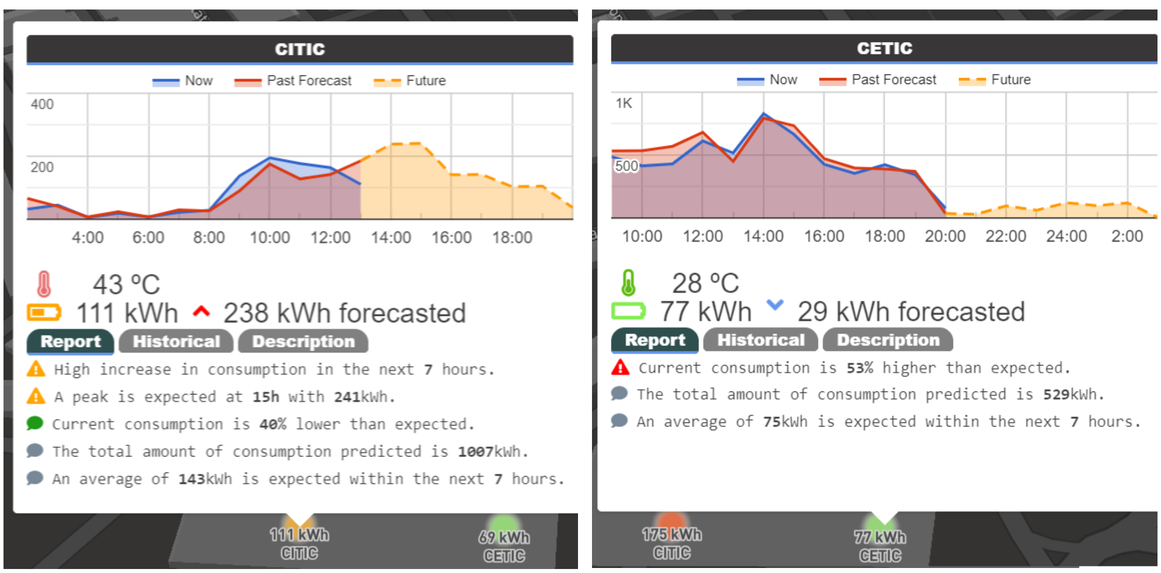
	The consumption is predicted to increase.
	The consumption is predicted to decrease.
	The consumption maintains its value.

(b)

Fig. 12. Set of icons for representing the (a) current status and (b) the future consumption of the buildings.

Apart from that, the individual popup describes important information about a building (Fig. 13). The specific popup of one building displays the current energy consumption on the central left-hand together with a battery icon which indicates the state of the building according to the range of energy use. The color and the filling of the battery change dynamically according to current consumption. The predicted consumption is located at the right side of the current consumption. This information is also presented with an animation (Fig. 12b) indicating if the prediction goes up, down or maintains its value. The specific forecasting value is shown next to that animated icon. Not only the current and the forecasting consumption is reflected in the panel, but also an historical evolution of the building (blue line), the past forecasted values (red) and the future consumption (orange). In this way, the user may observe the evolution of the actual consumption in the recent past, the past forecasting values help user to discern if the consumption at that period has been expected or, on the contrary, it differs from the predictions. In addition to this, the user can anticipate future behaviors in buildings by using all the foreseen information from forecasting models. The orange time-series describes an easy way to see at a glance what is going to take place.

Furthermore, at the bottom of the panel, there are three tabs about relevant information about the building. For instance, Fig. 13a depicts a warning as the consumption will grow considerably, as well as an expected peak within 2 hours. In this case, the current consumption was lower than predicted, so the system reports it as well, along with the total consumption and an average of the consumption per hour, although each specific prediction may be found in the main chart. On the contrary, Fig. 13b depicts the energy consumption of a building whose actual consumption is close to its minimum (green empty battery), and it is expect to decrease its consumption in the next hour (blue arrow). In this building, an alert is shown as the current consumption is significantly higher than expected. Nevertheless, there is no warnings about peaks nor high increases, as in Fig. 13a .

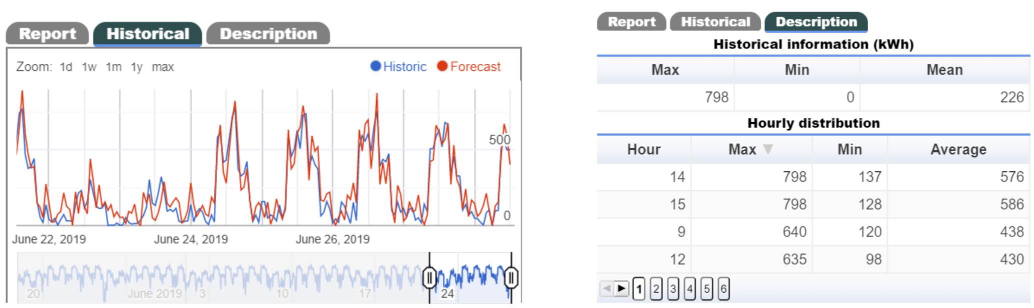


(a) (b)

Fig. 13. Example of the individual popup for two buildings with different information, alerts and energy consumption range.

Lastly, some buildings also collect the temperature conditions, that information is depicted with a thermometer which is filled in accordance with the registered maximum and minimum temperature of the building, in the same way as the battery icon. Both icons flash in red when these are close to the maximum value.

The last feature of VIMOEN to detail is illustrated in Fig. 14. It is essential to offer the user the maximum information possible, as a consequence, the historical consumption could be examined. This feature is made up of two parts, the whole consumption stored on the bottom, and the selected period which is depicted on the top. And to fully exploit the data, an historical report about the consumption is rendered in the third tab (Fig. 14b). This tab gathers the maximum, minimum and the mean recorded up to now, along with the detailed distribution of the consumption in all hours, even sorting the to make the most of the depicted information, e.g., at 14:00 is the period with the higher maximum.



Max	Min	Mean
798	0	226

Hourly distribution

Hour	Max	Min	Average
14	798	137	576
15	798	128	586
9	640	120	438
12	635	98	430

(a) (b)

Fig. 14. Example of (a) the historical consumption (focusing on the last week) and (b) description of the hourly distribution energy consumption.

The combination of VIMOEN with the experience, skills and dedication of the users produces a very powerful tool which is useful and helpful for monitoring and management of the energy consumption in distributed facilities.

6. Conclusion

This work describes a complete methodology to solve the problem of energy consumption forecasting along with the many-times forgotten application in real life. In our study, we apply data pre-processing to transform data from sensors into a more usable way which includes noisy treatment and data imputation. Our experimentation compares several forecasting models in order to obtain the best approach. In this study Elman Neural Network has proven to be the most accurate model to predict energy consumption. In addition to this, different time steps were tested to enrich and improve the functionality of the system, as well as to confirm the validity of the model to make medium term predictions. Thus, the best prediction model is nominated to make predictions in VIMOEN, a system for energy visualization in institutions and organizations whose facilities are distributed throughout several geographic zones.

VIMOEN is designed to easily incorporate new facilities and updated knowledge about new structures. As such, it is a novel prototype for visualization and monitoring the energy consumption which comprises the visualization of historical data with a view to detect and control energy problems for the future, with the intention of contributing and assisting in decision-making processes, and therefore, enabling everyday saving in energy consumption (reported in real time as shown in Fig. 14a). What's more, an advantage of our design is its simplicity. However, this strength limits the wealth of knowledge which could be displayed, since it is harder to understand dashboards that are overstuffed with information. Our choice is made to ensure an intuitive, easy-to-use interface and suitable representation of the knowledge to assist user interaction with the system.

VIMOEN provides a solution for examining and displaying the energy consumption of a set of geographically-distributed buildings. This task, which is costly, complex, time consuming and usually carried out manually using spreadsheets, changes into an easy visual exercise thanks to VIMOEN. As it enables the visualization of all distributed buildings of an organization and incorporates their past, current and future consumptions.

For its implementation, we used Mapbox, which allows the creation of three-dimensional (3D) effects with the fill extrusion layer. Even though, building models with a high level of details can be a costly and time-consuming process, it could be a very useful approach for future work. Thus, future research will be oriented at developing three-dimensional (3D) models of the buildings with a low level of detail. And indeed there are some recent methodologies for creation of mock-up 3D models of buildings using an image-based automated process (Oskouie, Becerik-Gerber & Soibelman, 2017). This would contribute to create a more accurate view of the whole campus. Along these lines, the operating rate of a building could be considered for visualization in those 3D models and thus serve to facilitate cost control, distinguishing between busy and inactive spaces. And besides, it would improve the energy prediction because of some researches point out that the occupancy is the most important factor affecting energy waste (Deb, Eang, Yang & Santamouris, 2016).

7. Abbreviations

ANN	Artificial Neural Network.
BaRT	Bagged Regression Tree.
BoRT	Boosted Regression Tree.
C-SVM	Coarse Support Vector Machine.
ENN	Elman Neural Network
L-SVM	Linear Support Vector Machine.
MLP	Multi-Layer Perceptron.
RMSE	Root Mean Squared Error.
RT	Regression Tree.
SVM	Support Vector Machine.

8. Acknowledgments

This work has been developed with the support of the Department of Computer Science and Artificial Intelligence of the University of Granada, Approximate Reasoning and Artificial Intelligence Group (TIC111), and the project TIN201564776-C3-1-R.

9. Bibliography

Andiojaya, A. & Demirhan, H. (2019). A bagging algorithm for the imputation of missing values in time series. Expert Systems with Applications, 129, 10-26. http://dx.doi.org/10.1016/j.eswa.2019.03.044

Bull, A. D. (2011). Convergence rates of efficient global optimization algorithms. Journal of Machine Learning Research, 12, 2879-2904. Retrieved from http://arxiv.org/abs/1101.3501v3

Chae, Y. T., Horesh, R., Hwang, Y. & Lee, Y. M. (2016). Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. Energy and Buildings, 111, 184-194. http://dx.doi.org/10.1016/j.enbuild.2015.11.045

Chen, J., Shi, Q., Shen, L., Huang, Y. & Wu, Y. (2019). What makes the difference in construction carbon emissions between China and USA? Sustainable Cities and Society, 44, 604-613. http://dx.doi.org/10.1016/j.scs.2018.10.017

Davlea, L. & Teodorescu, B. (2016). A neuro-fuzzy algorithm for middle-term load forecasting. In 2016 International Conference and Exposition on Electrical and Power Engineering (EPE) (pp. 005-009). http://dx.doi.org/10.1109/ICEPE.2016.7781292

Deb, C., Eang, L. S., Yang, J. & Santamouris, M. (2016). Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks. Energy and Buildings, 121, 284-297. http://dx.doi.org/10.1016/j.enbuild.2015.12.050

Deb, C., Zhang, F., Yang, J., Lee, S. E. & Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. Renewable and Sustainable Energy Reviews, 74, 902-924. http://dx.doi.org/10.1016/j.rser.2017.02.085

Duque-Pintor, F., Fernández-Gómez, M., Troncoso, A. & Martínez-Álvarez, F. (2016). A New Methodology Based on Imbalanced Classification for Predicting Outliers in Electricity Demand Time Series. Energies, 9, 752-752. http://dx.doi.org/10.3390/en9090752

Fumo, N. & Rafe Biswas, M. A. (2015). Regression analysis for prediction of residential energy consumption. Renewable and Sustainable Energy Reviews, 47, 332-343. http://dx.doi.org/10.1016/j.rser.2015.03.035

Gelazanskas, L. & Gamage, K. A. A. (2014). Demand side management in smart grid: A review and proposals for future direction. Sustainable Cities and Society, 11, 22-30. http://dx.doi.org/10.1016/j.scs.2013.11.001

Granderson, J. & Lin, G. (2016). Building energy information systems: synthesis of costs, savings, and best-practice uses. Energy Efficiency, 9, 1369-1384. http://dx.doi.org/10.1007/s12053-016-9428-9

Huacón, C. F. & Pelegrin, L. (2017). SURV: A system for massive urban data visualization. In Undergraduate Research Technology Conference (URTC), 2017 IEEE MIT (pp. 1-4): IEEE. http://dx.doi.org/10.1109/URTC.2017.8284174

International Energy Agency. (2013). Transition to Sustainable Building. Strategies and Opportunities to 2050. In. Retrieved from http://www.iea.org/publications/freepublications/publication/transition-to-sustainablebuildings.html

Jung, H. C., Kim, J. S. & Heo, H. (2015). Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach. Energy and Buildings, 90, 76-84. http://dx.doi.org/10.1016/j.enbuild.2014.12.029

Kayacan, E., Ulutas, B. & Kaynak, O. (2010). Grey system theory-based models in time series prediction. Expert Systems with Applications, 37, 1784-1789. http://dx.doi.org/10.1016/j.eswa.2009.07.064

Lopes, M. A. R., Antunes, C. H. & Martins, N. (2012). Energy behaviours as promoters of energy efficiency: A 21st century review. Renewable and Sustainable Energy Reviews, 16, 4095-4104. http://dx.doi.org/https://doi.org/10.1016/j.rser.2012.03.034

Mapbox & LLC. (2018). Mapbox GL JS API. URL http://www.mapbox.com/mapbox-gl-js/api/.

Murshed, S., Al-Hyari, A., Wendel, J. & Ansart, L. (2018). Design and Implementation of a 4D Web Application for Analytical Visualization of Smart City Applications. ISPRS International Journal of Geo-Information, 7, 276. Retrieved from http://www.mdpi.com/2220-9964/7/7/276

Nagy, G. I., Barta, G., Kazi, S., Borbély, G. & Simon, G. (2016). GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach. International Journal of Forecasting, 32, 1087-1093. http://dx.doi.org/10.1016/j.ijforecast.2015.11.013

Oh, T. K., Lee, D., Park, M., Cha, G. & Park, S. (2018). Three-Dimensional Visualization Solution to Building-Energy Diagnosis for Energy Feedback. Energies, 11, 1736. Retrieved from http://www.mdpi.com/1996-1073/11/7/1736

Oskouie, P., Becerik-Gerber, B. & Soibelman, L. (2017). Automated Recognition of Building Façades for Creation of As-Is Mock-Up 3D Models. Journal of Computing in Civil Engineering, 31, 04017059. http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000711

Pahl, S., Goodhew, J., Boomsma, C. & Sheppard, S. R. (2016). The role of energy visualization in addressing energy use: Insights from the eViz project. Frontiers in psychology, 7, 92. Retrieved from http://www.frontiersin.org/articles/10.3389/fpsyg.2016.00092/full

Pérez-Lombard, L., Ortiz, J. & Pout, C. (2008). A review on buildings energy consumption information. Energy and Buildings, 40, 394-398. http://dx.doi.org/10.1016/j.enbuild.2007.03.007

Prasad, A. M., Iverson, L. R. & Liaw, A. (2006). Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. Ecosystems, 9, 181-199. http://dx.doi.org/10.1007/s10021-005-0054-1

Rodrigues, N., Netzel, R., Ullah, K. R., Burch, M., Schultz, A., Burger, B. & Weiskopf, D. (2017). Visualization of time series data with spatial context: communicating the energy production of power plants. In Proceedings of the 10th International Symposium on Visual Information Communication and Interaction (pp. 37-44): ACM. http://dx.doi.org/10.1145/3105971.3105982

Ruiz, L. G. B., Cuellar, M. P., Delgado, M. & Pegalajar, M. C. (2016). An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings. Energies, 9, 21. http://dx.doi.org/10.3390/en9090684

Ruiz, L. G. B., Rueda, R., Cuéllar, M. P. & Pegalajar, M. C. (2018). Energy consumption forecasting based on Elman neural networks with evolutive optimization. Expert Systems with Applications, 92, 380-389. http://dx.doi.org/10.1016/j.eswa.2017.09.059

Santamouris, M. & Kolokotsa, D. (2015). On the impact of urban overheating and extreme climatic conditions on housing, energy, comfort and environmental quality of vulnerable population in Europe. Energy and Buildings, 98, 125-133. http://dx.doi.org/10.1016/j.enbuild.2014.08.050

Sarma, P., Singh, H. K. & Bezboruah, T. (2018). A Real-Time Data Acquisition System for Monitoring Sensor Data. Retrieved from http://www.researchgate.net/profile/Pratiksha_Sarma/publication/326398953_A_Real-Time_Data_Acquisition_System_for_Monitoring_Sensor_Data/links/5b4a05b7aca272c609409d43/A-Real-Time-Data-Acquisition-System-for-Monitoring-Sensor-Data.pdf

Singh, V. & Tiwari, K. (2017). Prediction of GreenHouse Micro-Climate using Artificial Neural Network. APPLIED ECOLOGY AND ENVIRONMENTAL RESEARCH, 15, 767-778. Retrieved from http://www.aloki.hu/pdf/1501_767778.pdf

Xu, M., Watanachaturaporn, P., Varshney, P. K. & Arora, M. K. (2005). Decision tree regression for soft classification of remote sensing data. Remote Sensing of Environment, 97, 322-336. http://dx.doi.org/10.1016/j.rse.2005.05.008

Yuan, C., Liu, S. & Fang, Z. (2016). Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM (1,1) model. Energy, 100, 384-390. http://dx.doi.org/10.1016/j.energy.2016.02.001

6.5 A Time-Series Clustering Methodology for Knowledge Extraction in Energy Consumption Data

Referencia

Ruiz, L. G. B., Pegalajar, M. C., Arcucci, R, Molina-Solana, M., (2019). A Time-Series Clustering Methodology for Knowledge Extraction in Energy Consumption Data.

Estado

En revisión.

A Time-Series Clustering Methodology for Knowledge Extraction in Energy Consumption Data

L.G.B. Ruiz [a,*], M.C. Pegalajar [a], R. Arcucci [b], M. Molina-Solana [b]

[a] Department of Computer Science and Artificial Intelligence, University of Granada.

[b] Data Science Institute, Imperial College London, UK.

bacaruiz@ugr.es* / mcarmen@decsai.ugr.es / m.molina-solana@imperial.ac.uk / y.guo@imperial.ac.uk

Abstract. One of the essential aims for incorporating intelligent systems in cities and buildings is the energy savings and pollution reduction that can be attained. To achieve this goal, energy modelling and better understanding of how energy is been consumed are key factors. As a result, a methodology for knowledge acquisition in energy-related data is proposed here by the use of Time-Series Clustering (TSC) techniques. In our experimentation, we utilize data from the buildings at University of Granada (Spain) and we compare several clustering methods for getting the best model as well as several algorithms for obtaining the best grouping. Thus, our methodology is able to obtain non-trivial knowledge from raw energy data. In contrast to previous studies in this field, an automatic recursive strategy is proposed here for searching and analysing periodicity in these series.

Keywords: *Time-Series Clustering; Energy Efficiency; Knowledge Extraction; Data Mining*

1. Introduction

Buildings account nowadays for a high percentage of global energy consumption (along its associated carbon emissions) in developed countries (Levesque, et al., 2018), and particularly in Europe and US, where buildings consume about 40% of primary energy (Berardi, 2015; D'Agostino, Cuniberti, & Bertoldi, 2017). Interestingly, approximately 90% of building emissions are caused by lighting and HVAC operating systems. They are responsible for maintaining the inner comfort by means of managing heating, ventilation and lighting conditions. However, a major concern is that those systems are prone to noise, light pollution and health problems (Belley, Gaboury, Bouchard, & Bouzouane, 2015). An additional complication is that energy demand is expected to increase in the near future due to the climate change caused by inefficient usage of energy, population growth and increase of time spent in buildings (Cao, Dai, & Liu, 2016).

For those reasons, new studies and proposals have been put forward to achieve energy savings and address this problem in our daily environment. A common approach in literature is to consider occupant's behavior and their interactions. This has proved to be a key factor in energy savings. There are some studies proposing to monitor energy consumption and offer support and advice (Ashouri, Haghighat, Fung, Lazrak, & Yoshino, 2018), suggesting energy profiles based on different energy efficiency levels, occupancy and the greenness of household behavior (T. Zhang, Siebers, & Aickelin, 2012), and using occupant behavior modeling for improving occupant information about energy use (T. Hong, Sun, Chen, Taylor-Lange, & Yan, 2016).

Overall, rapid advances in sensing technologies and wireless networks open up the possibility of accessing to massive amount of data and consequently opening new lines of study. Many problems can be found in this regard, such as energy consumption modelling (Belaïd, Roubaud, & Galariotis, 2019), energy storage (Romanchenko, Kensby, Odenberger, & Johnsson, 2018), power quality (M. Hong, Yu, Yu, & Loparo, 2016), fault diagnosis (Capozzoli, Lauro, & Khan, 2015), anomaly detection (Capozzoli, Piscitelli, Brandi, Grassi, & Chicco, 2018), energy management (Molina-Solana, Ros, Ruiz, Gómez-Romero, & Martin-Bautista, 2017), energy prediction (Ruiz, Rueda, Cuéllar, & Pegalajar, 2018), or load profiling (T. Zhang, et al., 2012).

Decision-making in this context is a complex task generally carried out by experts who manually analyze raw energy-related data and propose energy planning decisions (Strantzali & Aravossis, 2016). As a result, diverse studies have been carried out so far: Lopez et al. (López-Ruiz, Bergillos, & Ortega-Sánchez, 2016) proposed a new methodology for stochastic prediction of climate-related variables which are directly involved in the decision making process to manage, distribute and maintain energy converters; Duque et al. (Duque-Pintor, Fernández-Gómez, Troncoso, & Martínez-Álvarez, 2016) developed an a priori technique for outlier detection in imbalanced data using a bagging approach; a generic energy consumption model is proposed in (Dietmair & Verl, 2009) for decision making and energy efficiency minimization in any given production system; Chen et al. (Chen, Taylor, & Wei, 2012) designed a model for representing energy consumption decision-making in building and thus exploring how energy consumption occupants' behaviors are linked to the structural properties of peer networks. There are also several articles describing the implementation of diverse forecasting models supporting decision-making processes and preventing unnecessary waste in energy consumption (similarly to (Duque-Pintor, et al., 2016)). Those models provide intelligence within a building for enhancing energy efficiency, saving costs and reducing the environmental impact of waste, residues and atmospheric emissions. For example, a forecasting framework to examine data from both wind and solar energy using statistical methods were proposed in (Andrade & Bessa, 2017), linear regression was used in (Fumo & Rafe Biswas, 2015) for residential energy forecasting, an a tree-based approach using a generalized additive tree ensemble was employed in (Nagy, Barta, Kazi, Borbély, & Simon, 2016) for solar and wind power prediction as well as deep learning neural networks (Torres, Troncoso, Koprinska, Wang, & Martínez-Álvarez, 2018) applying big data and time-series methods. For a deeper discussion in energy consumption forecasting techniques we refer the reader to (Deb, Zhang, Yang, Lee, & Shah, 2017).

Specifically, our former works were based on energy consumption models for solving this issue. Firstly, we employed Non-Linear Autoregressive Neural Networks to predict energy consumption in university buildings (Ruiz, Cuellar, Delgado, & Pegalajar, 2016). Secondly, we improved our models by means of Elman Neural Networks and evolutionary optimization archiving better and more precise results (Ruiz, et al., 2018). Thirdly, we parallelized the evolutionary adaptive search algorithm improving neural network-based models and thus enhance time cost of the models (Ruiz, Capel, & Pegalajar, 2019). All of them supply a powerful tool for improving energy efficiency. Even though predictive models enable to anticipate future events and to determine forthcoming energy consumption based on trends and historical data, they are not able to describe how buildings are consuming. This is the reason why descriptive models are necessary in order to obtain a complete system. Descriptive models utilize data to uncover behaviors and hidden patterns. These models enable the visualization, detection and discovery of this sort of knowledge, which can inform decision-making processes.

Nevertheless, recent literature has focused on forecasting models, and not so much on descriptive analytics. We hypothesize this is because arranging data into groups when working with unlabeled datasets is a more creative and complex task to accomplish. Those groups (or clusters) are aggregated by minimizing within-group-object similarity and maximizing between-group-object dissimilarity according to a specific criterion (Warren Liao, 2005). Those sort of techniques are known as Time-Series Clustering strategies (X. Zhang, Liu, Du, & Lv, 2011). The parameterization and adjustment of the clustering strategy is often tangled and surrender to uncertainty. Some studies have tested the effect of similarity measures in clustering applications for identifying patterns (Iglesias & Kastner, 2013), and some others adapted TSC for predictive purposes (for instance, (Simmhan & Noor, 2013) designed an incremental clustering method and a similarity measure for determining clusters which is conceived to reduce forecasting errors. Regarding visualization of clusters, we can highlight (Wijk & Selow, 1999) for energy objectives and (Diansheng, Jin, MacEachren, & Ke, 2006) for geo-visual pattern exploration (they analyzed complex patterns across multivariate, spatial and temporal dimensions via grouping, arranging and visualization). Several methodologies for approximation have also been suggested such as (Espinoza, Joye, Belmans, & Moor, 2005) for short-term load forecasting, profile discovery and customer segmentation using periodic autoregression models; Gao and Ali (Gao & Malkawi, 2014) for energy performance benchmarking of commercial building; Gabaldon et al. (Gabaldon, et al., 2010) to obtain a decision-making system for electricity price modeling using self-organizing maps and statistical Ward's linkage clustering so as to identify electricity market prices; Dias and Ramos' approach (Dias & Ramos, 2014) is similar to the one below, they analyzed cycles in energy prices using Markov-based dynamic clustering. However, there is no clear methodology for start-to-end knowledge discovery in buildings, including how energy data should be treated or which approach should be adopted. As a result, this study proposed a TSC methodology to analyze and extract patterns in buildings, in thus to better understand how buildings are consuming energy and can inform about possible features and behaviours hidden in data. In this way, one may observe similar behavior throughout the year. For instance, when this analysis is performed, the following questions can be answered: *are the consumptions in August and December similar? Is there any repetitive behavior every day? Does consumption exhibit any monotonous trend somewhere? is there any peak in consumption which is related to any hour of the day?* The final goal is to implement energy saving practices.

The rest of the manuscript is organized as follows: Section 2 describes the proposed methodology for clustering analysis, methods and metrics employed; Section 3 reports a complete discussion of the results obtained and Section 4 gives some conclusions and directions of future work.

2. Methodology

2.1. Proposed scheme

Our methodology is depicted in Figure 1, and its pipeline is structured in three main stages: load, analysis and visualization.

The first stage consists in loading and processing energy consumption data collected in buildings. Those data are commonly collected from sensors and stored directly without any previous treatment, so that the database gathers raw cumulative consumption data recorded by the building's energy counter. Then if those data were represented, they would reveal a growing trend, and thus to calculate energy consumption e_t at time t we subtract the cumulative consumed energy at the moment immediately prior $D_{t-1}$from current moment D_t (Ruiz, et al., 2018).

Since meters send information every hour, the time granularity of the data will be hourly. Without losing generality, we will assume that sensor data is hourly. Once raw data have been transformed into a more usable form, noise and missing treatment is carried out. To achieve this target two techniques were applied. On the one hand, energy consumption is treated using moving average filter and a sliding window procedure to clean breaks and inconsistent data points (Smith, 1997) as described by equation (1), where w is the size of the window used. On the other hand, a linear interpolation is computed using the nearest neighbour points to impute missing values. Finally, preprocessed data is stored and database updated.

$$y(n) = \frac{1}{w} \cdot \left(x(n) + x(n-1) + \cdots + x\big(n - (w-1)\big) \right) \quad \textbf{(1)}$$

After that, each building is analyzed separately. The first analysis is called «full-length analysis» because it is elaborated using the whole time-series. This covers an initial search of period. This task is performed by using the autocorrelation function which is a stochastic process that measures the correlation between two points in time (Box, Jenkins, Reinsel, & Ljung, 2015). Thus, the time-series of each building is split into different time-series according to the identified period, with this process recurrently repeated until the analyzed time-series does not show any period, i.e., if a 24h-period is found, the hourly time-series is split into 24 distinct sequences and this task repeated not for the whole time-series but for each piece of the 24 series and this division indicates that the consumption has a periodicity of 24 values, in other words, daily. This step is designed to ensure that periods in the time-series are found even if they do not follow daily, weekly or yearly periodicity. This stage ends when the whole energy consumption has been analyzed and each part has been studied.

Thereafter, all this knowledge is summarized and presented to the user by means of clustering according to the observed period knowledge. Several approaches have been developed to represent those groups, along with the pattern extraction of each cluster. Energy consumption related to each cluster is determined by its pattern and boundaries. Finally, pattern distribution is performed considering time information. Accordingly, cluster information is visualized for its analysis, in such a way that knowledge can be extracted about energy patterns. This would help to identify how that pattern is distributed along the year, whether it is a cycling pattern or whether it is a typical behavior of a particular period.

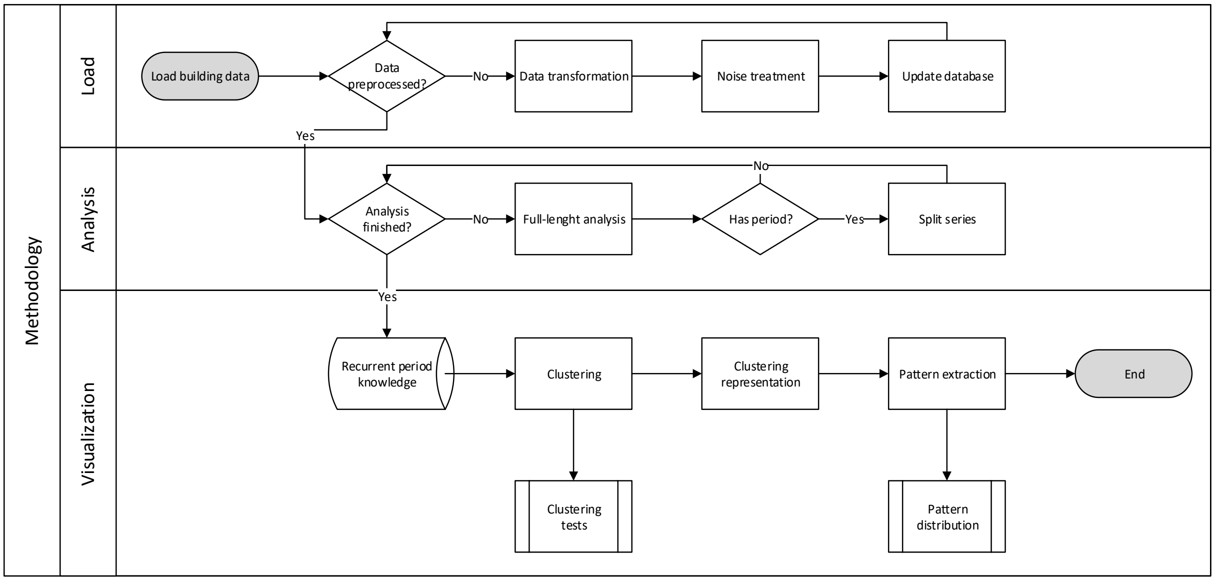


Figure 1. Overall scheme of the proposed methodology for a specific building.

Figure 1 illustrated the proposed steps for analysing energy consumption of a given building according to its energy features and without considering any other information related to other buildings (i.e. this provides knowledge about just one building). Since in a real-world scenario the ability to combine and manage diverse knowledge is essential for optimizing models and accomplishing better results, we designed a global strategy (depicted in Figure 2).

The first step is loading building data for all buildings. Then building analysis (as described before) is performed to extract patterns. Once all buildings have been processed, a global clustering is carried out to group all buildings. Finally, global clustering information is represented in a meaningful way to enable knowledge extraction and exploration of relationships among buildings, similar behaviours and to generalize on energy consumption performance.

Note that to apply this scheme there are two ways of treating data to make it compatible. The first one is to normalize data in a specific range, e.g., between [0, 1] as energy consumption are not usually in the same range among different buildings; in this way one may identify similarities not in actual-quantitative but in qualitative form due to the actual range of data is lost. The second one is to perform the clustering analysis without previous data transformation. In this manner, similar evolution in time of consumption will be clustered differently if the building's energy consumption range is different. In this case study, time-series are normalized so one can compare data among buildings. Those results will be indicated in following sections.

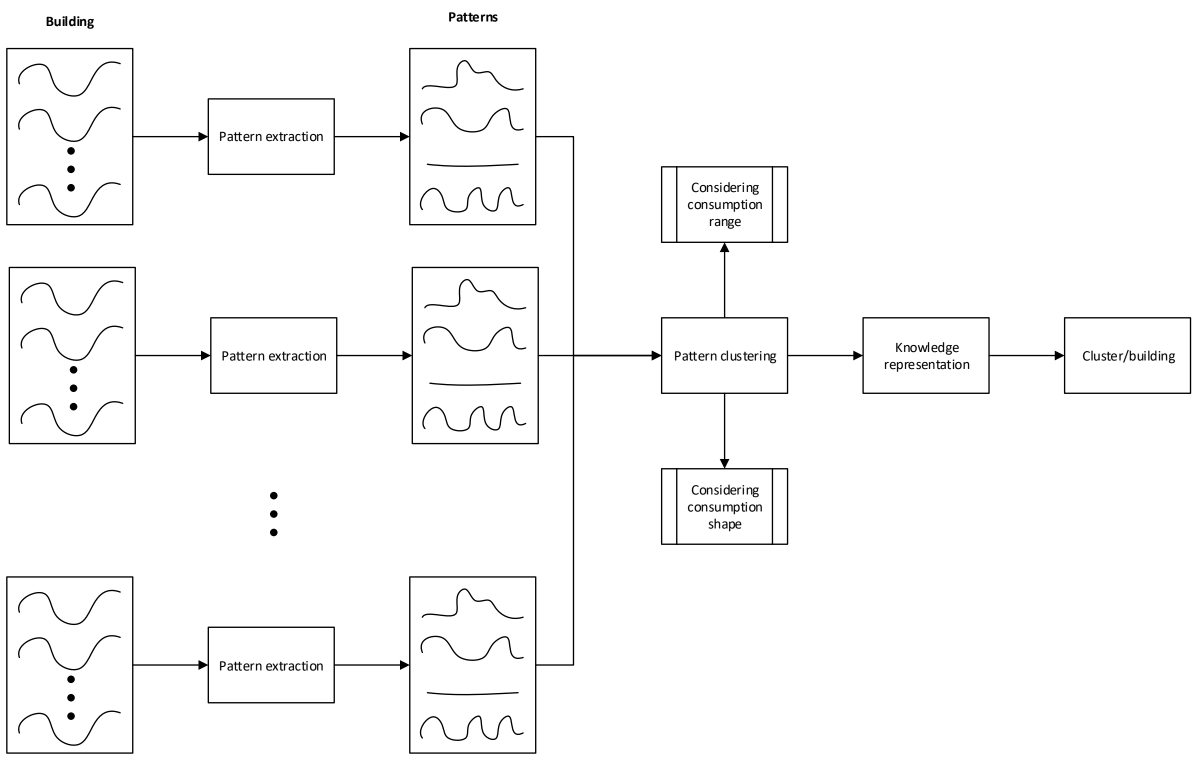


Figure 2. Global scheme to perform Time-Series-Clustering for a set of buildings.

2.2. Clustering methods

To perform time series clustering, we have selected four well-known techniques.

A. Hierarchical clustering

The first one is hierarchical clustering which partitions a multilevel hierarchy of groups by generating a cluster tree (known as "dendrogram"), in which groups at each level are joined as clusters at the next level. This enables an adaptation of the clustering scale for obtaining better solutions according to the desired aim. This method needs to build a distance function between samples which is employed in the kernel of this technique (Alves, Santos, & Schimit, 2016). In our case the agglomerative hierarchical cluster tree is used to create the matrix that encodes the clustering information. Besides, there are diverse algorithms for computing the distance between groups: unweighted average distance, centroid distance, farthest distance, weighted center mass distance, shortest distance, inner squared distance and weighted average distance. All those methods where tested and no substantive differences were observed in grouping, therefore the minimum variance algorithm (inner squared distance) has been set for comparing with all the other techniques.

B. Gaussian Mixture

Gaussian Mixture (GM) models is the second technique applied. Those models are composed of varied multivariate normal density components where each component has a mean, covariance matrix and the size of the cluster. The mixture component controls volume, orientation and the shape of the cluster (Maugis, Celeux, & Martin-Magniette, 2009). One of the main differences between GM and the other algorithms is that this algorithm works properly when groups are not roundly shaped or overlapped. GM addresses this issue by assigning samples to clusters according to specific probabilities. Once the algorithm has grouped data to a set of multivariate normal density components then the probability, or membership, of each point to clusters can be calculated. GM clustering models are considered a useful tool for grouping data with the assumption that data points are Gaussian distributed. In this way, not only spherical structures can be modelled, but also more complex configurations according to two parameters: mean and standard deviation (Celeux & Govaert, 1995). In our case, the Expectation-Maximization algorithm has been applied to fit GM models' parameters.

C. K-Means

k-Means is one of the best-known methods in literature for its efficiency and simplicity. This algorithm performs clustering on data by splitting data into disjointed clusters. As a result, each observation is assigned to only one cluster. k-Means is an unsupervised method which attempts to find the optimal set of points that minimizes the distances from each point to its nearest center (Duwairi & Abu-Rahmeh, 2015). It is composed of two main phases. In the first one, it computes each centroid and in the second step it takes each observation to the nearest centroid. Once the clustering is done it recalculates the new positions of all centroids by using the previous distance in order to achieve the minimum distance from each object to its centroid among all clusters (Dhanachandra, Manglem, & Chanu, 2015). In this work, the improved k-Means++ has been implemented. This algorithm uses a heuristic to search centroid seeds, and it achieves better time cost and better solutions than the original algorithm (Arthur & Vassilvitskii, 2007).

D. K-Medoids

And fourthly, the k-Medoids is a clustering method commonly employed in applications where robustness to outliers is an essential requirement. This algorithm is analogous to k-Means whose target is to split a set of points into several groups or clusters so that those clusters minimize the sum of distances between an observation and a centre (or medoid). The main difference between these two methods is that k-Medoids returns actual data points as representatives of the cluster. In this respect, we tested several algorithms to find the medoids: Partitioning Around Medoids (PAM) which is the classical algorithm for the k-Medoids problem and the Clustering Large Applications algorithm known as CLARA (Kaufman & Rousseeuw, 2009) and two variants of the Lloyd's algorithm for small and large scale dataset similar to k-means (Park & Jun, 2009), the first one tends to return better solutions than CLARA or the large approach, yet it is not a good option for big datasets.

2.3. Metrics

Organizing observations into groups requires defining some methods for estimating the distance as well as the similarity among points. As a result, a distance matrix is obtained and it is used for partitioning the observation into clusters. This distance should be established according to the final intended use and may vary depending on the need of an application. In our particular case, five different functions have been used and tested to compute distance between each pair of points.

A. Squared Euclidean Distance

The Squared Euclidean distance derives from the ordinary straight-line distance between two points known as Euclidean distance. It is calculated by computing the square root of the sum of the squares of the differences between those points (Fabregas, Gerardo, & Tanguilig III, 2017):

$$d(x,c) = (x - c) \cdot (x - c)' \tag{2}$$

Where x is an observation and c is a cluster, both are points, i.e., distance between points is defined with the same equation being c another point. This is the simplest and one of the most used distances.

B. Cosine Distance

The cosine distance is a measure of similarity between two vectors, and it is most frequently employed in high-dimensional positive spaces, e.g., text mining. The cosine distance is computed from one minus the cosine of the included angle between observations defined as follows (Yousef & Moghadam Charkari, 2015):

$$d(x,c) = 1 - \frac{x \cdot c'}{\sqrt{(x \cdot x') \cdot (c \cdot c')}} \tag{3}$$

C. Hamming Distance

The Hamming distance is commonly used in information theory for measuring distances between two string of equal length. This distance between two observations is the percentage of the vector components that differ (Titsias & Yau, 2017).

$$d(x,c) = \frac{1}{p} \cdot \sum_{j=1}^{p} I(x_j \neq c_j) \tag{4}$$

Where p is the length of the sample and $I(\cdot)$ is the indicator function which indicates membership of an element in a subset.

D. City Block Distance

City Block distance is also known as Manhattan distance. It represents distance between point in a "city road grid" and it determines the sum of absolute differences between observations (Prabhakar & Rajaguru, 2016):

$$d(x, c) = \frac{1}{p} \cdot \sum_{j=1}^{p} |x_j - c_j| \quad (5)$$

E. Correlation Distance

The last distance used in this work is the Correlation Distance. This one is popular on statistics, probability theory and time series applications, also called as covariance distance. This measurement examines the dependence between two paired vectors. The main advantage of this one is that it can measure dependence between two points of not necessarily equal dimension. It is computed from one minus the sample correlation between points, treated as sequence of values) (Jiang, Wang, & Zhang, 2019).

$$d(x, c) = 1 - \frac{(x - \bar{x}) \cdot (c - \bar{c})}{\sqrt{(x - \bar{x})^2} \cdot \sqrt{(c - \bar{c})^2}} \quad (6)$$

Where

$$\bar{x} = \frac{1}{p} \cdot \sum\nolimits_{j}^{p} x_j \, , \bar{c} = \frac{1}{p} \cdot \sum\nolimits_{j}^{p} c_j \quad (7)$$

Solution Evaluation

On the other hand, there is also another metric which should be defined which is not a distance. In order to evaluate the cohesion and separation among clusters the silhouette coefficient must be computed. This coefficient is a measure for scoring solutions. This measurement varies from -1 to 1. A higher silhouette coefficient indicates that the cluster is compact and poorly matched to other groups. This coefficient is described in (Gutiérrez-Batista, Campaña, Vila, & Martin-Bautista, 2018) as follows:

$$S = \frac{a - b}{\max(a, b)} \quad (8)$$

Where a is the average distance from the current point to the other points in the same cluster, and b is the minimum average distance from a specific point to points in another cluster. This value is minimized over clusters.

3. Results

This section summarizes the results from the range of tests we performed and their exploratory settings. We developed the code in Matlab R2017a and run them on an Intel® Core™ i7-4500U CPU @ 1.80 GHz, 2.40 GHz machine with 8 GBs of RAM. In order to avoid biased results, we executed each test 8 times. This configuration was chosen to make the most of the available resources because this processing unit has 4 cores and simulations can run in parallel. Note that only the most noteworthy visual results are presented in this section. All results can nevertheless be found online at: https://drive.google.com/drive/folders/1eOmjmeGjVdNjFGNQ_2D31ar5dZVbRaov.

3.1. Data collection

The used dataset contains energy consumption information obtained from buildings of the University of Granada, located in Spain. Those data were collected from building meters which gather sensors' measures. The whole campus of the University of Granada assembles diverse kind of buildings, such as laboratories, research centres, classrooms, lecture rooms, faculties, schools and so on. UGR has a smart management system which monitors and records that information from distributed sensors installed in the different buildings.

Given deployment done at different periods, not all facilities have the same information, yet energy consumption is recorded in most of them. In some of the buildings, data ages up to seven years ago. To comply with Data Protection Laws, buildings have been anonymized. Historical accumulated energy consumption data is stored in kWh on an hourly basis.

3.2. Discussion

After preprocessing raw energy data as described in Figure 1, the first step is to study periodicity for a specific building consumption. This period is found recurrently as described above. Thus, Figure 3a indicates that B1 presents a 24-hour period (autocorrelation diagram measures a high correlation at time 24 and 48, i.e., this building fulfils a daily period). Once the full-length analysis is done, then the time-series is split into k series according to its period, so that 24 new analyses are performed, one for each hour in this case. Figure 3b and c depict the autocorrelation analysis for the 1st and the 9th hour, respectively (no period found in the first hour, and period of 7 for the 9th). Notice that in these two figures the algorithm is working with pieces of the original time-series which correspond to each hour.

These 7 values mean the 7 days of the week, so a weekly period is found. Since Figure 3a is made up of hours, Figure 3b and Figure 3c displays days, as a consequence Figure 3d and Figure 3e will work in weeks, and Figure 3f in years. Thus, a period of 52 weeks is manifested in Figure 3d which means a yearly period in the 1st day of the week and this fact is not discover in the 6th day as Figure 3e illustrates. Finally, as there is no enough data (4 years in this case), just only 4 lags are indicated in Figure 3f and there is no evidence of period in yearly time scale. In this case, all periods are aligned with the expected grid.

With more data available, or a higher granularity (e.g., seconds, minutes), recurrent depth could be higher than the current max of 4.

Also, the correlation matrix reveals some meaningful information about the specific depth level of the time series. Accordingly, leaning on the Figure 4 results, strong correlation is shown between two periods of the day. The first one between the 22nd hour of the day and the 8th (or even 9th) hour where a high correlation is manifested; the second one is found between the 9th and the 22nd hour, that fact means that there are similar consumption behaviours among the hours of the same range, it seems to be the nocturnal interval and the diurnal season.

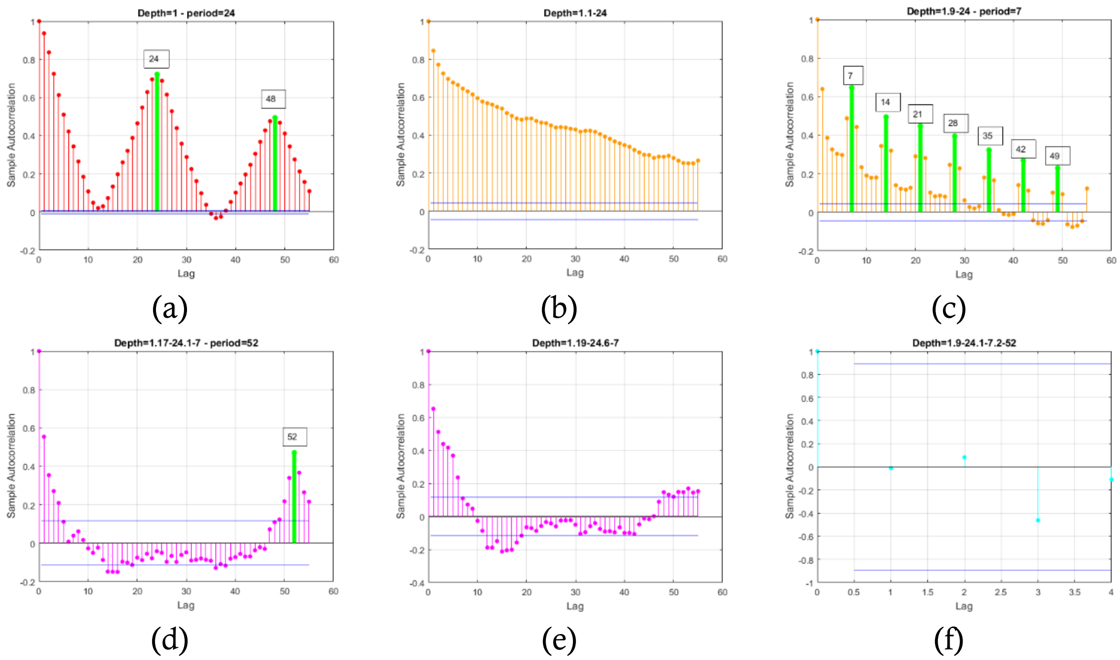


Figure 3. Autocorrelation diagrams of the recurrent analysis carried out for B1. Example of the (a) hourly data where 24-hour period is found, (b) the 1st hour out of 24 of the same building, (c) the 9th hour where period 7 is found, (d) the 1st piece out of 7 parts of the 17th hour where period 52 is found, (e) the 6th piece out of 7 from the hour 19, (f) the 2nd piece out 52 from the 1st part out of 7 from the 9th hour.

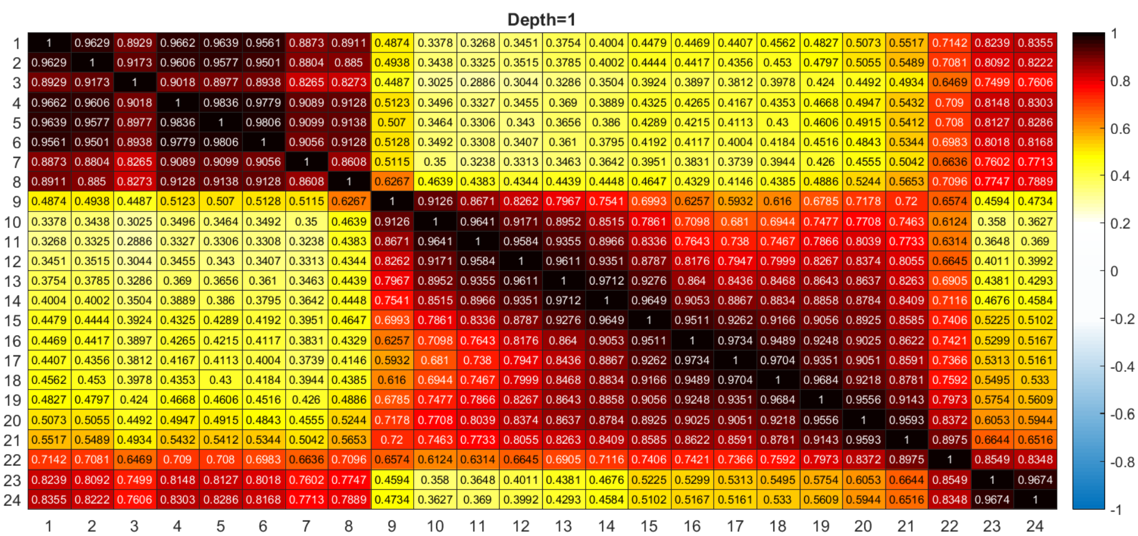
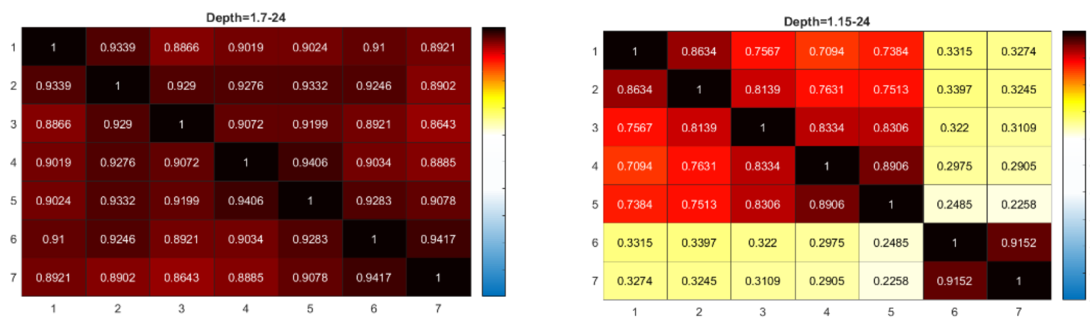


Figure 4. Correlation matrix example of the whole time series of building B1.

Another pair of examples are illustrated in Figure 5, where two hours are analyzed. The first one, depicts the autocorrelation matrix for the 7th piece of the consumption if it is split into 24 parts, i.e., the 7th hour. As can be seen there is a high correlation among the seven days, and there is no appreciable difference between the 7 a.m. on Monday and the same hour on Sunday. On the other hand, observing Figure 5b one can find that there is more correlation among working days (from 1 to 5, i.e., Monday to Friday) and also there is the same strong correlation between Saturday and Sunday as there are weak correlations between working days (1 to 5) and non-working days in that period.

Note that depending on the hour this weakness become bigger or smaller. But due to space restriction no more examples are shown in this context, although those instances are enough for covering this part of the analysis. All results can be found online at: https://drive.google.com/drive/folders/1eOmjmeGjVdNjFGNQ_2D31ar5dZVbRaov



(a) (b)

Figure 5. Correlation matrix example of the whole time-series of building B2 for both 7th and 15th hours of the day.

As a first step towards performing cluster analysis, one may appreciate those dependencies which was just mentioned in an initial overview of those hours grouping the whole sequence in Figure 6 where just 500 days are displayed for better visualization. Looking at the dendrogram plot it looks clear that there are two big groups. On the one hand those hours marked in red made up of 12 hours (from 22 p.m. to 9 a.m.) which correspond with a period with low activity registered, and also one may notice that these limits (9 a.m. and 22 p.m.) are the further hours, most probably because at that time people are usually going to work or leaving it. However, this is a reasonable assumption that should be left in expert's hands as we do not count with occupancy data to validate it. On the other hand, the second group is composed of the rest of the hours, but it is divided into two clusters, from 10 a.m. to 14 p.m. (green cluster) and from 15 p.m. to 21 p.m. (blue cluster). Curiously, groups are linearly split in time, i.e., there are three well differentiated periods, night and early morning (red), morning before lunch (green) and afternoon (blue). Note that in Spain lunch is regularly held between 2 p.m. and 3 p.m. Figure 7 illustrates the mean for each cluster, so that each group gathers different consumptions. The green one represents those hours whose consumption is usually the highest one. The blue cluster is the afternoon period (with an intermediate consumption). And finally, the red line indicates those evening and night hours in which expenditure is considerably lower.

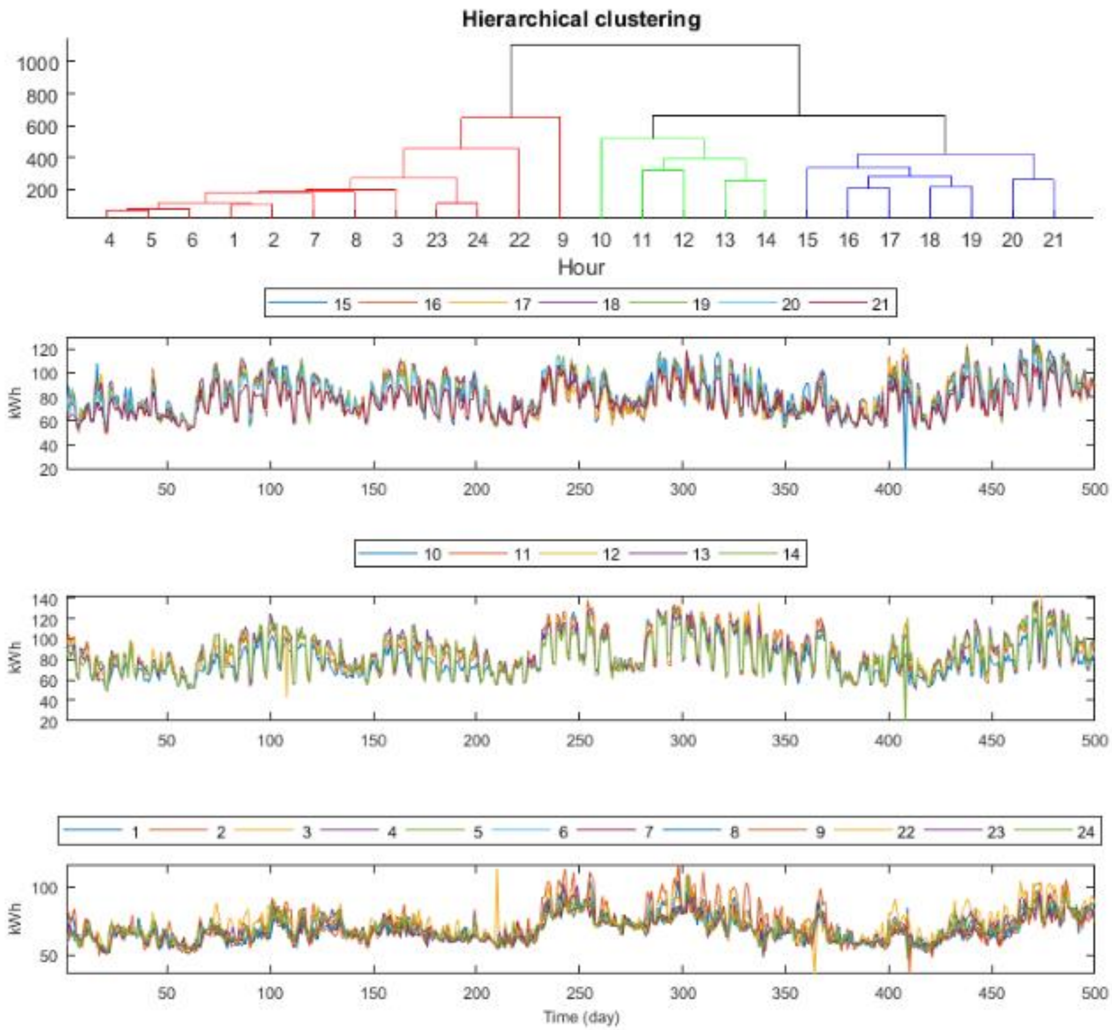


Figure 6. Hierarchical clustering for the whole time-series. Dendrogram of the hierarchical cluster tree for hourly sequences (first figure) and the representation of three clusters, blue (second plot), green cluster (third plot) and red group (fourth plot).

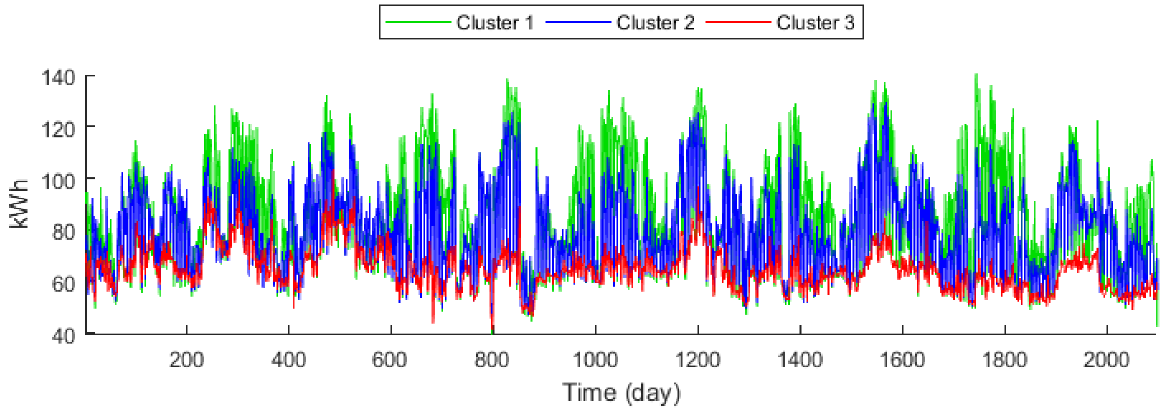


Figure 7. Example of the average consumption for each cluster. Cluster 1, hours from 10 a.m. to 14 p.m. (green). Cluster 2, hours from 15 p.m. to 21 p.m. (blue). Cluster 3, hours from 22 p.m. to 9 a.m. (red).

At this point, one may retrieve a great deal of information from data. However, this approximation has not yet exploited the full capabilities of clustering techniques. To do so, the first step is to determine which is the clustering technique that provides the best outcomes. For that reason, four clustering techniques have been employed for grouping the energy consumption information.

Table 1 contains corresponding values of Silhouette coefficient for eight buildings in the same way as we did in preceding works (Ruiz, et al., 2019; Ruiz, et al., 2016; Ruiz, et al., 2018). These four methods were k-Medoids, k-Means, Hierarchical clustering and Gaussian Mixture models, with different number of clusters.

A first glance, GM is the model which achieves the worst results in almost all scenarios. Besides, this model shows a much more varying behavior. Although it reaches the best results for k=2 and k=3 in building B7, this is probably because this building manifests a much more differentiated tendency split into two and three groups, and this fact make easier for GM to cluster that consumption. This situation also happens in building B6 with 15 and 20 clusters. But, in the rest of the test performed that technique presents no such good coefficient, because it cannot converge to a local minimum. Consequently, for its irregular behavior and its bad performance in the current test battery, we exclude this method from further experiments.

Hierarchical clustering is the following method that achieves the worst results, although it shows similar score to k-Means and k-Medoids in some cases, and scores the best with building B1 with 7 clusters and building B4 with 4 clusters. This method furnishes very interesting information which may be useful for a deeper analysis on data grouping so that it could be used for visualizing how samples are arranged into the same group. Another difference between this method and the other two is that Hierarchical clustering gets worse Silhouette coefficient as the number of clusters is increased, this fact may be better observed in Figure 8.

Table 1. Silhouette coefficient values for k-Medoids, k-Means, Hierarchical and Gaussian Mixtures Clustering methods setting different number of clusters in eight different buildings.

Table 2. Silhouette coefficient values for k-Medoids, k-Means, Hierarchical and Gaussian Mixtures Clustering methods setting different number of clusters in eight different buildings.

Test	Number of clusters (k)											
	2	3	4	5	6	7	8	9	10	15	20	30
Building 1												
kMedoids	0.44127	0.35265	0.30122	0.26834	0.27238	0.25816	0.25339	0.23917	0.23484	0.21769	0.18486	0.16420
kMeans	**0.44204**	**0.35576**	**0.31374**	**0.29206**	**0.29336**	0.27016	**0.25963**	**0.25804**	**0.24602**	**0.22701**	**0.20575**	**0.17528**
Hierar.	0.43060	0.32504	0.28742	0.27955	0.28575	**0.27572**	0.21764	0.22055	0.19920	0.19099	0.14739	0.13154
GM	0.19641	0.14545	0.01452	0.09493	0.04693	0.01430	0.03068	0.08340	0.03042	0.01557	0.03884	-0.00465
Building 2												
kMedoids	**0.54697**	**0.45307**	0.40098	0.35662	0.32823	**0.31363**	0.30173	0.28986	0.29241	**0.27390**	0.25902	**0.25166**
kMeans	0.54281	0.44581	**0.42437**	**0.37527**	**0.35739**	0.31020	**0.31899**	**0.31863**	**0.30360**	0.26743	**0.26892**	0.24914
Hierar.	0.54400	0.38352	0.40556	0.31041	0.28870	0.27540	0.27851	0.26796	0.27038	0.22822	0.22938	0.24324
GM	0.27122	0.19579	0.24141	0.09761	0.17816	0.10074	-0.02699	0.10827	0.10431	0.10783	0.14430	0.09147
Building 3												
kMedoids	**0.61912**	**0.62109**	**0.60515**	**0.61472**	0.57661	0.57033	0.55373	**0.54617**	0.51530	0.47496	0.45810	0.44791
kMeans	0.61907	0.61972	0.60413	0.60519	**0.58930**	**0.57078**	**0.55780**	0.53415	**0.52344**	**0.47626**	**0.47197**	**0.45490**
Hierar.	0.60721	0.59974	0.58629	0.58983	0.57992	0.56603	0.50803	0.51087	0.48930	0.47062	0.44617	0.44444
GM	0.27079	0.22623	0.14682	0.35104	0.19483	0.15633	0.32806	0.26766	0.27965	0.05403	-0.19083	-0.26192
Building 4												
kMedoids	**0.53243**	0.54704	0.50326	**0.52292**	**0.52287**	**0.52819**	**0.52608**	**0.52764**	**0.53045**	0.52292	0.51525	**0.50706**
kMeans	0.53198	**0.54737**	0.50942	0.50895	0.49593	0.50516	0.48512	0.49836	0.49885	**0.52508**	**0.52395**	0.48659
Hierar.	0.51755	0.52434	**0.51087**	0.51346	0.48603	0.49942	0.49912	0.50271	0.50232	0.51948	0.49417	0.49029
GM	-0.18505	-0.21564	0.04310	0.12873	0.24303	0.31653	0.24840	0.21877	0.10188	0.13289	0.10820	0.06728
Building 5												
kMedoids	**0.51936**	**0.45138**	0.39333	0.39654	**0.36939**	0.37233	**0.37022**	0.35519	**0.36631**	0.32009	**0.33462**	0.31133
kMeans	0.51837	0.43078	**0.39575**	**0.40058**	0.36074	**0.37968**	0.36834	**0.36179**	0.35555	**0.34489**	0.32664	**0.31894**
Hierar.	0.49333	0.37871	0.37684	0.36061	0.35644	0.34762	0.34789	0.29630	0.30062	0.30190	0.29178	0.28849
GM	0.42979	0.39011	0.38452	0.19536	0.28463	0.26423	0.26833	0.21045	0.21994	0.27239	0.18394	0.19542
Building 6												
kMedoids	0.55598	0.48409	**0.47569**	**0.48308**	0.47217	0.43059	0.39512	**0.37566**	**0.37127**	0.31538	0.29840	**0.28812**
kMeans	**0.55765**	0.48473	0.45510	0.47234	0.45427	**0.43064**	**0.40914**	0.37481	0.36905	0.31129	0.30443	0.28245
Hierar.	0.55702	0.48896	0.46872	0.47142	**0.47395**	0.39681	0.34916	0.31690	0.31206	0.30152	0.29835	0.24944
GM	0.48531	**0.51226**	0.39349	0.27960	0.09788	0.13895	0.25847	0.17072	0.23002	**0.34352**	**0.31322**	0.18810
Building 7												
kMedoids	0.70427	0.51683	**0.49585**	0.46256	0.40832	0.36795	0.32396	0.31757	0.31429	0.27442	0.24990	0.23758
kMeans	0.70436	0.51545	0.49484	**0.46406**	**0.41210**	**0.39529**	**0.35761**	**0.32591**	**0.33314**	**0.28732**	**0.26567**	**0.25215**
Hierar.	0.69915	0.48837	0.47744	0.45658	0.35772	0.31063	0.31536	0.31468	0.31560	0.25125	0.24454	0.20498
GM	**0.71866**	**0.69913**	0.47914	0.43430	0.33915	0.30568	0.33884	0.11113	0.07371	0.07404	0.04922	-0.00444
Building 8												
kMedoids	0.46312	0.37383	0.35105	**0.35520**	**0.35745**	**0.34988**	**0.35390**	**0.35677**	**0.36334**	0.34902	0.33695	**0.31194**
kMeans	**0.46432**	**0.37736**	**0.36167**	0.35180	0.35682	0.34944	0.34927	0.35209	0.35614	**0.35909**	**0.33697**	0.30796
Hierar.	0.39864	0.35169	0.36064	0.33213	0.32437	0.34098	0.33162	0.32821	0.32677	0.32354	0.29340	0.25812
GM	0.23177	0.24488	0.18190	0.14170	0.12377	0.01289	0.04458	-0.00589	-0.06063	-0.16940	-0.14097	-0.00324

Best results are obtained with k-Means and k-Medoids. Those two methods prove a very similar behavior and their results are closely related (see Figure 8). In many cases, k-Means obtain better score just for few thousandths. This situation may be caused because k-Means does not employ actual points and then data may be accomplished for a better grouping. However, k-Means is not as robust to outliers as k-Medoids as the representative point of each cluster is a data point inside the cluster and hence a particular day may be representative of a consumption pattern. For this reason and since there are no significant differences between them, then the k-Medoids method have been chosen for optimizing its parameters in order to further improve its score and therefore obtain better energy consumption groups.

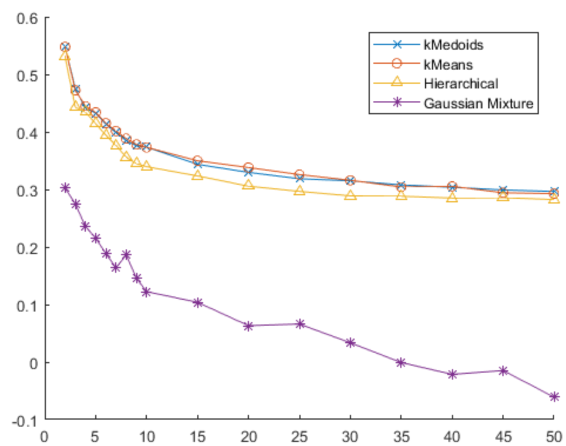


Figure 8. Average of the Silhouette coefficient of the four techniques for different number of clusters.

Table 2 reports the Silhouette Coefficients of using k-Medoid method as clustering technique but employing 5 different distances (Squared Euclidean, Cosine, Hamming, Correlation and City Block) and 4 variations of the k-Medoids approach (PAM, Small, CLARA, Large). In the same way, these tests have been executed 8 times and the average of its results was computed.

From Table 2, one may observe that the distance which attains the best results in all cases is the Square Euclidean. The next best distance is the City Block which achieves positive coefficients up to 0.3 in all instances. This one is followed by the Cosine distance which has very similar results to Correlation distance, but in this case those two distances manifest a Silhouette coefficient about null. In other words, there is no good clustering when data are on the border of two natural groups. Finally, Hamming distance, as expected, gets the worst results. In all test it indicates that samples of the cluster have been assigned to the wrong cluster.

Table 3. Comparison of Silhouette Coefficients among different k-Medoids methods using Squared Euclidean, Cosine, Hamming, Correlation and Cityblock distances.

Test	Method			
	PAM	Small	CLARA	Large
Building 1				
SEuclidean	**0.370767**	**0.374352**	**0.355470**	**0.378641**
Cosine	0.003703	0.026772	0.036079	0.015752
Hamming	-0.354830	-0.362750	-0.380248	-0.373617
Correlation	-0.118380	-0.122095	-0.122368	-0.118845
Cityblock	0.337596	0.326648	0.327708	0.340988
Building 2				
SEuclidean	**0.433886**	**0.426087**	**0.407519**	**0.436019**
Cosine	0.134271	0.111313	0.132805	0.150079
Hamming	-0.327733	-0.415255	-0.349293	-0.332920
Correlation	-0.039631	-0.024842	-0.031563	-0.020744
Cityblock	0.350316	0.353753	0.349109	0.352138
Building 3				
SEuclidean	**0.515248**	**0.564987**	**0.533020**	**0.512428**
Cosine	0.314941	0.304684	0.254124	0.321091
Hamming	-0.322504	-0.322221	-0.395272	-0.344482
Correlation	0.153388	0.061183	0.134331	0.146379
Cityblock	0.370124	0.419148	0.406508	0.366430
Building 4				
SEuclidean	**0.619327**	**0.611023**	**0.604774**	**0.613401**
Cosine	0.116974	0.199454	0.166437	0.141759
Hamming	-0.463068	-0.437350	-0.465271	-0.436506
Correlation	-0.226191	-0.219497	-0.224021	-0.220027
Cityblock	0.570980	0.557400	0.586834	0.553310
Building 5				
SEuclidean	**0.505906**	**0.497783**	**0.478890**	**0.476475**
Cosine	0.121742	0.051292	0.160880	0.105356
Hamming	-0.334818	-0.359787	-0.398228	-0.332779
Correlation	-0.070437	-0.040160	0.013851	-0.080705
Cityblock	0.414498	0.417235	0.411978	0.436053
Building 6				
SEuclidean	**0.544521**	**0.499667**	**0.491220**	**0.497870**
Cosine	-0.149090	-0.219445	-0.258559	-0.187748
Hamming	-0.708095	-0.671854	-0.664089	-0.618530
Correlation	-0.224318	-0.220000	-0.244297	-0.222777
Cityblock	0.420878	0.416523	0.441077	0.452321
Building 7				
SEuclidean	**0.453018**	**0.467313**	**0.467227**	**0.445960**
Cosine	-0.017338	-0.031254	-0.044617	-0.038263
Hamming	-0.577349	-0.661232	-0.684352	-0.694103
Correlation	0.256202	0.249247	0.259983	0.251625
Cityblock	0.385672	0.386050	0.376613	0.389255
Building 8				
SEuclidean	**0.507361**	**0.493785**	**0.475112**	**0.491974**
Cosine	0.020378	0.011505	0.019555	0.018918
Hamming	-0.539601	-0.578973	-0.590736	-0.588974
Correlation	-0.158841	-0.150763	-0.138819	-0.149151
Cityblock	0.473578	0.465422	0.439933	0.470881

Furthermore, as may be appreciated on Table 2, several experiments have been carried out through different k-Medoids methods. Table 3 summarizes the results by selecting the best distance (the Squared Euclidean function). Although Table 3 confirms that there is no clear method obtaining the best Silhouette Coefficient, the Partitioning Around Medoid algorithm achieves the best grouping in half of the cases. This is then followed by the Small algorithm which responds in a very similar way to PAM algorithm, yet only reaches the best score in two instances. Also, it is important to note that the Small algorithm is not recommendable for large scale applications as it performs a more exhaustive search of medoids. Thirdly, the Large algorithm acquire the best coefficient in two cases, but its performance is less precise than the previous methods as it examines only a random sample of cluster members. Finally, the worst algorithm to find medoids is CLARA, the other large scale method. All these results and its differences are better appraised in Figure 9 where there is a little difference between PAM and Small algorithms, followed by the large scale and CLARA methods.

Table 4. Summary of the results using the best distance for k-Medoids algorithms.

Test	Method			
	PAM	Small	CLARA	Large
Building 1	0.370767	0.374352	0.355470	**0.378641**
Building 2	0.433886	0.426087	0.407519	**0.436019**
Building 3	0.515248	**0.564987**	0.533020	0.512428
Building 4	**0.619327**	0.611023	0.604774	0.613401
Building 5	**0.505906**	0.497783	0.478890	0.476475
Building 6	**0.544521**	0.499667	0.491220	0.497870
Building 7	0.453018	**0.467313**	0.467227	0.445960
Building 8	**0.507361**	0.493785	0.475112	0.491974

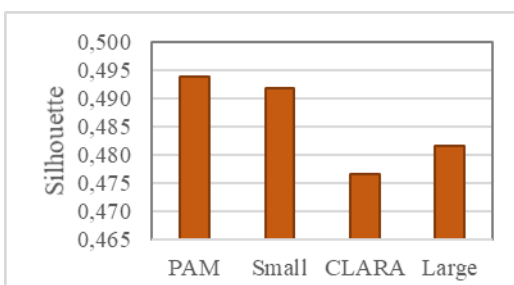


Figure 9. Silhouette coefficient mean of all test buildings using the best distance obtained.

As a result, our final proposal for an algorithm is to perform the clustering k-Medoids using the PAM algorithm and the Squared Euclidean distance as metric for evaluating distances among points. And hence, with the aim of contrasting the first solution, Table 4 compares the best results obtained using any of the four methods and the k-Medoids with its parameters optimized. It can be noted that the method has successfully passed all test performed except building B7 with 3 clusters whose coefficient is 2.73% worse. However, the optimized k-Medoids reaches an average improvement rate up to 34.35%. In the best case it achieves an improvement rate up to 57.77%, obtained with building B1 and 8 clusters. And the worst enhancement ratio is acquired with building B4 and using 10 clusters.

Table 5. Comparison between the first approximation carried out and the best k-Medoids algorithm using Squared Euclidean distance, for different number of clusters (k).

k	Building 1		Building 2		Building 3		Building 4		Building 5		Building 6		Building 7		Building 8	
	Table 1	PAM	Table 1	PAM	Table 1	PAM	Table 1	PAM	Table 1	PAM	Table 1	PAM	Table 1	PAM	Table 1	PAM
2	0.4420	**0.6407**	0.5470	**0.7256**	0.6191	**0.7700**	0.5324	**0.6408**	0.5194	**0.6978**	0.5576	**0.7312**	0.7187	**0.8697**	0.4643	**0.6534**
3	0.3558	**0.5290**	0.4531	**0.6191**	0.6211	**0.7339**	0.5474	**0.6899**	0.4514	**0.6355**	0.5123	**0.6528**	**0.6991**	0.6800	0.3774	**0.5392**
4	0.3137	**0.4488**	0.4244	**0.5674**	0.6051	**0.7212**	0.5109	**0.6221**	0.3958	**0.5659**	0.4757	**0.6310**	0.4959	**0.6731**	0.3617	**0.4771**
5	0.2921	**0.4266**	0.3753	**0.5241**	0.6147	**0.7270**	0.5229	**0.6051**	0.4006	**0.5537**	0.4831	**0.6493**	0.4641	**0.6346**	0.3552	**0.5139**
6	0.2934	**0.4329**	0.3574	**0.4720**	0.5893	**0.7027**	0.5229	**0.6086**	0.3694	**0.5079**	0.4740	**0.6526**	0.4121	**0.6013**	0.3574	**0.5218**
7	0.2757	**0.4213**	0.3136	**0.4625**	0.5708	**0.6883**	0.5282	**0.6132**	0.3797	**0.5204**	0.4306	**0.6020**	0.3953	**0.5389**	0.3499	**0.4810**
8	0.2596	**0.4096**	0.3190	**0.4582**	0.5578	**0.6669**	0.5261	**0.6178**	0.3702	**0.5079**	0.4091	**0.5497**	0.3576	**0.4800**	0.3539	**0.4921**
9	0.2580	**0.3939**	0.3186	**0.4332**	0.5462	**0.6621**	0.5276	**0.6132**	0.3618	**0.5145**	0.3757	**0.5186**	0.3259	**0.4671**	0.3568	**0.5041**
10	0.2460	**0.3773**	0.3036	**0.4260**	0.5234	**0.6206**	0.5305	**0.5992**	0.3663	**0.4665**	0.3713	**0.5193**	0.3331	**0.4472**	0.3633	**0.5004**
15	0.2270	**0.3443**	0.2739	**0.4215**	0.4763	**0.5740**	0.5251	**0.6156**	0.3449	**0.4419**	0.3435	**0.4392**	0.2873	**0.3905**	0.3591	**0.4842**
20	0.2057	**0.3039**	0.2689	**0.4049**	0.4720	**0.5574**	0.5240	**0.6013**	0.3346	**0.4651**	0.3132	**0.4119**	0.2657	**0.3702**	0.3370	**0.4697**
25	0.1832	**0.2747**	0.2578	**0.3856**	0.4606	**0.5503**	0.5184	**0.5950**	0.3243	**0.4544**	0.2928	**0.4012**	0.2548	**0.3752**	0.3155	**0.4378**
30	0.1753	**0.2745**	0.2517	**0.3917**	0.4549	**0.5485**	0.5071	**0.5894**	0.3189	**0.4438**	0.2881	**0.4017**	0.2522	**0.3237**	0.3119	**0.4326**

Once all experiments have been done, the final step is to visualize clustering results and draw conclusions about how the energy is being consumed. With this aim in mind, Figure 10 illustrates an example of the distribution of the consumption for building B1 by setting up 5 clusters. Thus, 5 different energy patterns are found. As can be seen, the first 8 hours of the day the consumption held constant, then it increases between 10h and 12h where reaches a peak and it falls slightly till 15h. At that point consumption starts growing again but with less strength and soon falls once more until night has come. This behavior is clear in this building with different levels of growth. Cluster 3, green line, introduces the energy pattern corresponding to the highest consumption days and this pattern is very common in days from Monday to Thursday in almost all months, with also some Fridays of January, June, July, September and November presenting this very same pattern. However, it is unusual in May and

October, but usual only in some Wednesdays and Tuesdays of March and April. This pattern could be noted as High-level consumption.

Cluster 4 is the next higher consumption in this building. In this case, all working days (from Monday to Friday) follow that consumption pattern. It seems that the number of Fridays decreases from September to December and even in July some Sundays had that consumption. This pattern could be called as Middle-high-level consumption.

The previous clusters represent high part of the working days of the week. Now, cluster 2 which may be entitled as Middle-level consumption is representative of Sundays and Saturdays of July, and that consumption is repeated over the year during all the week. This means, on the one hand, that there are some days during each month where consumption increases on weekends; and, on the other hand, consumption decreases from Monday to Friday so that they consume like weekend. It depends on how one looks at this, as Cluster 5 and 1 manifest the energy behavior of most of the weekends. Also holydays should present similar behavior, this seems to be the case for Thursdays of February or Tuesday of May.

More information could be extracted from this data by increasing the number of clusters, however, as Figure 8 indicated, 5 to 10 groups is the better option in this case. And also for space constraints no all buildings were shown and analyzed in this manuscript. We hope the results presented here are representative of the proposed methodology. All results may be consulted and downloaded from Google Drive at: https://drive.google.com/drive/folders/1eOmjmeGjVdNjFGNQ_2D31ar5dZVbRaov

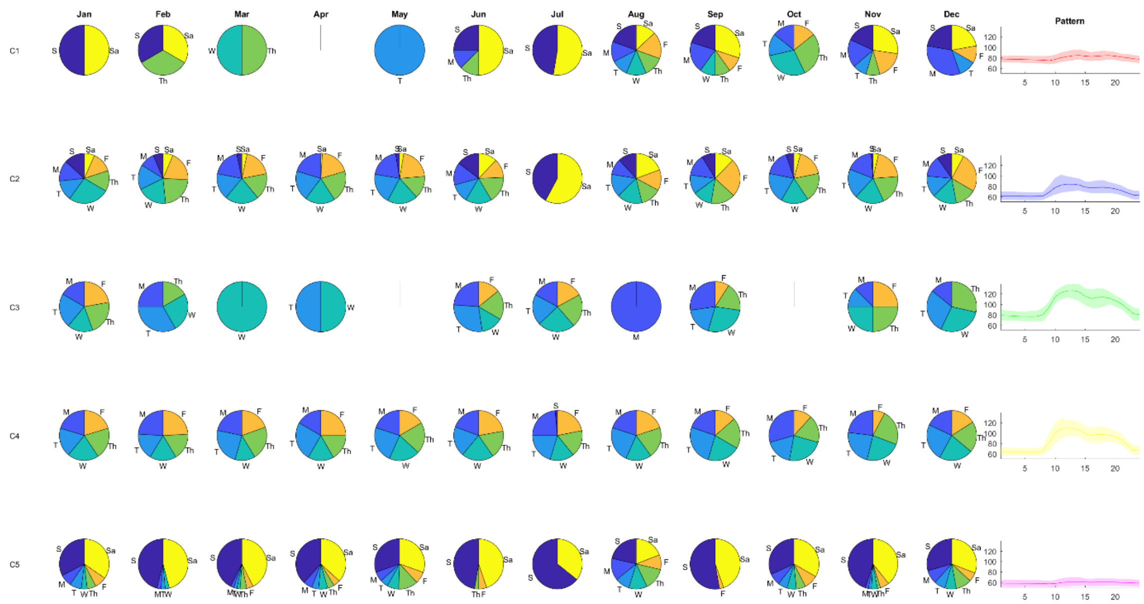


Figure 10. Example of the week-day distribution over the monthly period according to its cluster pattern for a specific building

4. Conclusions

In this manuscript we put forward a methodology for analysing energy consumption in a set of buildings using times series techniques. In particular, we focused on clustering techniques so as to extract patterns in consumption and thus, to understand in a better way how buildings are consuming. It is important to emphasize that one of the key factors for integrating systems and making buildings more intelligent is the energy savings which may be obtained, and a clear understanding of building behavior is necessary if one wants to follow energy strategies.

As a result, our methodology proposed a framework for analysing and discovering energy patterns in consumption by using clustering techniques. Our clustering analysis is a helpful tool for this purpose in which a whole process is carried out from raw data obtained in buildings' sensors to knowledge in useful way. The main idea is to find several frequent patterns in data and use them for describing a building energy profile. According to that building information, one may discover that energy consumption in grouped into several categories and depending on the season and the weekday those consumptions are clustered differently.

Furthermore, this study lays the foundations for forthcoming work which will develop a fuzzy system for evaluating the performance of a specific building in order to determine and classify similar behavior in different facilities. Figure 10 have already anticipated how to develop such a system, as 5 membership functions were defined (low, middle-low, middle, middle-high and high consumption). Future work will be also focused on taking advantage of this inter-cluster and intra-cluster information in order to improve preceding works on energy forecasting and finally to implement a complete decision-support system for an adequate energy management and thus to combine all the available knowledge for achieving energy saving.

5. Acknowledgments

This work has been developed with the support of the Department of Computer Science and Artificial Intelligence of the University of Granada, TIC111, and the project TIN201564776-C3-1-R.

6. Abbreviations

TSC	Time-Series Clustering
GM	Gaussian Mixtures.
PAM	Partitioning Around Medoids.
CLARA	Clustering LARge Applications.

7. Bibliography

Alves, W. A. L., Santos, S. D., & Schimit, P. H. T. (2016). Hierarchical Clustering Based on Reports Generated by Scriptlattes. In (pp. 28-35). Cham: Springer International Publishing. [http://dx.doi.org/10.1007/978-3-319-51133-7_4]

Andrade, J. R., & Bessa, R. J. (2017). Improving Renewable Energy Forecasting with a Grid of Numerical Weather Predictions. *IEEE Transactions on Sustainable Energy, PP*, 1-1. [http://dx.doi.org/10.1109/TSTE.2017.2694340]

Arthur, D., & Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms* (pp. 1027-1035): Society for Industrial and Applied Mathematics. [http://dl.acm.org/citation.cfm?id=128338]

Ashouri, M., Haghighat, F., Fung, B. C. M., Lazrak, A., & Yoshino, H. (2018). Development of building energy saving advisory: A data mining approach. *Energy and Buildings, 172*, 139-151. [http://dx.doi.org/10.1016/j.enbuild.2018.04.052]

Belaïd, F., Roubaud, D., & Galariotis, E. (2019). Features of residential energy consumption: Evidence from France using an innovative multilevel modelling approach. *Energy Policy, 125*, 277-285. [http://dx.doi.org/10.1016/j.enpol.2018.11.007]

Belley, C., Gaboury, S., Bouchard, B., & Bouzouane, A. (2015). Nonintrusive system for assistance and guidance in smart homes based on electrical devices identification. *Expert Systems with Applications, 42*, 6552-6577. [http://dx.doi.org/10.1016/j.eswa.2015.04.024]

Berardi, U. (2015). Building Energy Consumption in US, EU, and BRIC Countries. *Procedia Engineering, 118*, 128-136. [http://dx.doi.org/10.1016/j.proeng.2015.08.411]

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*: John Wiley & Sons. [https://books.google.es/books?hl=es&lr=&id=rNt5CgAAQBAJ&oi=fnd&pg=PR7&dq=Time+Series+Analysis:+Forecasting+and+Control.+3rd+ed.+Englewood+Cliffs,+NJ:+Prentice+Hall,+1994&ots=DJ65tVo-OF&sig=ujT9CshDrDbsWLIhxcgFBpUBJ2M#v=onepage&q=autocorrelation&f=false]

Cao, X., Dai, X., & Liu, J. (2016). Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy and Buildings, 128*, 198-213. [http://dx.doi.org/10.1016/j.enbuild.2016.06.089]

Capozzoli, A., Lauro, F., & Khan, I. (2015). Fault detection analysis using data mining techniques for a cluster of smart office buildings. *Expert Systems with Applications, 42*, 4324-4338. [http://dx.doi.org/10.1016/j.eswa.2015.01.010]

Capozzoli, A., Piscitelli, M. S., Brandi, S., Grassi, D., & Chicco, G. (2018). Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings. *Energy, 157*, 336-352. [http://dx.doi.org/10.1016/j.energy.2018.05.127]

Celeux, G., & Govaert, G. (1995). Gaussian parsimonious clustering models. *Pattern Recognition, 28*, 781-793. [http://dx.doi.org/10.1016/0031-3203(94)00125-6]

Chen, J., Taylor, J. E., & Wei, H.-H. (2012). Modeling building occupant network energy consumption decision-making: The interplay between network structure and conservation. *Energy and Buildings, 47*, 515-524. [http://dx.doi.org/10.1016/j.enbuild.2011.12.026]

D'Agostino, D., Cuniberti, B., & Bertoldi, P. (2017). Energy consumption and efficiency technology measures in European non-residential buildings. *Energy and Buildings, 153*, 72-86. [http://dx.doi.org/10.1016/j.enbuild.2017.07.062]

Deb, C., Zhang, F., Yang, J., Lee, S. E., & Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews, 74*, 902-924. [http://dx.doi.org/10.1016/j.rser.2017.02.085]

Dhanachandra, N., Manglem, K., & Chanu, Y. J. (2015). Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm. *Procedia Computer Science, 54*, 764-771. [http://dx.doi.org/10.1016/j.procs.2015.06.090]

Diansheng, G., Jin, C., MacEachren, A. M., & Ke, L. (2006). A Visualization System for Space-Time and Multivariate Patterns (VIS-STAMP). *IEEE Transactions on Visualization and Computer Graphics, 12*, 1461-1474. [http://dx.doi.org/10.1109/TVCG.2006.84]

Dias, J. G., & Ramos, S. B. (2014). Dynamic clustering of energy markets: An extended hidden Markov approach. *Expert Systems with Applications, 41*, 7722-7729. [http://dx.doi.org/10.1016/j.eswa.2014.05.030]

Dietmair, A., & Verl, A. (2009). A generic energy consumption model for decision making and energy efficiency optimisation in manufacturing. *International Journal of Sustainable Engineering, 2*, 123-133. [http://dx.doi.org/10.1080/19397030902947041]

Duque-Pintor, F., Fernández-Gómez, M., Troncoso, A., & Martínez-Álvarez, F. (2016). A New Methodology Based on Imbalanced Classification for Predicting Outliers in Electricity Demand Time Series. *Energies, 9*, 752-752. [http://dx.doi.org/10.3390/en9090752]

Duwairi, R., & Abu-Rahmeh, M. (2015). A novel approach for initializing the spherical K-means clustering algorithm. *Simulation Modelling Practice and Theory, 54*, 49-63. [http://dx.doi.org/10.1016/j.simpat.2015.03.007]

Espinoza, M., Joye, C., Belmans, R., & Moor, B. D. (2005). Short-term load forecasting, profile identification, and customer segmentation: a methodology based on periodic time series. *IEEE Transactions on Power Systems, 20*, 1622-1630. [http://dx.doi.org/10.1109/TPWRS.2005.852123]

Fabregas, A. C., Gerardo, B. D., & Tanguilig III, B. T. (2017). Enhanced initial centroids for k-means algorithm. *International Journal of Information Technology and Computer Science, 1*, 26-33. [http://pdfs.semanticscholar.org/9e9b/97a1f2d8b28e8b0de4abab25365e2e40776a.pdf]

Fumo, N., & Rafe Biswas, M. A. (2015). Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews, 47*, 332-343. [http://dx.doi.org/10.1016/j.rser.2015.03.035]

Gabaldon, A., Guillamon, A., Ruiz, M., Valero, S., Alvarez, C., Ortiz, M., & Senabre, C. (2010). Development of a methodology for clustering electricity-price series to improve customer response initiatives. *IET generation, transmission & distribution, 4*, 706-715. [http://dx.doi.org/10.1049/iet-gtd.2009.0112]

Gao, X., & Malkawi, A. (2014). A new methodology for building energy performance benchmarking: An approach based on intelligent clustering algorithm. *Energy and Buildings, 84*, 607-616. [http://dx.doi.org/10.1016/j.enbuild.2014.08.030]

Gutiérrez-Batista, K., Campaña, J. R., Vila, M.-A., & Martin-Bautista, M. J. (2018). Building a contextual dimension for OLAP using textual data from social networks. *Expert Systems with Applications, 93*, 118-133. [http://dx.doi.org/10.1016/j.eswa.2017.10.012]

Hong, M., Yu, X., Yu, N., & Loparo, K. A. (2016). An Energy Scheduling Algorithm Supporting Power Quality Management in Commercial Building Microgrids. *IEEE Transactions on Smart Grid, 7*, 1044-1056. [http://dx.doi.org/10.1109/TSG.2014.2379582]

Hong, T., Sun, H., Chen, Y., Taylor-Lange, S. C., & Yan, D. (2016). An occupant behavior modeling tool for co-simulation. *Energy and Buildings, 117*, 272-281. [http://dx.doi.org/https://doi.org/10.1016/j.enbuild.2015.10.033]

Iglesias, F., & Kastner, W. (2013). Analysis of Similarity Measures in Times Series Clustering for the Discovery of Building Energy Patterns. *Energies, 6*, 579-597. [http://dx.doi.org/10.3390/en6020579]

Jiang, G., Wang, W., & Zhang, W. (2019). A novel distance measure for time series: Maximum shifting correlation distance. *Pattern Recognition Letters, 117*, 58-65. [http://dx.doi.org/10.1016/j.patrec.2018.11.013]

Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis* (Vol. 344): John Wiley & Sons. [http://onlinelibrary.wiley.com/doi/book/10.1002/9780470316801]

Levesque, A., Pietzcker, R. C., Baumstark, L., De Stercke, S., Grübler, A., & Luderer, G. (2018). How much energy will buildings consume in 2100? A global perspective within a scenario framework. *Energy, 148*, 514-527. [http://dx.doi.org/10.1016/j.energy.2018.01.139]

López-Ruiz, A., Bergillos, R. J., & Ortega-Sánchez, M. (2016). The importance of wave climate forecasting on the decision-making process for nearshore wave energy exploitation. *Applied Energy, 182*, 191-203. [http://dx.doi.org/10.1016/j.apenergy.2016.08.088]

Maugis, C., Celeux, G., & Martin-Magniette, M. L. (2009). Variable selection for clustering with Gaussian mixture models. *Biometrics, 65*, 701-709. [http://dx.doi.org/10.1111/j.1541-0420.2008.01160.x]

Molina-Solana, M., Ros, M., Ruiz, M. D., Gómez-Romero, J., & Martin-Bautista, M. J. (2017). Data science for building energy management: A review. *Renewable and Sustainable Energy Reviews, 70*, 598-609. [http://dx.doi.org/10.1016/j.rser.2016.11.132]

Nagy, G. I., Barta, G., Kazi, S., Borbély, G., & Simon, G. (2016). GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach. *International Journal of Forecasting, 32*, 1087-1093. [http://dx.doi.org/10.1016/j.ijforecast.2015.11.013]

Park, H.-S., & Jun, C.-H. (2009). A simple and fast algorithm for K-medoids clustering. *Expert Systems with Applications, 36*, 3336-3341. [http://dx.doi.org/10.1016/j.eswa.2008.01.039]

Prabhakar, S. K., & Rajaguru, H. (2016). Code Converters with City Block Distance Measures for Classifying Epilepsy from EEG Signals. *Procedia Computer Science, 87*, 5-11. [http://dx.doi.org/10.1016/j.procs.2016.05.118]

Romanchenko, D., Kensby, J., Odenberger, M., & Johnsson, F. (2018). Thermal energy storage in district heating: Centralised storage vs. storage in thermal inertia of buildings. *Energy Conversion and Management, 162*, 26-38. [http://dx.doi.org/10.1016/j.enconman.2018.01.068]

Ruiz, L. G. B., Capel, M. I., & Pegalajar, M. C. (2019). Parallel memetic algorithm for training recurrent neural networks for the energy efficiency problem. *Applied Soft Computing, 76*, 356-368. [http://dx.doi.org/10.1016/j.asoc.2018.12.028]

Ruiz, L. G. B., Cuellar, M. P., Delgado, M., & Pegalajar, M. C. (2016). An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings. *Energies, 9*, 21. [http://dx.doi.org/10.3390/en9090684]

Ruiz, L. G. B., Rueda, R., Cuéllar, M. P., & Pegalajar, M. C. (2018). Energy consumption forecasting based on Elman neural networks with evolutive optimization. *Expert Systems with Applications, 92*, 380-389. [http://dx.doi.org/10.1016/j.eswa.2017.09.059]

Simmhan, Y., & Noor, M. U. (2013). Scalable prediction of energy consumption using incremental time series clustering. In *2013 IEEE International Conference on Big Data* (pp. 29-36): IEEE. [http://dx.doi.org/10.1109/BigData.2013.6691774]

Smith, S. W. (1997). Moving average filters. *The scientist and engineer's guide to digital signal processing*, 277-284. [http://www.dspguide.com/ch15.htm]

Strantzali, E., & Aravossis, K. (2016). Decision making in renewable energy investments: A review. *Renewable and Sustainable Energy Reviews, 55*, 885-898. [http://dx.doi.org/10.1016/j.rser.2015.11.021]

Titsias, M. K., & Yau, C. (2017). The Hamming Ball Sampler. *Journal of the American Statistical Association, 112*, 1598-1611. [http://dx.doi.org/10.1080/01621459.2016.1222288]

Torres, J. F., Troncoso, A., Koprinska, I., Wang, Z., & Martínez-Álvarez, F. (2018). Deep Learning for Big Data Time Series Forecasting Applied to Solar Power. In (pp. 123-133). Cham: Springer International Publishing. [http://dx.doi.org/10.1007/978-3-319-94120-2_12]

Warren Liao, T. (2005). Clustering of time series data—a survey. *Pattern Recognition, 38*, 1857-1874. [http://dx.doi.org/10.1016/j.patcog.2005.01.025]

Wijk, J. J. V., & Selow, E. R. V. (1999). Cluster and calendar based visualization of time series data. In *Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis'99)* (pp. 4-9). [http://dx.doi.org/10.1109/INFVIS.1999.801851]

Yousef, A., & Moghadam Charkari, N. (2015). SFM: A novel sequence-based fusion method for disease genes identification and prioritization. *Journal of Theoretical Biology, 383*, 12-19. [http://dx.doi.org/10.1016/j.jtbi.2015.07.010]

Zhang, T., Siebers, P.-O., & Aickelin, U. (2012). A three-dimensional model of residential energy consumer archetypes for local energy policy design in the UK. *Energy Policy, 47*, 102-110. [http://dx.doi.org/10.1016/j.enpol.2012.04.027]

Zhang, X., Liu, J., Du, Y., & Lv, T. (2011). A novel clustering method on time series data. *Expert Systems with Applications, 38*, 11891-11900. [http://dx.doi.org/10.1016/j.eswa.2011.03.081]

CONCLUSIONES Y TRABAJOS FUTUROS

7 Conclusiones y trabajo futuro

Esta sección recoge de forma resumida todos los resultados alcanzados durante el desarrollo de la presente tesis doctoral a modo de conclusiones obtenidas de la misma. Además, se presentan algunas vías de investigación y trabajos futuros que podrían ser la continuación de esta tesis.

En primer lugar, hemos estudiado el problema de la eficiencia energética en edificios de dominio público. Concretamente, en las instalaciones asociadas al campus de la Universidad de Granada, el cual recoge distintos tipos de edificios como aulas de docencia, centros de investigación, centros de procesamiento, escuelas técnicas, facultades, entre otros. La UGR está compuesta de 25 centros docentes distribuidos en 7 campus en dos continentes —Europa y África— y tres ciudades —Granada, Ceuta y Melilla—.

Así, el primer objetivo conseguido ha sido desarrollar un procedimiento que nos permitiese tratar y pre-procesar toda la información procedente de los edificios. Para ello, propusimos una metodología para analizar y extraer información energética con el objetivo final de predecir el consumo energético de cada instalación. Los modelos estudiados han sido RNA los cuales demostraron ser una herramienta útil y precisa para predecir el consumo.

En segundo lugar, la hibridación de RNAs y AE probaron ser una solución potente para aumentar en gran medida la precisión que los modelos predictivos podían llegar a alcanzar. Junto con esto, el enfoque paralelo desarrollado del AG evidenció una gran ventaja en términos de tiempo de ejecución respecto a su diseño secuencial sin alterar la calidad de las soluciones obtenidas. Concluyendo así, que los AE es un método muy eficiente para optimizar los pesos de las RNAs en un problema de eficiencia energética.

En tercer lugar, se estudiaron métodos basados en agrupamiento para obtener patrones de consumo energético e información referente a los comportamientos más comunes de los edificios. Los resultados indicaron que nuestra propuesta es una exitosa herramienta para el análisis del consumo y descubrir conocimiento no obvio.

El último objetivo propuesto fue desarrollar un prototipo de software que recogiese de forma conjunta todo lo desarrollado en esta tesis. Uno de los nichos todavía por explotar en la literatura sobre la Eficiencia Energética es la implantación de los modelos teóricos en un escenario real. Como resultado, nuestra propuesta comprende la visualización del consumo energético de todos los edificios la UGR, y así centralizar en un solo software toda la información de los edificios distribuidos de la Universidad. Además, dotar de inteligencia al software por medio de los modelos predictivos y proporcionando en tiempo real la predicción del consumo junto con advertencias sobre posibles problemas en el futuro. Los resultados obtenidos demuestran que se puede obtener conocimiento preciso y muy valioso de las predicciones realizadas para la toma de decisiones, como, por ejemplo, detectando comportamientos anómalos en el consumo.

Del trabajo desarrollado en esta tesis doctoral aparecen nuevas y prometedoras líneas de investigación que podrían ser la continuación de la misma.

1. En primer lugar, la teoría de Conjuntos Difusos nos permite operar con información imprecisa, lo que permite representaciones de los datos expresados en un lenguaje más cercano a la comprensión humana [41]. Se propone utilizar estos métodos para describir patrones de consumo y clasificación de edificios según el mismo.

2. Por otro lado, las dependencias y asociaciones en los datos pueden ser descubiertos mediante el uso de algoritmos de minería de reglas de asociación. Las reglas de asociación difusas [42] expresan relaciones entre términos lingüísticos, tales como “salas llenas en los días soleados normalmente requieren menos calefacción”.

3. También existen técnicas centradas en el trabajo con datos temporales que son aplicables a los sensores de transmisión y los datos de energía continuos. Se propone por consiguiente aplicar estas técnicas para la búsqueda de eventos —habituales o anómalos—; y en la síntesis de series temporales mediante descripciones lingüísticas [43].
4. También se propone profundizar en la combinación de los métodos de agrupamiento con las técnicas de predicción ya que estos pueden proveer de información muy útil a los modelos predictivos y mejor su precisión.

Conclusions and future work

This section summarises all the results attained during the development of this doctoral thesis as a way of conclusions achieved from this project. Furthermore, some research lines and future work are proposed.

First of all, we studied the energy efficiency problem in public buildings. Particularly, the facilities related to the campus of the University of Granada. The UGR is made up of several kind of buildings, such as, lecture rooms, research centres, data centres, engineering schools, faculties, among others. The UGR is composed of 25 teaching centres distributed throughout 7 campuses in two continents —Europe and Africa— and three cities —Granada, Ceuta and Melilla—.

Hence, the first aim we attained was to develop a procedure which allowed us to treat and pre-process properly all the information from those buildings. To do so, we proposed a methodology for analysing and capturing energy data, with the last goal of predicting energy consumption in each facility. The studied models were Artificial Neural Networks which proved to be a very useful and accurate tool to forecast consumption.

Secondly, the hybridization of Artificial Neural Networks and Evolutionary Algorithms demonstrated to be a powerful solution to improve prediction in terms of accuracy. In addition to this, the parallel design of the Genetic Algorithm developed evidenced a great advantage in terms of time cost in comparison with its sequential design, and all this without affecting the quality of the solutions. Therefore, we could conclude that the Evolutionary Algorithms were an efficient method to optimize the weights of the Artificial Neural Networks within an energy efficiency problem.

Thirdly, we studied clustering-based methods to obtain energy consumption patterns and knowledge related to the most common energy behaviours of the buildings. The results indicated that our proposal is a successful tool so as to analyze the consumption and discover non-obvious knowledge.

The last objective proposed was to develop a software prototype that encompasses jointly all the work done in this thesis. One of the gaps about Energy Efficiency that is currently for exploitation in literature is the application of the theoretical model in a real scenario. As a result, our proposal comprises the energy visualization of all buildings of our University, and in this way, to centralize in a single software all the information related to the distributed buildings of the UGR. Besides, to provide an intelligent software by using the predictive models which give the prediction of the consumption in real time along with some warnings regarding possible problems in the near future. The results accomplished show that accurate valuable knowledge can be obtained by using the predictions for decision making, for instance, detecting unusual behaviours in the consumption.

From the work developed in this doctoral thesis new and promising lines of research emerge which may contribute to future work.

1. Firstly, the Fuzzy Sets theory is able to deal with imprecise information, which allows us to represent data in a closer language to the human comprehension. Therefore, we propose to employ these methods to describe patterns in consumption and classify buildings according to that information.
2. Secondly, dependencies and associations in data may be discover by means of data mining techniques, e.g., association rules. The fuzzy association rules express relationships among linguistic terms, such as, "jam-packed rooms in sunny days typically require less heating".
3. There are also some techniques focused on time-series that may be applied on transmission sensors and energy-related data. Consequently, we propose to apply these techniques so as to seek unusual or common

events; and also to summarize time-series by means of linguistic descriptions.

4. Finally, we propose to go into the combination of clustering and forecasting techniques as this hybridization may provide useful information to the predictive models, and thus, to improve their accuracy.

BIBLIOGRAFÍA

8 Bibliografía

1. Esling, P. and C. Agon, *Time-series data mining.* ACM Computing Surveys (CSUR), 2012. **45**(1): p. 1-34. [http://dx.doi.org/10.1145/2379776.2379788]
2. Martínez de Pisón, F., et al., *Minería de datos en series temporales para la búsqueda de conocimiento oculto en históricos de procesos industriales.* Actas del III Taller Nacional de Minería de Datos y Aprendizaje, TAMIDA, 2005: p. 31-38. [http://www.lsi.us.es/redmidas/CEDI/papers/143.pdf]
3. Saez, M., et al., *Métodos de series temporales en los estudios epidemiológicos sobre contaminación atmosférica.* Revista española de salud pública, 1999. **73**(2): p. 133-143. [http://ref.scielo.org/b8cf59]
4. Delgado, M., M.C. Pegalajar, and M.P. Cuéllar, *Memetic evolutionary training for recurrent neural networks: An application to time-series prediction.* Expert Systems, 2006. **23**(2): p. 99-114. [http://dx.doi.org/10.1111/j.1468-0394.2006.00327.x]
5. Delgado, M., M.C. Pegalajar, and M.P. Cuéllar, *Evolutionary training for Dynamical Recurrent Neural Networks: an application in finantial time series prediction.* Mathware and Soft Computing, 2006. **13**(2): p. 89-110. [http://eudml.org/doc/41874]
6. International Energy Agency. *Transition to Sustainable Building. Strategies and Opportunities to 2050.* 2013; [http://www.iea.org/publications/freepublications/publication/transition-to-sustainablebuildings.html]
7. Molina-Solana, M., et al., *Data science for building energy management: A review.* Renewable and Sustainable Energy Reviews, 2017. **70**: p. 598-609. [http://dx.doi.org/10.1016/j.rser.2016.11.132]
8. Cuéllar, M.P., M. Delgado, and M.C. Pegalajar, *Topology Optimization and Training of Recurrent Neural Networks with Pareto-Based Multi-objective Algorithms: A Experimental Study*, in *Computational and Ambient Intelligence: 9th International Work-Conference on Artificial Neural Networks, IWANN 2007, San Sebastián, Spain, June 20-22, 2007. Proceedings*, F. Sandoval, et al., Editors. 2007, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 359-366. [http://dx.doi.org/10.1007/978-3-540-73007-1_44]

9. Yuce, B., et al., *Utilizing artificial neural network to predict energy consumption and thermal comfort level: An indoor swimming pool case study.* Energy and Buildings, 2014. **80**: p. 45-56. [http://dx.doi.org/10.1016/j.enbuild.2014.04.052]

10. García, S., et al., *Evolutionary-based selection of generalized instances for imbalanced classification.* Knowledge-Based Systems, 2012. **25**(1): p. 3-12. [http://dx.doi.org/10.1016/j.knosys.2011.01.012]

11. Brockwell, P.J., R.A. Davis, and M.V. Calder, *Introduction to time series and forecasting*. Vol. 2. 2002: Springer. [http://dx.doi.org/10.1007/978-3-319-29854-2]

12. Chatfield, C., *The analysis of time series: an introduction*. 2003: Chapman and Hall/CRC.

13. Pengfei, L., et al. *The sharing services of map location based on network API*. in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2017. [http://dx.doi.org/10.1109/IGARSS.2017.8128395]

14. Qiu, X., et al., *Empirical Mode Decomposition based ensemble deep learning for load demand time series forecasting.* Applied Soft Computing, 2017. **54**: p. 246-255. [http://dx.doi.org/10.1016/j.asoc.2017.01.015]

15. Pegalajar, M.C., et al. *Artificial Neural Networks and Fuzzy Logic for Specifying the Color of an Image Using Munsell Soil-Color Charts*. 2018. Cham: Springer International Publishing.

16. Jia, C., et al., *Improving the Performance of Distributed TensorFlow with RDMA.* International Journal of Parallel Programming, 2017. [http://dx.doi.org/10.1007/s10766-017-0520-3]

17. Ekici, S., S. Yildirim, and M. Poyraz, *A transmission line fault locator based on Elman recurrent networks.* Applied Soft Computing, 2009. **9**(1): p. 341-347. [http://dx.doi.org/10.1016/j.asoc.2008.04.011]

18. Kaur, H. and S. Ahuja, *Time Series Analysis and Prediction of Electricity Consumption of Health Care Institution Using ARIMA Model*, in *Proceedings of Sixth International Conference on Soft Computing for Problem Solving: SocProS 2016, Volume 2*, K. Deep, et al., Editors. 2017, Springer Singapore: Singapore. p. 347-358. [http://dx.doi.org/10.1007/978-981-10-3325-4_35]

19. Nagy, G.I., et al., *GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach.* International Journal of Forecasting, 2016. **32**(3): p. 1087-1093. [http://dx.doi.org/10.1016/j.ijforecast.2015.11.013]

20. Yuan, C., S. Liu, and Z. Fang, *Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model.* Energy, 2016. **100**: p. 384-390. [http://dx.doi.org/10.1016/j.energy.2016.02.001]

21. Fumo, N. and M.A. Rafe Biswas, *Regression analysis for prediction of residential energy consumption.* Renewable and Sustainable Energy Reviews, 2015. **47**: p. 332-343. [http://dx.doi.org/10.1016/j.rser.2015.03.035]

22. Haydari, Z., et al. *Time-series load modelling and load forecasting using neuro-fuzzy techniques*. in *2007 9th International Conference on Electrical Power Quality and Utilisation*. 2007. [http://dx.doi.org/10.1109/EPQU.2007.4424201]

23. Lo Sciuto, G., et al., *Geometric Shape Optimization of Organic Solar Cells for Efficiency Enhancement by Neural Networks*, B. Eynard, et al., Editors. 2017, Springer International Publishing: Cham. p. 789-796. [http://dx.doi.org/10.1007/978-3-319-45781-9_79]

24. Miki, Y., et al., *Classification of teeth in cone-beam CT using deep convolutional neural network.* Computers in Biology and Medicine, 2017. **80**(September 2016): p. 24-29. [http://dx.doi.org/10.1016/j.compbiomed.2016.11.003]

25. Boomilingam, T. and M. Subramaniam, *Medical Image Feature Classification from Rule Mining and Retrieval via Weighted Swarm Neural Network.* Journal of Medical Imaging and Health Informatics, 2017. **7**(1): p. 111-117. [http://dx.doi.org/10.1166/jmihi.2017.1992]

26. Michelsanti, D., et al., *Fast Fingerprint Classification with Deep Neural Network.* Visapp - International Conference on Computer Vision Theory and Applications, 2017. [http://vbn.aau.dk/en/publications/fast-fingerprint-classification-with-deep-neural-network(e8c9084c-62fd-4cec-9fd8-f36a636fcb78).html]

27. Pino-Mejías, R., et al., *Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO2 emissions.* Energy, 2017. **118**: p. 24-36. [http://dx.doi.org/10.1016/j.energy.2016.12.022]

28. Piao, W., et al., *Development of a protocol to optimize electric power consumption and life cycle environmental impacts for operation of wastewater treatment plant.* Environmental Science and Pollution Research, 2016. **23**(24): p. 25451-25466. [http://dx.doi.org/10.1007/s11356-016-7771-1]

29. Beyer, L., A. Hermans, and B. Leibe, *DROW: Real-Time Deep Learning based Wheelchair Detection in 2D Range Data.* 2016. [http://arxiv.org/abs/1603.02636]

30. Neto, A.H. and F.A.S. Fiorelli, *Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption.* Energy and Buildings, 2008. **40**(12): p. 2169-2176. [http://dx.doi.org/10.1016/j.enbuild.2008.06.013]

31. Aydinalp-Koksal, M. and V.I. Ugursal, *Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption*

in the residential sector. Applied Energy, 2008. **85**(4): p. 271-296. [http://dx.doi.org/10.1016/j.apenergy.2006.09.012]

32. Mehr, H.D., H. Polat, and A. Cetin. *Resident activity recognition in smart homes by using artificial neural networks*. in *2016 4th International Istanbul Smart Grid Congress and Fair (ICSG)*. 2016. [http://dx.doi.org/10.1109/SGCF.2016.7492428]
33. Belshaw, M., et al. *Towards a single sensor passive solution for automated fall detection*. in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2011. [http://dx.doi.org/10.1109/IEMBS.2011.6090506]
34. McCulloch, W.S. and W. Pitts, *A logical calculus of the ideas immanent in nervous activity.* The bulletin of mathematical biophysics, 1943. **5**(4): p. 115-133.
35. Hebb, D., *The Organization of Behavior. emphNew York*. 1949, Wiley.
36. Pegalajar, M., *Inferencia gramatical mediante redes neuronales artificiales*. 1998, Universidad de Granada.
37. Gurney, K., *An introduction to neural networks*. 2014: CRC press. [http://dx.doi.org/10.1201/9781315273570]
38. Koza, J.R., *Genetic Programming: On the Programming of Computers by Means of Natural Selection.* 1992. 1-187. [http://www.cs.bham.ac.uk/~wbl/biblio/cache/bin/cache.php?koza:book,http___mitpress.mit.edu_books_genetic-programming,http://mitpress.mit.edu/books/genetic-programming]
39. Cuéllar, M.P., *Sistemas evolutivos para entrenamiento y optimización de modelos neuronales recurrentes dinámicos. Aplicación en modelado y predicción de series de datos*. 2006, Universidad de Granada.
40. Eshelman, L.J., *The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination*, in *Foundations of Genetic Algorithms*, R. Gregory J.E, Editor. 1991, Elsevier. p. 265-283. [http://www.sciencedirect.com/science/article/pii/B9780080506845500203]
41. Hüllermeier, E., *Fuzzy methods in machine learning and data mining: Status and prospects.* Fuzzy Sets and Systems, 2005. **156**(3): p. 387-406. [http://dx.doi.org/10.1016/j.fss.2005.05.036]
42. Delgado, M., et al., *A formal model for mining fuzzy rules using the RL representation theory.* Information Sciences, 2011. **181**(23): p. 5194-5213. [http://dx.doi.org/10.1016/j.ins.2011.07.029]
43. Castillo-Ortega, R., et al., *Linguistic summarization of time series data using genetic algorithms.* Proc. of the 7th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-2011), 2011(July): p. 416-423. [http://air.unimi.it/handle/2434/161284]