

Máquinas de aprendizaje extremo aleatorias para la predicción del consumo eléctrico de edificios

Gonzalo Vergara¹, Juan I. Alonso-Barba¹, Emilio Soria-Olivas², and José A. Gámez¹

¹ Grupo de investigación SIMD, I³A, Universidad de Castilla-La Mancha, Albacete
² Grupo de investigación IDAL, E.T.S.E, Universidad de Valencia, Burjassot

Resumen El estudio de la eficiencia energética en edificios es un campo de investigación activo. La modelización y predicción de magnitudes relacionadas con la energía permite analizar el consumo eléctrico dando lugar a ventajas tanto ambientales como económicas. En este estudio, se utilizan y comparan diferentes técnicas basadas en redes neuronales artificiales con el fin de predecir la potencia activa consumida por diferentes edificios. Los datos disponibles son de tipo temporal, ambiental y eléctrico y corresponden a 30 edificios del campus de la Universidad de León (España). Segmentamos los edificios en función de su consumo eléctrico utilizando el análisis de componentes principales y la técnica de *clustering* k-medoides. Posteriormente aplicamos técnicas propias del aprendizaje máquina como el perceptrón multicapa (*MLP*) y las máquinas de aprendizaje extremo (*ELM*). Como novedad introducimos una técnica basada en un *ensemble* de máquinas *ELM*. Nuestro análisis muestra que el método propuesto mejora los resultados de *MLP* en días laborables donde el consumo eléctrico es mayor.

Keywords: Ensembles, Máquinas de Aprendizaje Extremo, Perceptrón Multicapa, Predicción, Consumo eléctrico.

1. Introducción

Actualmente se estima que el consumo de edificios supone el 40 % del total de energía consumida en países desarrollados [20]. La agencia internacional de la energía (*IEA*) ha expresado que el sector de los edificios es uno de los más interesantes para invertir en eficiencia energética. Desde el punto de vista económico es posible reducir el consumo energético en 1509 Mtep³ en el año 2050 [9]. Este hecho implica también importantes ventajas ambientales. Si la demanda energética de los edificios se reduce como dictaminan las normativas de la Unión Europea 32/2006 [18] y 27/2012 [19], las emisiones de dióxido de carbono (CO₂) pueden disminuir de modo importante también. En este caso se estima que 12.6 Gt de CO₂ dejarían de emitirse a la atmósfera en el año 2050 [9].

³ 1 tep = 1 tonelada equivalente de petróleo = 11.63 MWh

En la literatura se encuentra una gran variedad de técnicas de minería de datos aplicables a la predicción del consumo energético. Entre ellas destaca la regresión lineal (*LR*) que se aplica a la predicción del consumo eléctrico mensual en grandes edificios públicos [12], los modelos autorregresivos integrados de media móvil (*ARIMA*) permiten detección de fallos en equipos electrónicos [2]. Las redes neuronales artificiales (*ANN*) se han utilizado extensamente para el estudio del consumo energético en edificios minimizando el consumo de energía debido al aire acondicionado en un edificio de oficinas [11] o la energía de edificios con luz natural [24]. También se han aplicado al estudio de la influencia del calentamiento de un edificio en el consumo eléctrico [6]. Las máquinas de vectores soporte (*SVM*) se utilizan en [26] para predecir el consumo eléctrico en grandes edificios a partir de datos simulados.

En este estudio utilizaremos datos correspondientes a 30 edificios de la Universidad de León [1,13]. Estos edificios tienen diferentes usos y se pueden clasificar en tres tipos: *académicos* como la facultad de Biología o Filosofía (edificios 13 y 6), de *soporte* como el Colegio Mayor (edificio 25) y de *investigación* como el Centro Tecnológico.

Las máquinas de aprendizaje extremo (*ELM*) se usan tanto en problemas de clasificación como de regresión [8,22]. La principal ventaja de las *ELM* es la ausencia del ajuste en la capa oculta del modelo neuronal por lo que su tiempo de aprendizaje es mucho menor que un *MLP* clásico. *ELM* ha probado su capacidad de predicción en diferentes problemas [8]. En este trabajo planteamos la construcción de un modelo de agregación de *ELMs* utilizando una modificación de la técnica de *bagging*. Con esta técnica que denominamos máquinas de aprendizaje extremo aleatorias (*RELM*) conseguimos mejorar los resultados de *MLP* en días laborables. En días festivos el consumo es casi constante. A continuación exponemos el método para ajustar los parámetros de los modelos y presentamos sus resultados para distintos edificios.

2. Datos Analizados

Según el código técnico de la edificación [14] la ciudad de León se sitúa en la zona climática E1 y se encuentra a 914.7 m sobre el nivel del mar. Esta zona tiene condiciones climáticas severas (tanto en invierno como en verano) y por ello el consumo eléctrico del equipamiento debido a tareas de calentamiento, ventilación y aire acondicionado (*HVAC*) es alto. La mayor parte de los datos corresponden al año 2010 en el que la temperatura mínima en la ciudad de León fue de -10°C y la temperatura máxima de 34°C .

Los conjuntos de datos disponibles contienen variables de tipo ambiental y eléctrico. Las variables ambientales se adquirieron desde la misma estación meteorológica para todos los edificios mientras que las variables de tipo eléctrico se tomaron independientemente para cada edificio [1]. Las variables consideradas en este estudio han sido: hora, día, mes, año, temperatura ambiente ($^{\circ}\text{C}$), humedad relativa (%), radiación solar (W/m^2), día laborable, potencia activa horaria (kW) y desviación estándar de potencia horaria (kW).

2.1. Preprocesado de datos

Los datos se registraron cada dos minutos durante 13 meses (desde el uno de marzo de 2010 hasta el 31 de marzo de 2011) lo que suma 285120 patrones para cada edificio. Con el fin de obtener una medida horaria para cada variable hemos calculado la media aritmética de los valores obtenidos durante cada hora para cada una de las variables. Sin embargo las instancias con valores perdidos se han eliminado y, solamente, se han tenido en cuenta las horas que contienen 30 muestras.

Para agrupar los edificios de acuerdo a su comportamiento eléctrico hemos llevado a cabo un análisis de componentes principales (*PCA*) [10] utilizando el siguiente conjunto de variables derivadas de la potencia activa de cada edificio: mínimo, máximo, media, rango, desviación típica y coeficiente de variación. Sobre el espacio transformado por el *PCA* hemos elaborado un *clustering* del tipo *k-medoids*. Las dos primeras componentes principales explican el 95.73 % de la variación de las variables. Hemos encontrado que la mejor elección del número de *clusters* es tres en términos del índice *silhouette* [16] (Fig. 1). Estos tres grupos tienen 4, 10 y 16 edificios respectivamente. Desde un punto de vista descriptivo observamos que el primer grupo de edificios tiene un consumo alto (100 kW) y baja variación (30 %), el segundo grupo se caracteriza por un consumo medio (30 kW) y alta variación (90 %) y el tercer grupo tiene bajos consumos (10 kW) y altas variaciones de potencia (100 %).

Para comparar las diferentes técnicas hemos seleccionado como representante de cada grupo un edificio. En la Fig. 1 vemos los diferentes edificios agrupados y el edificio representante de cada grupo. Los tres edificios seleccionados y sus códigos han sido: edificio de Biología (13), edificio de Filosofía (6) y Colegio Mayor (25).

La variable hora presenta una discontinuidad entre las 23 y las 0 h. Para salvar este inconveniente es habitual la siguiente transformación trigonométrica desdoblado la hora en dos nuevas variables [5]:

$$h_x = \sin\left(\frac{2\pi h}{24}\right) \quad \text{y} \quad h_y = \cos\left(\frac{2\pi h}{24}\right)$$

donde h es la hora. Además los conjuntos de datos se han reordenado de modo aleatorio con el fin de romper la dependencia causal entre instancias consecutivas. Finalmente todas las variables se han estandarizado.

3. Técnicas utilizadas

En ciertos casos las técnicas autorregresivas (*AR*) o de regresión lineal (*LR*) pueden ser satisfactorias ya que son fáciles de implementar y de bajo coste computacional [17]. Cuando las técnicas lineales no son suficientes se recomienda el uso de técnicas no lineales como el perceptrón multicapa (*MLP*) [3,7]. Es bien conocido que el *MLP* es robusto y tiene buenos resultados predictivos en gran variedad de aplicaciones [15,21], sin embargo, su principal inconveniente

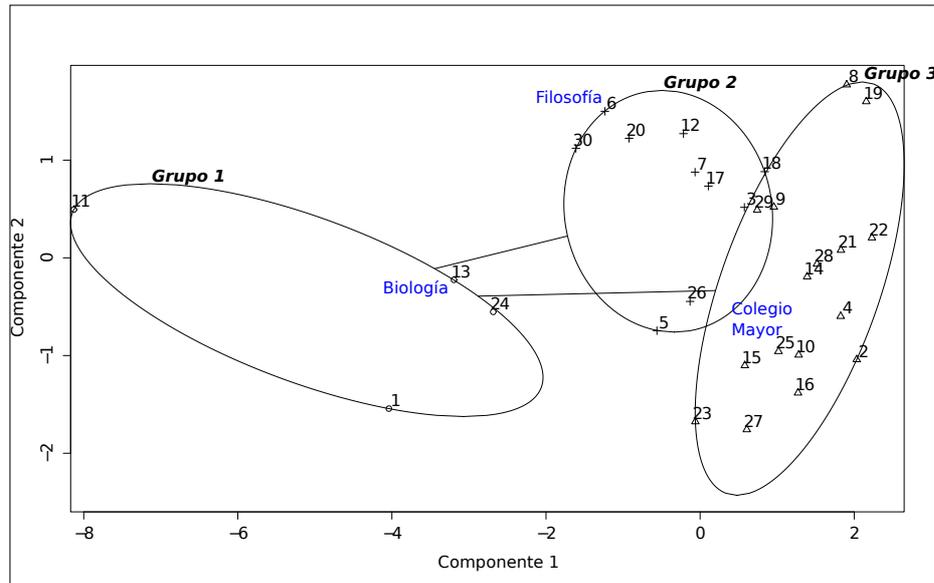


Figura 1. Agrupamiento de los 30 edificios en términos de las dos primeras componentes principales.

es el largo tiempo de entrenamiento. Por otra parte *ELM* ha demostrado buenos resultados en problemas de regresión [8]. En nuestro problema de consumo eléctrico de edificios la agregación de *ELM* (*RELM*) ha obtenido mejores resultados que *MLP* en el conjunto de días laborables tanto en términos de error de predicción como de tiempo de computación. Si bien es cierto que en términos de error global *MLP* sigue obteniendo menores errores.

A continuación describimos brevemente las técnicas utilizadas.

3.1. Perceptrón multicapa

La arquitectura de red utilizada en este estudio es el perceptrón multicapa. Este modelo consiste en un conjunto de unidades computacionales conocidas como neuronas artificiales [7]. Las salidas de las neuronas de una capa alimentan a las neuronas de la siguiente capa. Cada valor de entrada x_i a la neurona j se multiplica por un coeficiente w_{ij} llamado peso sináptico que representa la intensidad de la conexión entre neuronas. A este valor se le aplica una función no lineal llamada función de activación siendo habitual utilizar la función tangente hiperbólica ϕ . Con estos elementos la salida o_j de la j -ésima neurona está dada por:

$$o_j = \phi \left(\sum_{i=0}^n w_{ij} \cdot x_i \right)$$

donde n es el número de variables del conjunto de datos. Las neuronas de una red específica se agrupan formando capas hasta dar lugar a una red completamente conectada. El proceso de entrenamiento consiste en ajustar los pesos sinápticos del *MLP*; en nuestro caso hemos utilizado el algoritmo de *Levenberg-Marquardt* que consiste en una combinación del método de gradiente descendente y del método de *Gauss-Newton*. Este algoritmo tiene una convergencia estable y rápida [25].

3.2. Máquinas de aprendizaje extremo

El algoritmo *ELM* se debe a Huang [8] y toma como base la estructura del perceptrón de una capa oculta con H neuronas cuyos pesos son inicializados aleatoriamente. Los pesos de salida se pueden obtener analíticamente mediante la matriz pseudoinversa [8]. Así pues, dado un conjunto de N vectores de entrada existen los parámetros β_i , \mathbf{w}_i y b_i con:

$$\sum_{i=1}^H \beta_i \phi(\mathbf{w}_i \mathbf{x}_j + b_i) = \mathbf{t}_j,$$

donde $j = 1, \dots, N$ y \mathbf{t}_i es el i -ésimo vector *target*. Esta ecuación se puede expresar de forma matricial como:

$$\mathbf{H}\beta = \mathbf{T},$$

donde \mathbf{H} es la matriz $N \times H$ de salidas de la capa oculta, β es la matriz $H \times n$ de pesos de salida y \mathbf{T} es una matriz $N \times n$ de N *targets*. El entrenamiento consiste en un problema de mínimos cuadrados con pesos óptimos $\hat{\beta} = \mathbf{H}^\dagger \mathbf{T}$, donde $\mathbf{H}^\dagger = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ es la pseudoinversa de *Moore-Penrose* [8].

3.3. Máquinas de aprendizaje extremo aleatorias

En este trabajo, inspirándonos en el método *random forest* [4] hemos aplicado un *bagging* modificado. Mantendremos un número de variables fijo y con el resto de ellas se hace una selección aleatoria. El proceso de aprendizaje tiene dos etapas:

1. Selección de diferentes variables para entrenar los modelos *ELM*. Fijamos algunas variables y aplicamos *bagging* al conjunto de las variables fijas más las variables aleatorias. En esta etapa seleccionaremos dos parámetros: factor F que se define como el número de neuronas dividido por el número de variables del modelo y el número de *bootstraps* (B).
2. Ajuste del tercer parámetro utilizando un conjunto de aprendizaje para días laborables. Este parámetro que define la agregación de los modelos *ELM* es el percentil del modelo y lo ajustamos de manera que el *MAE* del conjunto de validación sea el menor.

En el diagrama 2 podemos observar el proceso de ajuste de parámetros de los modelos *RELM*.

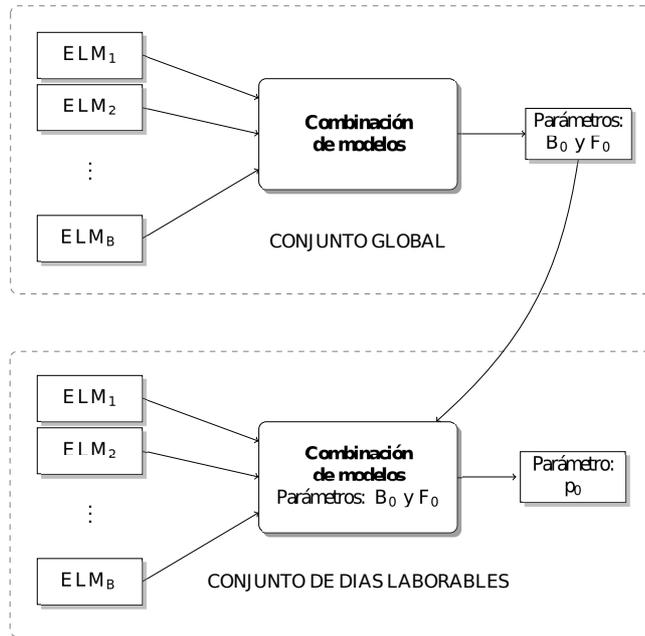


Figura 2. Ajuste de los parámetros en la técnica RELM.

4. Evaluación experimental

En esta sección describimos el proceso de aprendizaje para los modelos *MLP*, *ELM* y *RELM*. El método de validación empleado es un *hold-out* de 10 iteraciones (reordenaciones) con un 55% de las instancias para el conjunto de entrenamiento, el 15% para el conjunto de validación y el 30% para el conjunto de *test*. Este proceso se ha realizado para los tres edificios seleccionados de Biología, Filosofía y Colegio Mayor. Las variables predictoras contienen únicamente la información de la hora anterior ya que tomar información previa a una hora no mejora los resultados al aplicar *MLP* y *ELM* a nuestro problema [22]. Las predicciones se han realizado a una hora para todos los modelos con el fin de comparar sus resultados.

Las medidas de error que hemos utilizado han sido el error absoluto medio (*MAE*) utilizado comúnmente en datos de consumo eléctrico por contener *outliers* y la raíz del error cuadrático medio (*RMSE*) [23]:

$$MAE = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{N}, \quad RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}}$$

donde y_i son los datos reales, \hat{y}_i son las predicciones y N es el número de instancias. Los modelos se han construido para cada una de las 10 reordenaciones y se ha calculado el *MAE* del conjunto de validación de cada reordenación. Además se ha promediado el *MAE* del conjunto de *test* de las 10 reordenaciones.

Los parámetros utilizados para las diferentes técnicas han sido los siguientes. Los modelos *ELM* se han entrenado utilizando desde 100 hasta 2000 neuronas con un incremento de 100 neuronas. Se toma la primera aproximación del número de neuronas H_0 de modo que minimice el error promedio del conjunto de validación y posteriormente se realiza un ajuste fino desde $H_0 - 100$ hasta $H_0 + 100$ con un incremento de 20 neuronas. Los modelos *MLP* para una y dos capas ocultas se han entrenado utilizando la función sigmoide en cada capa oculta y una función lineal en la capa de salida. Se hace variar el número de neuronas por capa desde 2 hasta 20 neuronas con un incremento de 2 neuronas. Se selecciona la configuración de neuronas con menor *MAE* del conjunto de validación.

Los modelos *RELM* se han entrenado tomando en la fase 1 las variables fijas (proyecciones de la hora y potencia activa en la hora actual) más una selección aleatoria del resto de variables. Barremos el factor F entre 50 y 100 con un incremento de 10 y el número de *bootstraps* B entre 10 y 100 con un incremento de 10. En la fase 2 y tras seleccionar la mejor configuración F_0 y B_0 hemos probado distintas agregaciones de los modelos *ELM* (media, mediana, media de los cuartiles primero y tercero, etc.) ajustamos el parámetro p hasta encontrar el modelo cuyas salidas proporcionen el menor *MAE* de validación del conjunto de días laborables que finalmente ha sido el percentil $p_0 = 55$.

Todos los experimentos incluyendo las medidas de los tiempos de computación se han llevado a cabo con un procesador Intel® Core™ i7 3770 a 3.4 GHz utilizando OS X. Los cálculos y modelos se han ejecutado utilizando Matlab® v2010 y v2012.

4.1. Comparación entre técnicas

En esta sección mostramos los resultados dados por los modelos *MLP*, *ELM* y *RELM* para los tres edificios seleccionados. Además se comparan las prestaciones de cada una de las técnicas atendiendo a los errores globales y locales así como al tiempo de computación de cada modelo.

La Tabla 1 muestra los errores de cada modelo para distintos conjuntos.

Tabla 1. Errores de los modelos *ELM* para potencia en distintos edificios. Además del número óptimo de nodos H de cada modelo se presenta el *MAE* correspondiente al conjunto de *test*. Los valores de potencia se acompañan de la desviación estándar.

Edificio	Potencia (kW)	H	MAE_t
Biología	73.06 ± 30.75	800	4.34
Filosofía	27.84 ± 26.74	1300	2.90
Colegio Mayor	14.22 ± 7.44	600	1.96

En la Tabla 2 aparecen los resultados para el edificio de Biología. En primer lugar *ELM* da valores de error un 50% mayores que el resto de métodos y necesita un gran número de neuronas para entrenar. El perceptrón tiene los mejores resultados en los valores globales de *MAE* o *RMSE*. Sin embargo los

Tabla 2. Resultados de los modelos para el edificio de Biología ($P = 73.06 \pm 30.75$ kW). Se presentan los parámetros del modelo k , el tiempo medio de aprendizaje t y los errores MAE y $RMSE$ de diferentes conjuntos. En negrita se resaltan los mejores valores.

Modelo	k	t (s)	MAE_e	MAE_v	MAE_t	$RMSE_t$	$MAE_{L,v}$	$MAE_{L,t}$
<i>ELM</i>	800	1.4 ± 0.5	3.87	4.82	4.34	6.01	18.34	16.78
<i>1LP</i>	20	290 ± 80	2.99	2.94	2.88	4.04	6.06	7.19
<i>2LP</i>	10/12	280 ± 30	2.85	2.82	2.81	3.92	6.95	4.97
<i>RELM</i>	130/20	181 ± 12	2.71	3.21	3.33	4.93	5.08	4.79

Tabla 3. Resultados de los modelos para el edificio de Filosofía ($P = 27.84 \pm 26.74$ kW). Se presentan los parámetros del modelo k , el tiempo medio de aprendizaje t y los errores MAE y $RMSE$ de diferentes conjuntos. En negrita se resaltan los mejores valores.

Modelo	k	t (s)	MAE_e	MAE_v	MAE_t	$RMSE_t$	$MAE_{L,v}$	$MAE_{L,t}$
<i>ELM</i>	1300	3.4 ± 1.1	2.23	3.50	2.90	4.23	27.81	24.10
<i>1LP</i>	18	270 ± 80	1.98	2.02	1.96	3.03	4.40	6.27
<i>2LP</i>	16/20	197 ± 22	1.91	1.79	1.80	2.81	3.85	3.14
<i>RELM</i>	140/20	198 ± 9	1.64	2.08	2.05	3.38	4.03	2.91

Tabla 4. Resultados de los modelos para el edificio de Colegio Mayor ($P = 14.22 \pm 7.44$ kW). Se presentan los parámetros del modelo k , el tiempo medio de aprendizaje t y los errores MAE y $RMSE$ de diferentes conjuntos. En negrita se resaltan los mejores valores.

Modelo	k	t (s)	MAE_e	MAE_v	MAE_t	$RMSE_t$	$MAE_{L,v}$	$MAE_{L,t}$
<i>ELM</i>	600	0.9 ± 0.3	1.75	2.04	1.96	2.71	7.10	4.71
<i>1LP</i>	18	105 ± 10	1.56	1.49	1.55	2.28	2.93	2.54
<i>2LP</i>	6/20	131 ± 13	1.50	1.46	1.48	2.26	2.95	2.63
<i>RELM</i>	80/10	55 ± 4	1.49	1.60	1.70	2.58	2.70	2.05

modelos *RELM* obtienen los mejores resultados de *MAE* en el conjunto de días laborables tanto de validación como de *test*. Este hecho ocurre también en el edificio de Colegio Mayor (Tabla 4) incluso con una mayor diferencia en el caso de *RELM* con una disminución del error de test del 20 % respecto al perceptrón con una capa oculta con el segundo mejor resultado.

En el edificio de Filosofía (Tabla 3) los resultados son algo diferentes. Si bien *RELM* obtiene mayor precisión en el conjunto global de entrenamiento y de test de días laborables, es el perceptrón con dos capas ocultas quien obtiene el mejor resultado en el conjunto de validación de días laborables. En cuanto a los tiempos de aprendizaje los modelos *RELM* resultan ser más rápidos que los del perceptrón salvo en el edificio de Filosofía con tiempo similar al del perceptrón con dos capas ocultas. También vemos que el número de factores en *RELM* es unas 10 veces mayor que el número de neuronas por capa en *2LP*.

5. Conclusiones

En este trabajo hemos utilizado técnicas de minería de datos para construir modelos de predicción a una hora de la energía consumida en diferentes edificios pertenecientes a la Universidad de León.

Hemos descrito las técnicas utilizadas que tienen en común la estructura de red neuronal, con una o dos capas ocultas, e introducimos el método *RELM* basado en una modificación del *bagging*. La técnica *RELM* consiste en una agregación de *ELMs* en donde un subconjunto de variables se mantiene en todos los modelos y el resto de variables se va tomando de modo aleatorio. En segundo lugar hemos evaluado el comportamiento de los modelos *ELM*, modelos *MLP* (con una y dos capas ocultas) y modelos *RELM* independientemente.

Finalmente hemos comparado las tres técnicas incidiendo en aquellas con mejores resultados. *ELM* es la técnica más sencilla de implementar pero da los peores resultados. *MLP* con dos capas ocultas o en algún caso con una capa oculta da los mejores resultados a nivel global. Sin embargo, si atendemos al conjunto de días laborables con los mayores consumos en edificios académicos (Biología y Filosofía) concluimos que la técnica *RELM* es la más precisa en la predicción del consumo eléctrico. Además mostramos que, en nuestro problema, *RELM* es una técnica más rápida que *MLP* debido por una parte a estar compuesta por *ELMs* y, por otra parte, a que obtiene buenos resultados con un número pequeño de *bootstraps*.

Agradecimientos. Este trabajo ha sido parcialmente sostenido con fondos de la Junta de Comunidades de Castilla-La Mancha a través del proyecto PEII-2014-049-P. Los autores quieren mostrar su agradecimiento al grupo de investigación *SUPRESS* de la Universidad de León por su colaboración.

Referencias

1. Alonso, S.: Supervisión de la energía eléctrica en edificios públicos de uso docente basada en técnicas de minería de datos visual. Ph.D. thesis, Departamento de

- Ingeniería Eléctrica, Electrónica, de Computadores y Sistemas. Universidad de Oviedo. (2012)
2. Bian, X., Xu, Q., Li, B., Xu, L.: Equipment fault forecasting based on a two-level hierarchical model. In: Automation and Logistics, 2007 IEEE International Conference on. pp. 2095–2099 (2007)
 3. Bishop, C.: Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc. (2006)
 4. Breiman, L.: Random forests. *Machine Learning* 45, 5–32 (2001)
 5. Carpinteiro, O., Alves da Silva, A., Feichas, C.: A hierarchical neural model in short-term load forecasting. In: IJCNN (6). pp. 241–248 (2000), <http://doi.ieeecomputersociety.org/10.1109/IJCNN.2000.859403>
 6. Ekici, B., Aksoy, U.: Prediction of building energy consumption by using artificial neural networks. *Advances in Engineering Software* 40(5), 356–362 (2009)
 7. Haykin, S.: *Neural Networks and Learning Machines*. Prentice Hall, 3 edn. (2009)
 8. Huang, G., Zhu, Q., Siew, C.: Extreme learning machine: Theory and applications. *Neurocomputing* (70), 489 – 501 (2006)
 9. I.E.A. International Energy Agency: Energy Performance Certification of Buildings (2013)
 10. Jolliffe, I.: *Principal Component Analysis*. Springer Verlag (1986)
 11. Kusiak, A., Li, M., Tang, F.: Modeling and optimization of HVAC energy consumption. *Applied Energy* 87(10), 3092–3102 (2010)
 12. Ma, Y., Yu, J., Yang, C., Wang, L.: Study on power energy consumption model for large-scale public building. In: Proceedings of the 2nd international workshop on intelligent systems and applications. pp. 1–4 (2010)
 13. Mateo, F., Carrasco, J.J., Millán-Giraldo, M., Sellami, A., Escandell-Montero, P., Martínez-Martínez, J.M., Soria-Olivas, E.: Machine learning techniques for short-term electric power demand prediction. In: 21st European Symposium on Artificial Neural Networks, ESANN 2013, Bruges, Belgium, April 24-26, 2013 (2013)
 14. Ministerio de Fomento, Gobierno de España: Código Técnico de la Edificación (2010), <http://www.codigotecnico.org>
 15. Paliwal, M., Kumar, U.: Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications* 36(1), 2–17 (2009)
 16. Rousseeuw, P.: Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* 20(0), 53 – 65 (1987)
 17. Soliman, S., Al-Kandari, A.: Electric load modeling for long-term forecasting. In: *Electrical Load Forecasting*, pp. 353–406. Butterworth-Heinemann, Boston (2010)
 18. Unión Europea: Directiva 2006/32/CE del Parlamento europeo y del Consejo de 5 de abril sobre la eficiencia del uso final de la energía y los servicios energéticos (2006), <http://www.boe.es/doue/2006/114/L00064-00085.pdf>
 19. Unión Europea: Directiva 2012/27/UE del Parlamento europeo y del Consejo de 25 de octubre relativa a la eficiencia energética (2012), <http://www.boe.es/doue/2012/315/L00001-00056.pdf>
 20. U.S. Department of Energy: Buildings Energy Data Book (2010), <http://buildingsdatabook.eren.doe.gov/DataBooks.aspx>
 21. Vellido, A., Lisboa, P., Vaughan, J.: Neural networks in business: a survey of applications (1992–1998). *Expert Systems with Applications* 17(1), 51–70 (1999)
 22. Vergara, G., Cózar, J., Romero-González, C., Gámez, J.A., Soria-Olivas, E.: Comparing elm against mlp for electrical power prediction in buildings. In: Ferrández Vicente, J.M., Álvarez Sánchez, J.R., de la Paz López, F., Toledo-Moreo,

- F.J., Adeli, H. (eds.) *Bioinspired Computation in Artificial Systems*, Lecture Notes in Computer Science, vol. 9108, pp. 409–418. Springer International Publishing (2015)
23. Willmott, C., Matsuura, K.: Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate Research* 30(1), 79 (2005)
 24. Wong, S., Wan, K., Lam, T.: Artificial neural networks for energy analysis of office buildings with daylighting. *Applied Energy* 87(2), 551–557 (2010)
 25. Yu, H., Wilamowski, B.: *The Industrial Electronics Handbook*, vol. 5. CRC (2011)
 26. Zhao, H., Magoulès, F.: Parallel support vector machines applied to the prediction of multiple buildings energy consumption. *Journal of Algorithms & Computational Technology* 4(2), 231–249 (2010)