

**MODELOS PARA LA PREDICCIÓN DEL AUTOCONSUMO EN SISTEMAS FOTOVOLTAICOS
CONECTADOS A RED****Llanos Mora López*, Mariano Sidrach-de-Cardona****

* Dpto. de Lenguajes y Ciencias de la Computación. Universidad de Málaga. 29071 Málaga (España)

llanos@uma.es

**Dpto. de Física Aplicada II. . Universidad de Málaga. 29071 Málaga (España)

msidrach@uma.es<https://doi.org/10.34637/cies2020.1.2065>

RESUMEN

En este trabajo se presentan los resultados obtenidos para la modelización y optimización de instalaciones fotovoltaicas de autoconsumo. Se han obtenido las curvas de autoconsumo y autosuficiencia para diferentes perfiles de consumo horario en función de la potencia pico instalada y el tamaño de la batería. El estudio se ha realizado para tres ciudades españolas con diferentes condiciones climáticas. Para la generalización de los resultados se proponen diferentes modelos de aprendizaje automático que permiten estimar estos parámetros. Las variables de entrada de estos modelos están relacionadas con la configuración de la instalación, su ubicación y el tipo de perfil de consumo. El modelo que arroja mejores predicciones en el parámetro de autosuficiencia es Random Forest, que en la validación cruzada tiene un error relativo del 5%. Para la predicción del autoconsumo, el modelo que mejor se comporta es el Perceptrón Multicapa, con un error absoluto promedio de 0.55 y un error relativo del 3%.

PALABRAS CLAVE: autoconsumo fotovoltaico, autosuficiencia fotovoltaica, evaluación, modelización.

ABSTRACT

The results obtained for the modeling and optimization of photovoltaic self-consumption facilities are presented. The study has been carried out for three Spanish cities with different climatic conditions. The self-consumption and self-sufficiency curves for different hourly consumption profiles have been obtained based on the installed peak power and the size of the battery. In order to generalize the obtained results, different models of machine learning are proposed to estimate these parameters. The input variables of these models are related to the configuration of the installation, its location and the type of consumption profile. The model with best predictions of self-sufficiency is Random Forest, which in cross-validation has a relative error of 5%. For the prediction of self-consumption, the model that performs best is the Multilayer Perceptron, with an average absolute error of 0.55 and a relative error of 3%.

KEYWORDS: photovoltaic self-consumption, photovoltaic self-sufficiency, evaluation, modelling

INTRODUCCIÓN

La energía fotovoltaica está llamada a jugar un papel cada vez más importante dentro del actual mix energético. Dentro de los sistemas fotovoltaicos, las instalaciones de autoconsumo permiten empoderar a los ciudadanos y responsabilizarlos de la producción y uso de la energía. En los últimos años, y gracias al crecimiento que se está produciendo en este tipo de instalaciones, ha surgido el concepto de vivienda autosuficiente y vivienda de energía cero. Una casa o un edificio de energía cero (conocido como ZEB, Zero Energy Building) es un concepto que se utiliza en edificios con un balance energético entre generación y consumo de energía cercano a cero o incluso positivo en un año típico (NREL, 2006), (NREL, 2010). El desarrollo masivo de este tipo de viviendas podría mitigar problemas económicos y ambientales, como las emisiones de CO₂ y la dependencia de fuentes de energía fósil.

La autogeneración de electricidad mediante tecnología fotovoltaica es una oportunidad y una herramienta esencial para lograr los objetivos relacionados con el cambio climático que España ha comprometido, reducir el impacto ambiental de la generación de electricidad y democratizar el uso y la gestión de la energía.

El correcto dimensionado, diseño y evaluación inicial de estos sistemas debe calcular, además de la energía generada en función de las condiciones climáticas, los porcentajes de autoconsumo y autosuficiencia; es decir, qué fracción de la producción fotovoltaica va directamente al consumo (autoconsumo) y qué fracción del consumo es cubierto directamente por el sistema fotovoltaico. Para este fin es necesario conocer el perfil de consumo y realizar un balance energético del sistema en el punto de conexión. Debido a que el intercambio de energía con la red es instantáneo, al menos, este balance energético debe hacerse con valores horarios, si queremos tener valores de estos parámetros que se aproximen a la realidad. La situación es más compleja en el caso de tener sistemas fotovoltaicos con acumulación.

Además, la dificultad de acceso, que ponen las distribuidoras, a los datos de los contadores digitales domésticos hace que, en la práctica, estos valores se estimen a partir de un perfil horario medio anual tipo.

El análisis y modelización de sistemas fotovoltaicos de autoconsumo nos permite determinar los diferentes escenarios que definen las condiciones de diseño y funcionamiento óptimos de este tipo de sistemas, la energía que se puede utilizar directamente y la energía que se va a intercambiar con el red de energía eléctrica.

En este trabajo presentamos los resultados obtenidos para la optimización de las instalaciones de autoconsumo. El estudio se ha realizado para diferentes ciudades españolas con diferentes condiciones climáticas. Las curvas de autoconsumo y autosuficiencia se han obtenido en base a la potencia pico instalada y el tamaño de la batería, para diferentes perfiles de consumo horario. Se proponen diferentes modelos de aprendizaje automático para predecir estos parámetros.

MATERIALES Y MÉTODOS

La energía generada por un sistema fotovoltaico depende, por un lado, de parámetros climatológicos, como la radiación solar que recibe y la temperatura ambiente y, por otro, de la tecnología de los módulos, como tipo de módulos, rendimiento, etc, y de la orientación e inclinación de éstos. Además, se deben tener en cuenta las diferentes pérdidas que se producen en el sistema. Para estimar la potencia generada por un generador fotovoltaico (P_m) se utiliza el método propuesto en (Osterwald, 1986) que utiliza como datos de entrada la irradiancia incidente, la temperatura de los módulos y la potencia pico del generador en condiciones estándar, y cuya expresión es:

$$P_m(W) = P_{m,ref} \cdot \frac{G}{G_{ref}} \cdot [1 + \gamma \cdot (T_m - T_{m,ref})] \quad (1)$$

donde, P_m es la potencia máxima del generador fotovoltaico en condiciones estándar (STC), G es la irradiancia global incidente en los módulos (W/m^2), G_{ref} es la irradiancia global en STC ($1000 W/m^2$), T_m es la temperatura del módulo, $T_{m,ref}$ es la temperatura del módulo en STC (25°) y γ es el coeficiente de temperatura de pérdidas de potencia ($\%/^\circ C$). Para calcular la temperatura de trabajo de los módulos fotovoltaicos se ha utilizado el modelo de King. Este modelo propone una relación entre la temperatura del módulo, la temperatura ambiente (T_{amb} , $^\circ C$), la irradiancia incidente (G , W/m) y la velocidad del viento (v , m/s) según la expresión:

$$T_m = T_{amb} + G \cdot e^{(m+n \cdot v)} \quad (2)$$

donde m es un coeficiente empírico adimensional que describe el impacto de la irradiancia en la temperatura del módulo y n es un coeficiente empírico que describe el efecto del viento en la temperatura del módulo. En este caso los coeficientes utilizados corresponden a módulos de Si-monocristalino y módulos instalados en una estructura al aire libre. Los parámetros de este modelo se describen en (Mora-Segado, 2015).

Las pérdidas del sistema fotovoltaico consideradas son:

- Pérdidas por reflectancia angular y variación del espectro de radiación incidente (L_r). En este trabajo, se supondrá que son del 3%.
- Pérdidas óhmicas por cableado (L_o). En este estudio, se supondrá que son 2%, lo que significa $L_o=0,98$.
- Pérdidas por dispersión de parámetros en el generador (L_p). En este estudio, se supondrá que son del 2%. $L_p=0,98$.
- Pérdidas por errores en el seguimiento del punto de máxima potencia del inversor (L_i). En este estudio, se supondrá que son del 1%. $L_i=0,99$.

La potencia a la entrada del inversor se estima mediante la expresión:

$$P_{in} = L_r L_o L_p L_i P_m \quad (3)$$

Finalmente, para determinar la potencia generada en corriente alterna a la salida del inversor (P_{out}), se debe tener en cuenta el rendimiento del inversor. Este rendimiento se calculará según la ecuación propuesta por Jantsch (M. Jantsch el., 1992) a partir de la potencia de entrada al inversor P_{in} normalizada a su potencia nominal, P_{nor} :

$$\eta_{inv} = \frac{P_{nor} - (b_0 + b_1 \cdot P_{nor} + b_2 \cdot P_{nor}^2)}{P_{in}} \quad (4)$$

$$P_{out} = \eta_{inv} P_{in} \quad (5)$$

donde b_i son los coeficientes de ajuste estimados empíricamente ($b_0=0.04$, $b_1=.002$, $b_2=0.03$).

Aunque este modelo está definido en términos de potencia, en este trabajo vamos a utilizarlo para realizar los balances energéticos horarios. Trabajos anteriores (Ayala et al. 2018) han demostrado la validez de esta hipótesis de trabajo cuantificando el error cometido respecto a un balance instantáneo.

Los parámetros a estimar y analizar para evaluar el funcionamiento de una instalación fotovoltaica de autoconsumo son similares a los propuestos en (Sartori et al., 2012):

- Energía generada por la instalación fotovoltaica
- Energía inyectada a la red eléctrica
- Energía importada de la red eléctrica
- Energía consumida

Además, en la evaluación de las instalaciones de autoconsumo fotovoltaico, dos parámetros son especialmente útiles según (Luthander et al., 2015):

Porcentaje de autoconsumo (SC_t): definido como la parte de energía producida por el sistema fotovoltaico que se consume directamente en la vivienda donde se encuentra la instalación, respecto a la producción total del sistema fotovoltaico:

$$SC_t (\%) = \frac{\sum E_{t,FV, self}}{\sum E_{t,FV}} \times 100 \quad (6)$$

Porcentaje de autosuficiencia (SS_t): definido como la parte de energía producida por el sistema fotovoltaico que se consume directamente en la vivienda donde se encuentra la instalación, respecto al consumo total de la vivienda.

$$SS_t (\%) = \frac{\sum E_{t,FV, sc}}{\sum C_t} \times 100 (\%) \quad (7)$$

El estudio se ha realizado para tres ciudades españolas diferentes con diferentes condiciones climáticas (Tabla 1).

Utilizando los parámetros meteorológicos de radiación incidente global horaria y temperatura horaria, se han obtenido, de acuerdo con la metodología anterior, los balances energéticos de una instalación de autoconsumo utilizando diferentes valores de potencia pico, capacidades de batería y diferentes perfiles de consumo.

Tabla 1. Parámetros usados para cada ubicación

Localidad	Gdm (kWh/m ²)	STD Gdm (kWh/m ²)	Temp (°C)	STD T (°C)	Consumo (kWh)	FV 1kWp (kWh)
Málaga	5.98	2.04	18.1	0.8	11	4.81
Madrid	5.09	2.46	14.2	1.6	14.3	4.07
Santander	4.61	2.39	14.9	0.9	11.8	3.71

Los perfiles de consumo horario utilizados se basan en los perfiles propuestos por el centro UKERK (UKERK). A partir de estos perfiles, se ha considerado el porcentaje de consumo producido cada hora respecto al consumo diario total (en valores medios anuales). Para la construcción de los perfiles horarios de los 365 días del año se han considerado los periodos propuestos por este centro, excepto el periodo que se denomina "verano alto". Los perfiles correspondientes a este periodo se han generado para aquellos días en los que la temperatura media diaria es superior a los 25 grados.

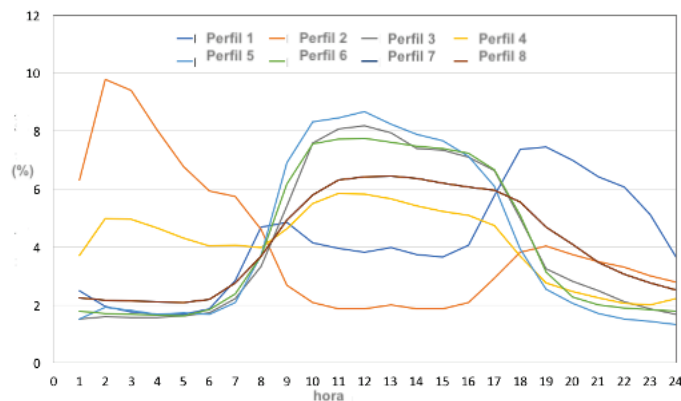


Figura 1. Perfiles de consumo utilizados

Una vez obtenidas las curvas de autoconsumo y autosuficiencia en función tanto de la potencia pico instalada, como de la capacidad de la batería para cada uno de los perfiles estudiados, se propone obtener una generalización del modelo de forma que se puedan obtener estos en función de unos pocos datos de entrada sin necesidad de hacer cada vez el balance energético horario. Para este fin se han utilizado modelos que se usan habitualmente en minería de datos.

La minería de datos se ha utilizado anteriormente para la predicción de la producción de sistemas fotovoltaicos; por ejemplo, en (Alfadda et al., 2017) y (Sharma et al., 2011) se propone el uso de modelos de soporte vectorial y en (Nageem y Jayabarathi, 2017) se incluyen factores de corrección que se incluyen en este tipo de modelos. Tienen en cuenta diferentes condiciones climáticas. También se han propuesto modelos híbridos que incluyen modelos eléctricos y estadísticos (Filipe et al., 2015) y modelos basados en redes neuronales y lógica difusa (Sivaneasan et al., 2017). En todos ellos lo que se hace es la predicción de la producción de un sistema, pero los valores de autoconsumo y autosuficiencia no se modelan en ningún caso.

Proponemos para este fin el uso de diferentes modelos, comparar los resultados obtenidos y decidir cuál es el mejor método de predicción. Los modelos utilizados han sido los siguientes:

- Regresión lineal: se utiliza con variables numéricas. En el caso de que existan variables independientes de tipo nominal, estas deben ser previamente transformadas en variables ficticias o *dummy*. La regresión lineal es un modelo que intenta ajustar variables independientes mediante una ecuación lineal.
- Perceptrón multicapa: es un modelo matemático útil para modelar relaciones no lineales entre datos de entrada y salida. Es un tipo de red neuronal que implica una generalización del perceptrón simple en el que se combinan varios perceptrones simples y permite tratar algunos problemas no lineales (Minsky y Papert, 1969). A partir de esta propuesta, (Rumelhart et al., 1986) presentó una modificación que permitía la retropropagación de los errores medidos en la salida de la red hacia las neuronas ocultas. Desde el punto de vista de su arquitectura se caracteriza porque tiene sus neuronas agrupadas en capas de diferentes niveles: una capa de entrada, una capa de salida y capas ocultas.
- M5' es un algoritmo derivado del método M5 definido por Quinlan (Quinlan, 1992). Fue propuesto por Wang y Witten (Wang y Witten, 1997). Se definió para predecir valores numéricos.

- REPTree : es un modelo de aprendizaje rápido de árboles de decisión. Para la clasificación de variables numéricas, el algoritmo primero ordena los valores de esas variables y comienza la ejecución. Luego, use las listas ordenadas para calcular la mejor manera de dividir en cada nodo del árbol. Esta forma de dividir minimiza la varianza. La medida que se utiliza es la entropía.
- Random Forest: forma parte de lo que se conoce como sistemas multclasificadores. Estos métodos suelen ser muy precisos (Sardá -Espinosa, 2017) y robustos en el caso de que haya ruido en los datos, además no producen sobreajuste. Sin embargo, son más difíciles de interpretar en comparación con los modelos basados en árboles de regresión simple. Es un algoritmo que induce una serie de árboles individuales (Breiman , 2001). Para ello, el número de variables que se utilizan para generar cada uno de esos árboles que se crearán a partir de cada nodo es limitado. Este número suele ser menor que el número de variables disponibles.

RESULTADOS

En la figura 2 se presentan las curvas del porcentaje de autoconsumo y autosuficiencia en función de la potencia pico instalada para la localidad de Málaga, para un perfil tipo 1 y para una instalación de autoconsumo sin acumulación. Como puede verse al aumentar la potencia pico instalada el autoconsumo va disminuyendo mientras que la autosuficiencia aumenta. El punto de corte de las dos curvas corresponde a la potencia pico a la cual la energía anual producida por el sistema FV iguala al consumo anual. En este punto, la vivienda en cuanto a energía eléctrica sería de cero emisiones. La curva de autosuficiencia va saturando y su valor teórico máximo sería el porcentaje de consumo que se produce en horas de Sol, ya que en un sistema sin acumulación, el consumo nocturno no puede ser satisfecho.

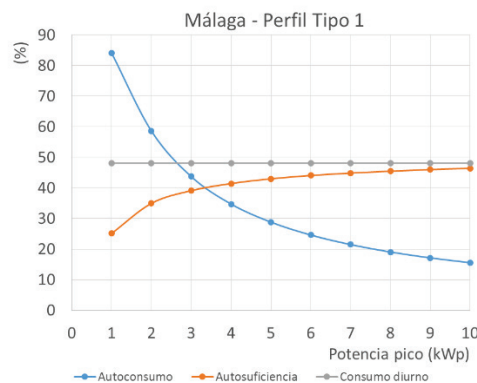


Fig. 2. Curvas de porcentajes anuales de autoconsumo y autosuficiencia en función de la potencia pico para una instalación en Málaga con un perfil tipo 1.

En las figuras 3 y 4 se presentan estos resultados, pero para una instalación FV con sistema de acumulación. Ahora los resultados dependen tanto de la potencia pico instalada como de la capacidad de la batería. Podemos ver que un mismo porcentaje de autosuficiencia puede ser cubierto con distintas configuraciones del sistema. Lo no linealidad de las curvas hace que siempre exista un valor por encima del cual aumentar el tamaño del generador o del sistema de acumulación no influye de forma significativa en el porcentaje de autosuficiencia, es decir, en el ahorro energético en la factura eléctrica. Estos resultados son distintos para cada localidad y para cada perfil de consumo, por lo que tiene interés el intento realizado en la generalización de los resultados.

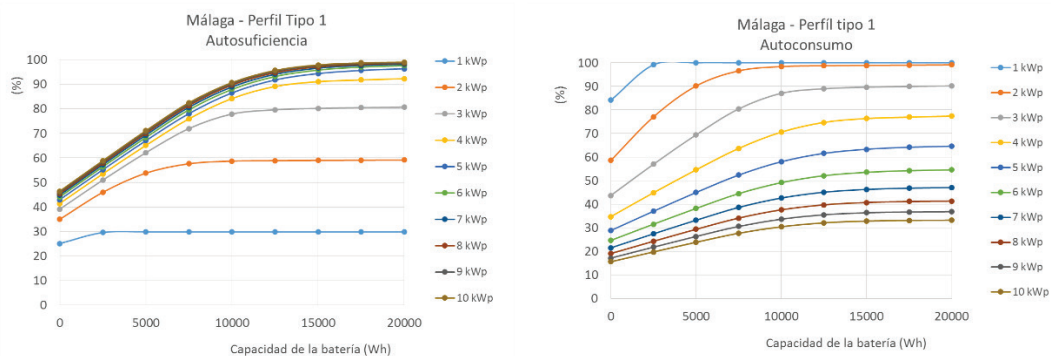


Fig. 3 y 4. Curvas de porcentajes anuales de autosuficiencia y autoconsumo en función de la potencia pico y de la capacidad de acumulación para un sistema FV en Málaga con un perfil de consumo tipo 1.

En este trabajo se han considerado aquellas variables independientes de los modelos que tienen una influencia directa en el valor de autosuficiencia. En concreto, las variables propuestas son:

- Tipo de perfil de consumo
- Valor medio anual de radiación global diaria
- Desviación típica de los valores medios diarios de radiación global
- Valor medio anual de temperatura diaria
- Desviación típica de los valores diarios de temperatura global
- Potencia pico de la instalación
- Capacidad de la batería.

Utilizando todas estas variables independientes se han calculado los valores de la autosuficiencia y del autoconsumo para cada una de las combinaciones posibles. En total, se obtuvieron 1056 valores diferentes para cada uno de estos dos parámetros. La estimación de los parámetros de cada uno de ellos, así como de las distintas métricas de error se ha hecho utilizando la herramienta Weka (Hall et al., 2009).

Por una parte, se han evaluado los errores obtenidos cuando se utiliza correlación cruzada. Por otra parte, el conjunto total de datos se ha dividido en conjunto de entrenamiento y conjunto de test, tal y como se hace en los modelos de minería de datos cuando no se utiliza validación cruzada. El conjunto de entrenamiento se utiliza para ajustar los modelos (obtener sus parámetros). El conjunto de test se utiliza para evaluar las distintas métricas de error. Se han utilizado un 80 % de las muestras para el conjunto de entrenamiento y el 20 % restante para el conjunto de test.

Para la estimación del autoconsumo se ha utilizado también como variable independiente el valor de la autosuficiencia, ya que los modelos propuestos para estimar la autosuficiencia y el autoconsumo pueden aplicarse en cascada, y utilizar para uno de ellos el valor obtenido del otro (el primero que se estime). En nuestro caso, el orden sería primero estimar el valor de autosuficiencia y después el de autoconsumo.

Los resultados obtenidos para la estimación de la autosuficiencia y el autoconsumo se muestran en las tablas 2 y 3, respectivamente.

Como puede observarse, todos los modelos propuestos, excepto la regresión lineal, tienen coeficientes de correlación superiores a 0,98. Entre ellos, el modelo con el que se obtienen mejores predicciones de autosuficiencia es Random Forest, que en validación cruzada tiene un error absoluto promedio inferior a 1 y un error relativo del 5%. También se obtienen muy buenos resultados para los modelos M5 'y REPTree. En todos los casos, el error relativo absoluto es inferior al 10%, lo que significa para una amplia gama de valores de autosuficiencia errores en puntos porcentuales del orden del 1%.

En la figura 5 se muestran los valores de autosuficiencia obtenidos mediante la simulación del comportamiento de la instalación frente a los valores obtenidos con los modelos Random Forest y REPTree.

Tabla 2. Métricas obtenidas para cada uno de los modelos analizados al estimar la autosuficiencia mediante validación cruzada

Métrica	RL	MP	RF	M5P	REPTree
R	0.798	0.9835	0.9986	0.9927	0.9925
MAE	11.1385	3.2984	0.9364	1.8181	1.6386
RMSE	14.0129	4.2624	1.5568	2.8995	2.8507
RAE	60.03%	17.78%	5.05%	9.80%	8.83%
rRSE	60.23%	18.32%	6.69%	12.46%	12.25%
N	1056	1056	1056	1056	1056

Tabla 3. Métricas obtenidas para cada uno de los modelos analizados al estimar el autoconsumo mediante validación cruzada

Métrica	RL	MP	RF	M5P	REPTree
R	0.8599	0.9992	0.9963	0.9963	0.9912
MAE	10.2565	0.8015	1.1807	1.4163	1.7489
RMSE	12.9934	1.0402	2.2602	2.2642	3.3616
RAE	51.73%	4.04%	5.95%	7.14%	8.82%
rRSE	50.94%	4.08%	8.86%	8.88%	13.18%
N	1056	1056	1056	1056	1056

Para la estimación del autoconsumo, el modelo que mejor funciona es el perceptrón multicapa, con un error absoluto promedio de 0,55 y un error relativo del 3%. Los modelos de árbol de decisión también tienen pequeños errores, especialmente Random Forest y M5P. En la Figura 6 se muestran los valores de autoconsumo frente a las predicciones de estos valores para los modelos de perceptrón multicapa y Random Forest.

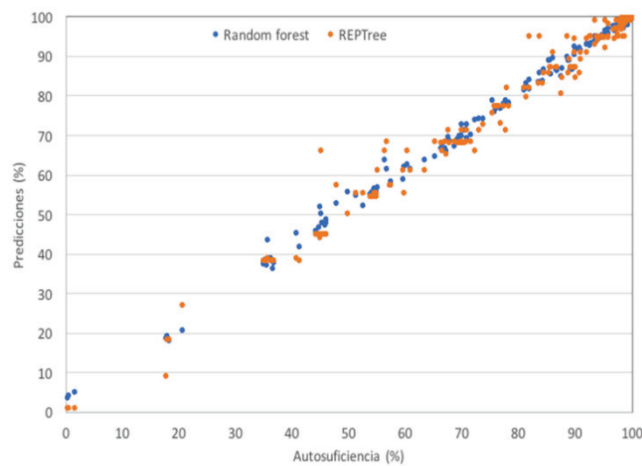


Figura 5. Valores de autosuficiencia frente a las predicciones de estos valores para los modelos Random Forest y REPTree.

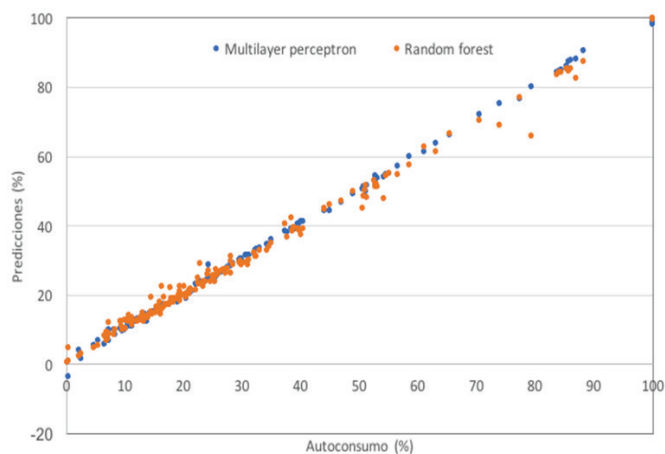


Figura 6. Valores de autoconsumo frente a las predicciones de estos valores para los modelos perceptrón multicapa y Random Forest.

CONCLUSIONES

Los resultados obtenidos muestran que sólo desde un balance horario de los datos de consumo a lo largo de un año y del estudio de la sensibilidad de los coeficientes de autoconsumo y autosuficiencia en función de la potencia pico instalada y del tamaño de la batería, es posible optimizar el dimensionado de los sistemas fotovoltaicos de autoconsumo. Sin embargo, este estudio demuestra que es posible una generalización de este proceso con buenos resultados y que simplifica el proceso de cálculo. Para este fin, se han evaluado diferentes modelos de minería de datos para la modelización de los parámetros de autoconsumo y autosuficiencia. Específicamente, se ha utilizado una regresión lineal, un perceptrón multicapa y varios tipos de árboles de decisión. En todos los casos, excepto en la regresión lineal, los resultados obtenidos permiten afirmar la validez de los diferentes modelos. Para el autoconsumo, el modelo con mejores resultados es el perceptrón multicapa, mientras que para la autosuficiencia el que obtiene resultados más precisos es Random Forest. El coeficiente de correlación es superior en ambos casos a 0,99, mientras que el error absoluto medio es inferior a 1 y el error medio relativo es del 3% para el autoconsumo y del 7% para la autosuficiencia. Los mejores resultados obtenidos en la estimación del autoconsumo con respecto a la autosuficiencia se explican porque en el primer caso se han utilizado los valores de autosuficiencia como variable independiente; esto es posible en situaciones reales ya que la estimación de estos parámetros se puede hacer en dos fases, ya que para cada una de ellas se ha estimado un modelo diferente.

AGRADECIMIENTOS

Este trabajo ha sido apoyado por el proyecto de investigación, RTI2018-095097-B-I00 (2108) del Ministerio de Ciencia, Innovación y Universidades, España. Se enmarca dentro de la actividad de la Fundación Renovables en el desarrollo de proyectos y ha sido en parte financiado por subvención del Ministerio para la Transición Ecológica.

REFERENCIAS

- Ayala-Gilardón, M. Sidrach-de-Cardona, L. Mora-López. (2018) Influence of time resolution in the estimation of self-consumption and self-sufficiency of photovoltaic facilities. *Applied Energy*. 229. pp.990-997
- Alfadda, R. Adhikari, M. Kuzlu and S. Rahman, (2017). Hour-ahead solar PV power forecasting using SVR based approach. IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, pp. 1-5.
- Breiman, L. Random forests. *Machine Learning*,(2001). 45 (1), pp. 5–32.
- Filipe, J.M. Bessa, R.J. Sumaili, J. Tomé, R. and Sousa, J.N. (2015). A hybrid short-term solar power forecasting tool. *18th International Conference on Intelligent System Application to Power Systems (ISAP)*, Porto, pp. 1-6.
- Hall, M. Frank, E. Holmes, G. Pfahringer, B. Reutemann, P. I. H.(2009). Wit- ten, The weka data mining software: an update, SIGKDD Explorations Newsletter 11, 10–18.
- Sardá-Espinosa, A. Subbiah, S. Bartz-Beielstein, T.(2017). Conditional inference trees for knowledge extraction from motor health condition data. *Engineering Applications of Artificial Intelligence*, 62 (Supplement C), pp.26-37.
- Sartori, I. Napolitano, A. and Voss, K. (2012). Net zero energy buildings: A consistent definition framework. *Energy and Buildings*, 48, pp.220–232.
- Sharma, N. Sharma, P. Irwin, D. and Shenoy, P.(2011). "Predicting solar generation from weather forecasts using machine learning," 2011 *IEEE International Conference on Smart Grid Communications*, Brussels, pp. 528-533.
- Sivaneasan, B. Yu, C.Y. Goh, K.P.(2017). Solar Forecasting using ANN with Fuzzy Logic Pre-processing, *Energy Procedia*, vol. 143, pp. 727– 732.
- Wang, Y. Witten, I.H.(1997). Induction of model trees for predicting continuous classes. In: *9th European Conference on Machine Learning* (poster papers). Springer, pp. 128 – 137.