

# MODELOS DE PREDICCIÓN A CORTO PLAZO PARA SISTEMAS FOTOVOLTAICOS

A. Muñoz-Jiménez, P. Zorzano-Santamaría, L. A. Fernández-Jiménez, E. Zorzano-Alba, A. Yanguas-Peña, y E. García-Garrido.

Departamento de Ingeniería Eléctrica

E.T.S.I.I., Universidad de La Rioja

C/ Luis de Ulloa, 20; 26002 Logroño, La Rioja (España)

Tel.: +34 941 299477, Fax: +34 941 299478, e-mail: pedrojose.zorzano@unirioja.es, andres.munoz@unirioja.es

## Resumen.

El incremento de precios de los combustibles convencionales, las amenazas derivadas del cambio climático, y las políticas energéticas impulsadas en los países desarrollados, han propiciado un fuerte incremento en la potencia instalada en plantas generadoras de energía eléctrica basadas en recursos renovables. Las instalaciones basadas en energía eólica son las que han experimentado un mayor incremento en los últimos años, seguidas por las basadas en energía solar fotovoltaica. A la par del aumento de instalaciones, se ha producido un aumento de la capacidad de la instalación media.

Este artículo presenta los resultados obtenidos con un conjunto de modelos de predicción de la generación eléctrica vertida a la red por una instalación fotovoltaica. El valor de predicción corresponde a la energía eléctrica generada para cada una de las 24 horas del día siguiente, suponiendo que el instante de realizar las predicciones es el de las 9:00.

Los modelos desarrollados se basan en el uso de herramientas de predicción numérica meteorológica y técnicas avanzadas de predicción, basadas fundamentalmente en redes neuronales artificiales. Los mejores modelos ofrecen unos errores de predicción inferiores al 15% de la potencia instalada, para los períodos de predicción que cubren las 24 horas del día siguiente.

## Palabras clave

Predicción a corto plazo, sistemas fotovoltaicos, generación distribuida, redes neuronales artificiales.

## 1. Introducción

Las energías renovables han captado la atención tanto de las autoridades como de la opinión pública en los últimos años. El aprovechamiento de las energías renovables implica importantes ventajas ambientales, entre ellas el hecho de que se tratan de energías con una contaminación mínima y además no dependen de fuentes energéticas exteriores. Además las instalaciones de generación eléctrica basadas en recursos renovables

pueden servir como factor de estabilización de la población en zonas con riesgo de despoblación.

La energía eólica ha sido la energía renovable que más se ha desarrollado en los últimos años. La potencia instalada en los parques eólicos existentes en todo el mundo, ha pasado de 7.5 GW en 1997 a 121 GW a finales de 2008, alcanzando así el 1.5% de la energía eléctrica generada en todo el mundo [1]. En un primer momento, los parques eólicos vertían a la red eléctrica la energía eléctrica generada, recibiendo la compensación económica correspondiente. Estas compensaciones podían incluir subvenciones o primas para hacer competitivas a este tipo de instalaciones y estimular así su proliferación. Pero la reducción de los costes de instalación y nuevas políticas energéticas, con nuevos incentivos, propiciaron que en algunos países la mayoría de los parques eólicos pudieran vender la energía eléctrica producida en el mercado eléctrico. Este es el caso de España, donde desde el año 2004 la mayoría de los parques eólicos actúan como un productor más en el mercado eléctrico.

Las reglas del mercado eléctrico obligan a que cada productor proporcione al operador del mismo, en los periodos establecidos, los valores de la energía que va a verter en la red para cada una de las horas cubiertas por la sesión correspondiente. La sesión más importante en el mercado eléctrico es la sesión diaria, para la que los productores deben presentar sus valores de generación para cada una de las 24 horas del día siguiente. La hora a la que se celebra la sesión del mercado diario difiere según el país; en el caso de España, esta sesión tiene lugar a las 10:00.

Si los valores de energía ofertada no se corresponden con la realidad, el productor es penalizado de acuerdo a las reglas del mercado. Obviamente, los valores de generación aportados por el gestor de un parque eólico no son sino predicciones de generación. Por tanto, cuanto menor sea el error cometido en estas predicciones menor

es la penalización y por tanto, mayor es el beneficio obtenido por la venta de la energía producida. La necesidad, por parte de los gestores de los parques eólicos, de disponer de predicciones de la generación eléctrica en sus instalaciones ha impulsado el desarrollo de modelos de predicción que determinen, con el menor error posible, la producción eléctrica horaria para el día siguiente. Así, en los últimos años se han desarrollado un conjunto de modelos de predicción de la generación eléctrica a corto plazo en un parque eólico por varios equipos de investigación [2].

La capacidad (potencia instalada) de las instalaciones fotovoltaicas conectadas a la red eléctrica ha ido creciendo en los últimos años. Se ha pasado de pequeñas instalaciones cuya potencia instalada era, como mucho, de decenas de kW, a las más modernas instalaciones con una potencia instalada de decenas de MW. La potencia instalada en estos parques solares fotovoltaicos recuerda a la de los parques eólicos hace poco más de una década. Por otra parte, el esperado abaratamiento de costes en paneles fotovoltaicos [3] propiciará que esta tecnología sea en un futuro próximo competitiva, en términos económicos, frente a otras fuentes de energía. Así, el incremento de la potencia en instalaciones fotovoltaicas, unido a un abaratamiento de los costes de producción de energía eléctrica en las mismas, permitirá en un futuro próximo que estas instalaciones puedan, con incentivos o sin ellos, participar en los mercados eléctricos para vender la energía producida.

Por ello, los gestores de los parques solares fotovoltaicos conectados a red necesitarán modelos de predicción que les permitan determinar de la manera más exacta posible la producción horaria de energía eléctrica con una anticipación cercana a los dos días. Varios grupos de investigación internacionales están trabajando en el desarrollo de herramientas de este tipo, aunque los modelos publicados predicen los valores de la radiación solar en vez de la producción de energía eléctrica [4].

Este artículo presenta los resultados obtenidos en la predicción de la generación de energía eléctrica, hora a hora, en un parque solar fotovoltaico real para el periodo que comprende las siguientes 39 horas. Los modelos desarrollados usan herramientas de predicción numérica meteorológica, y técnicas avanzadas de predicción basadas, fundamentalmente, en redes neuronales artificiales. Los mejores modelos ofrecen unos errores RMSE (raíz del error medio cuadrático) inferiores al 15% de la potencia instalada, para períodos de predicción que cubren las 24 horas del día siguiente.

## 2. Modelos de predicción

El objetivo perseguido con el trabajo de investigación descrito en el presente artículo es el desarrollo de un modelo de predicción de la energía media horaria para todas las horas del día siguiente generadas en una instalación solar fotovoltaica conectada a la red de distribución de energía eléctrica.

El fin del modelo de predicción es proporcionar un valor para la energía eléctrica generada por un parque solar fotovoltaico concreto. Concretamente, los modelos de predicción desarrollados predicen el valor futuro de la energía media horaria generada a partir de los valores pasados de esta misma variable y de los valores previstos para una serie de variables meteorológicas, obtenidos por medio de un modelo de predicción numérica meteorológica, NWP (Numerical Weather Prediction).

### A. Datos de entrada para los Modelos de Predicción

Los datos de entrada utilizados para el desarrollo de los modelos de predicción que se presentan en este artículo han sido los valores de energía eléctrica media horaria generada en un parque solar fotovoltaico real, conectado a la red eléctrica, situado en La Rioja, y los valores de predicción para la radiación y la temperatura al nivel de superficie terrestre correspondientes a la localización del parque solar, para las siguientes 48 horas. Estos últimos valores han sido obtenidos por medio del modelo MM5 [5], un modelo NWP, para horizontes de 0 a 48 horas en intervalos de 15 minutos.

La instalación fotovoltaica utilizada está compuesta de varios grupos de paneles solares con diferentes tecnologías (fijos, seguidor a un eje, seguidor a dos ejes). La potencia total de la instalación es de 36 kW pico. La serie temporal disponible, con los valores de generación de energía eléctrica sobre una base horaria, comprende 362 días (del 2 de Junio de 2007 al 27 de Mayo de 2008). La Fig. 1 representa la energía generada para los primeros 5 días de la serie temporal.

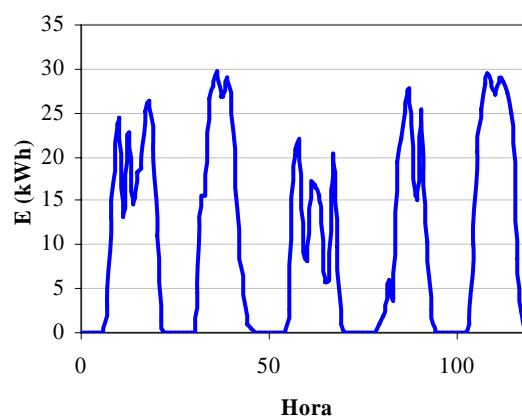


Fig. 1 Energía generada del 2 al 6 de Junio 2007

A la vista de la Fig. 1, se pueden observar las considerables diferencias entre los valores de la generación de energía eléctrica de un día para otro, lo cuál dificulta en gran medida el desarrollo de un modelo de predicción óptimo y fiable.

### B. Criterios de evaluación de los Modelos de Predicción

La adecuación de los modelos no se evalúa con los datos utilizados en su ajuste o entrenamiento, sino con otro conjunto de valores (valores de test) de las variables explicativas. Esta prueba permite la comparación de los resultados (valores de predicción) entre diferentes

modelos y de esta forma seleccionar el modelo que arroje los menores errores de predicción.

Para la comparación general entre los modelos se ha usado la raíz del error cuadrático medio, **RMSE**, ya sea en su valor real, ecuación (1), o en su valor normalizado a la potencia instalada, ecuación (2). En estas ecuaciones  $N$  representa el número total de datos y  $k$  el horizonte de predicción.

$$RMSE(k) = \sqrt{\frac{1}{N} \sum_{t=1}^N (e(t+k|t))^2} \quad (1)$$

$$RMSE_n(k) = \frac{1}{P_{inst}} \sqrt{\frac{1}{N} \sum_{t=1}^N (e(t+k|t))^2} \quad (2)$$

Además también se calculará el error medio, **ME**, y el error absoluto medio, **MAE**, cuyas expresiones son las indicadas en las ecuaciones (3)-(4) y (5)-(6), respectivamente. Estos criterios se han utilizado para comparaciones más extensas, en el presente artículo se mostrarán los valores de estos estadísticos para el modelo de predicción que mejores resultados proporciona, es decir, el que presenta menores valores del error de predicción.

$$ME(k) = \frac{1}{N} \sum_{t=1}^N e(t+k|t) \quad (3)$$

$$ME_n(k) = \frac{1}{P_{inst}} \frac{1}{N} \sum_{t=1}^N e(t+k|t) \quad (4)$$

$$MAE(k) = \frac{1}{N} \sum_{t=1}^N |e(t+k|t)| \quad (5)$$

$$MAE_n(k) = \frac{1}{P_{inst}} \frac{1}{N} \sum_{t=1}^N |e(t+k|t)| \quad (6)$$

### 3. Modelo MM5

El MM5 es un modelo de mesoescala, desarrollado por PSU/NCAR (Pennsylvania State University and National Center for Atmospheric Research), de área limitada, no hidrostático, basado en el uso de coordenadas sigma (niveles de presión), y diseñado para simular o predecir la circulación atmosférica en la escala regional. Ha sido desarrollado por medio de contribuciones de diferentes grupos de investigación internacionales durante los últimos años. El modelo MM5 está implementado como un programa informático capaz de proporcionar predicciones de variables meteorológicas con una gran resolución espacial pero con una validez temporal limitada (baja escala temporal). Este tipo de modelos NWP son conocidos como modelos mesoescalares.

El MM5 ha sido compilado en varios ordenadores, con sistema operativo Linux, en instalaciones de la Universidad de La Rioja para la predicción de los valores de radiación y temperatura al nivel de superficie para la localización del parque solar fotovoltaico. Con el

propósito de incrementar la validez temporal de las predicciones (horizontes de predicción hasta 48 horas), el modelo MM5 utiliza como datos de entrada los valores de predicción de una serie de variables meteorológicas obtenidas de un modelo global, el modelo GFS (Global Forecasting System del National Weather Services, USA). El modelo GFS proporciona predicciones para un conjunto de puntos repartidos por todo el mundo a intervalos de tres horas y con horizontes de predicción de 0 hasta 180 horas. Los datos del modelo GFS son descargados diariamente de servidores norteamericanos con las predicciones correspondientes a la asimilación (medida) de los datos atmosféricos a las 0:00.

### 4. Desarrollo de los Modelos

Los modelos elegidos para el desarrollo del sistema de predicción incluyen el clásico modelo persistente, modelos desarrollados bajo la metodología de Box-Jenkins (o de la familia *ARIMA*) [6], un modelo conocido como “modelo de los  $k$ -vecinos más próximos” (*k-nearest neighbors*) [7], modelos basados en redes neuronales artificiales y modelos basados en sistemas de inferencia neurodifusa (*neurofuzzy*). Para el desarrollo de los sistemas con redes neuronales artificiales se utilizó un proceso de optimización, basado en algoritmo genético, que permitía seleccionar las entradas (de entre las disponibles), el número de neuronas de capas ocultas y los parámetros del algoritmo de entrenamiento utilizado (retro-propagación de error).

#### A. Consideraciones Previas

Para el desarrollo de los modelos de predicción, los datos de los 362 días (8688 registros horarios) se dividieron en tres grupos, entrenamiento, validación cruzada y test. El grupo de entrenamiento engloba el 60 % de los datos, el grupo de validación un 20 % de los datos y el grupo de test el 20 % restante.

Concretamente, se han utilizado dos tipos de división o clasificación de los datos:

- **Tipo A:**
  - Entrenamiento: el primer 60 % de los datos de cada mes.
  - Validación: el siguiente 20% de los datos de cada mes.
  - Test: el 20% restante de los datos de cada mes.
- **Tipo B:**
  - Entrenamiento: el primer 60 % de todo el intervalo de los datos.
  - Validación: el siguiente 20% de todo el intervalo de los datos.
  - Test: el 20% restante de todo el intervalo de los datos.

En caso de haber dispuesto de una serie temporal superior a dos años, como datos de entrada, lo más recomendable sería realizar una sola clasificación, concretamente la *Tipo B*. Pero puesto que para el desarrollo de los modelos descritos, se disponía de

únicamente de los datos de un año, a priori, los mejores resultados se obtendrán con la clasificación *Tipo A*, ya que garantiza la disponibilidad de datos de todos los meses para los datos de entrenamiento, validación y test. En la actualidad, se están recopilando más datos para poder mejorar los modelos descritos en el presente artículo.

Para el ajuste de los modelos se usan los datos de entrenamiento y validación. En los modelos basados en redes neuronales artificiales y sistemas de inferencia neurodifusa el grupo de validación se usa para evitar que el modelo sufra un “sobreentrenamiento, es decir, para evitar que el modelo se ajuste demasiado a los datos de entrenamiento, a consecuencia de lo cuál no se obtendría un modelo de predicción general.

La adecuación de los modelos no se evalúa con los datos utilizados en su ajuste o entrenamiento, sino con otro conjunto de valores (valores de test) de las variables explicativas. Este ensayo permite la comparación de los resultados (errores de predicción) entre diferentes modelos y de esta forma seleccionar el modelo que arroje los mejores resultados (menores errores de predicción). Cuando se muestren los errores obtenidos con los diferentes modelos se especificarán para su correspondiente grupo de prueba, aunque para comparar modelos con diferente división de datos (*Tipo A* ó *B*) se incluirá también el error de los datos de prueba comunes de ambos tipos de división. Así, cuando hablemos de los diferentes modelos desarrollados se especificará el tipo de división de datos utilizado. En el caso de modelos *ARIMA*, o de redes neuronales recurrentes, es obvio que el único tipo de clasificación válido es el *Tipo B*, ya que en estos modelos los datos de entrenamiento deben estar temporalmente ordenados (ser correlativos).

#### B. Modelos de Predicción analizados

**Modelo persistente:** Se ha modificado el clásico modelo persistente (el valor de predicción es igual al último valor conocido) por un modelo persistente modificado por el que la energía eléctrica generada para cualquier hora de mañana es igual a la de la misma hora de ayer (horas entre las 9:00 y las 24:00) o de hoy (horas entre las 0:00 y las 9:00).

**Modelo Box-Jenkins o *ARIMA*:** con la metodología Box-Jenkins se construyeron dos modelos *ARIMA*. El primero de ellos (*ARIMA* 1) es el que ofrecía los mejores resultados para un horizonte de predicción de 1 hora, extendiéndose posteriormente los horizontes de predicción hasta abarcar todo el día siguiente. El segundo modelo (*ARIMA* 2) se desarrolló utilizando exclusivamente los datos de la serie temporal correspondiente al día anterior considerando exclusivamente la estacionalidad. Para el ajuste del modelo se utilizaron los datos de los grupos de entrenamiento y de validación.

**Modelo de los *k*-vecinos más próximos:** buscamos entre todos los datos previos de los que dispongamos, es decir, para todos los días anteriores a hoy, los registros cuyas

entradas (valores de predicción del modelo MM5) sean más similares en términos de distancia euclídea a las disponibles para cada una de las horas del día siguiente. Con el valor de energía generada para cada uno de esos registros se calcula la media aritmética entre ellos o la media ponderada (el factor de ponderación es la inversa de la distancia euclídea).

**Modelos basados en red neuronales artificiales:** se utilizó su potencial como “aproximadores universales” de funciones (incluyendo, evidentemente, las no lineales), para el desarrollo de los modelos. Para el ajuste de estos modelos se utilizaron los datos de los grupos de entrenamiento y de validación.

**Modelos de inferencia neurodifusa:** se desarrollaron diferentes tipos de sistemas de inferencia neurodifusa (*ANFIS*) [8], todos con partición del espacio de entrada por agrupamiento sustractivo. Se utilizaron los datos de los grupos de entrenamiento y validación en el desarrollo de los modelos.

En el presente artículo no se presentan todos los modelos analizados y desarrollados, sino que se incluyen los de mayor relevancia y mejores resultados.

## 5. Resultados Obtenidos

Tras el proceso de ajuste o entrenamiento de los modelos, y la aplicación de los mismos en la predicción de la energía generada a intervalos de una hora, se obtuvieron los resultados que se van a pormenorizar en los próximos párrafos.

El modelo persistente se aplicó directamente sobre los datos del grupo de test.

Los dos modelos *ARIMA* se diseñaron a partir de los datos de la energía generada, correspondientes a los grupos de entrenamiento y validación. Para el primer modelo, *ARIMA* 1, se desarrollaron modelos auto-regresivos puros (con órdenes de 1 a 3) y modelos auto-regresivos de media móvil (el de mayor orden fue el modelo *ARMA*(2,2)) sin considerar la estacionalidad; y además, se construyeron otros modelos similares considerando la estacionalidad (con periodo de estacionalidad de 24 muestras, es decir, 24 horas). Para el caso del segundo modelo, *ARIMA* 2, se construyeron modelos considerando sólo la estacionalidad. Una vez diseñados y evaluados los diversos modelos, se escogieron aquellos que proporcionaban un valor menor del error *RMSE* con los datos utilizados en el ajuste del modelo (coincidían en ser también los modelos con menor valor del criterio de información de Akaike, *AIC*, y aquellos cuyas gráficas de función de auto-correlación y auto-correlación parcial también sugerían que el modelo de predicción debería tener su misma estructura). El modelo *ARIMA* 1 escogido fue el *ARIMA* (3,0,2)(1,0,0)<sub>24</sub>, y el modelo *ARIMA* 2 fue el *ARIMA*(1,0,1)<sub>24</sub>. Para comparar los errores obtenidos por ambos modelos con el resto de modelos, se toma la media del error correspondiente a los horizontes de predicción que incluyen todo el día siguiente.

Para el modelo de los  $k$ -vecinos más próximos se han considerado de uno a diez vecinos, es decir, como mínimo tomando el registro anterior más aproximado o similar (en términos euclideos) y como máximo la media (aritmética o ponderada) de los diez registros más aproximados. Para comparar con el resto de modelos, se tomará el error correspondiente al modelo de los 10 vecinos más próximos con media ponderada por ser el que mejores resultados ofrecía.

En los modelos de predicción basados en los sistemas de inferencia neurodifusa se utilizó la partición del espacio de entrada por agrupamiento sustractivo, escogiendo los parámetros de éste de forma que el número de familias de pertenencia y reglas tuviese como valor máximo el de 5.

En el desarrollo de los modelos basados en redes neuronales artificiales se permitió que el proceso de optimización (con algoritmo genético) pudiera elegir las variables de entrada de las redes neuronales de entre todas las disponibles. También el proceso de optimización permitía seleccionar el número de neuronas en capas ocultas y los valores del factor de aprendizaje y momento correspondientes al algoritmo de entrenamiento. El número de generaciones se estableció en 50, al igual que el número de individuos por generación. Las tasas de cruce y mutación se establecieron en 0.9 y 0.02, respectivamente. La función de transferencia para las neuronas de capas ocultas fue la tangente hiperbólica. Las redes neuronales desarrolladas incluyen: perceptrones multicapa, MLP; modulares (dos MLP en paralelo), MNM; redes recurrentes de Elman; redes neuronales de función de base radial, RBF; y redes neuronales retardadas en el tiempo, TDNN, con unidades de retardo sólo en la capa de entrada.

La Tabla I muestra los resultados obtenidos con los diferentes modelos, concretamente se representan los errores  $RMSE$  en la predicción de la energía eléctrica generada para el grupo de test propio del modelo y para el grupo de test común entre la división de datos *Tipo A* y *B*. Es decir, para comparar modelos que tengan el mismo tipo de división de datos el mejor índice de comparación es el  $RMSE_{Propio}$  y para comparar modelos con distinto tipo de división de datos (uno *Tipo A* con otro *Tipo B*) el mejor índice es el error  $RMSE_{Común}$ . El nombre de cada modelo, en la Tabla I, hace referencia al tipo de modelo, así como a otras características referentes al tipo de clasificación de los datos.

De todos los modelos desarrollados con datos *Tipo A* destacamos los dos modelos con menor  $RMSE_{Propio}$ , y además los dos mejores entre los modelos desarrollados con datos *Tipo B* (con su  $RMSE_{Propio}$ ) y con *Tipo A* (con su  $RMSE_{Común}$ ). Así, el modelo más adecuado para la predicción de la energía eléctrica generada por parques solares fotovoltaicos es el **MLP 2A**. Concretamente, presenta un valor medio del  $RMSE$  para su grupo de test del 11.79 %, correspondiente al horizonte de predicción que engloba todo el día siguiente (horizontes de 16 a 39 horas).

El modelo **MLP 2A**, se trata de una red neuronal del tipo perceptrón multicapa (*MultiLayer Perceptron*), con dos capas ocultas. La estructura de esta red neuronal es la reflejada en la Fig. 2, en la que las capas a la izquierda y derecha de la figura corresponden a la capa de entradas y a la capa de salidas, respectivamente. Las capas ocultas son las dos capas intermedias. Cada capa sólo está unida (conexiones entre neuronas) a la capa precedente y a la capa que le sigue (si existe). Concretamente, el modelo estará compuesto por 7 neuronas en la primera capa oculta y otras 7 en la segunda capa oculta. La Tabla II recoge el error absoluto medio ( $MAE$ ) y el error medio ( $ME$ ) del modelo de predicción MLP 2A.

Tabla I. Resultados obtenidos

Modelo	RMSE <sub>Propio</sub>		RMSE <sub>Común</sub>	
	(Wh)	(%)	(Wh)	(%)
Persistente	7530.91	20.92		
ARIMA 1	7609.12	21.14		
ARIMA 2	6249.30	17.36		
10- Vecinos	6041.25	16.78		
MLP 1A	<b>4255.19</b>	<b>11.82</b>	4967.25	13.80
MLP 1B	5522.35	15.34	5452.58	15.15
MLP 2A	<b>4243.01</b>	<b>11.79</b>	<b>4740.19</b>	<b>13.17</b>
MLP 2B	4999.16	13.89	5008.43	13.91
MNM 1B	5341.96	14.84	5523.97	15.34
MNM 2A	4377.83	12.16	4903.67	13.62
Elman 1B	4964.37	13.79		
RBF 1B	5442.78	15.12	5648.83	15.69
RBF 2A	4429.43	12.30	4960.24	13.78
TDNN F	5315.29	14.76		
ANFIS 1B	5466.15	15.18		
ANFIS 2A	4375.21	12.15	4872.24	13.53
ANFIS 2B	5072.34	14.09		

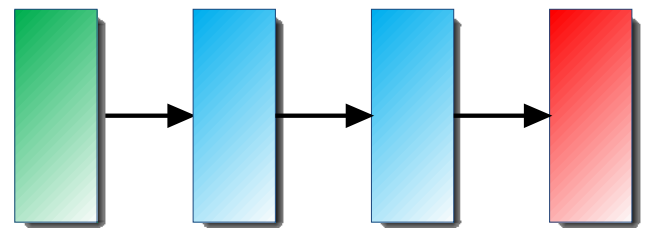


Fig. 2. Estructura de la red neuronal MLP 2A.

Tabla II. Errores MAE Y ME para el modelo MLP 2A

MAE <sub>Propio</sub>		MAE <sub>Común</sub>		ME <sub>Propio</sub>		ME <sub>Común</sub>	
(Wh)	(%)	(Wh)	(%)	(Wh)	(%)	(Wh)	(%)
2308.3	6.41	2524.3	7.01	99.5	0.28	-42.8	-0.12

Para este modelo de predicción basado en la red neuronal MLP 2A el error absoluto medio oscila entre 6.41 y 7.01%, y el error medio cometido con todo el grupo de datos de test ha sido inferior al 0.28%, lo cuál quiere decir, que la suma de la energía prevista a lo largo de todo el período de predicción (24 horas del día siguiente) y la energía real producida se diferencian muy poco, tal como refleja la Tabla III.

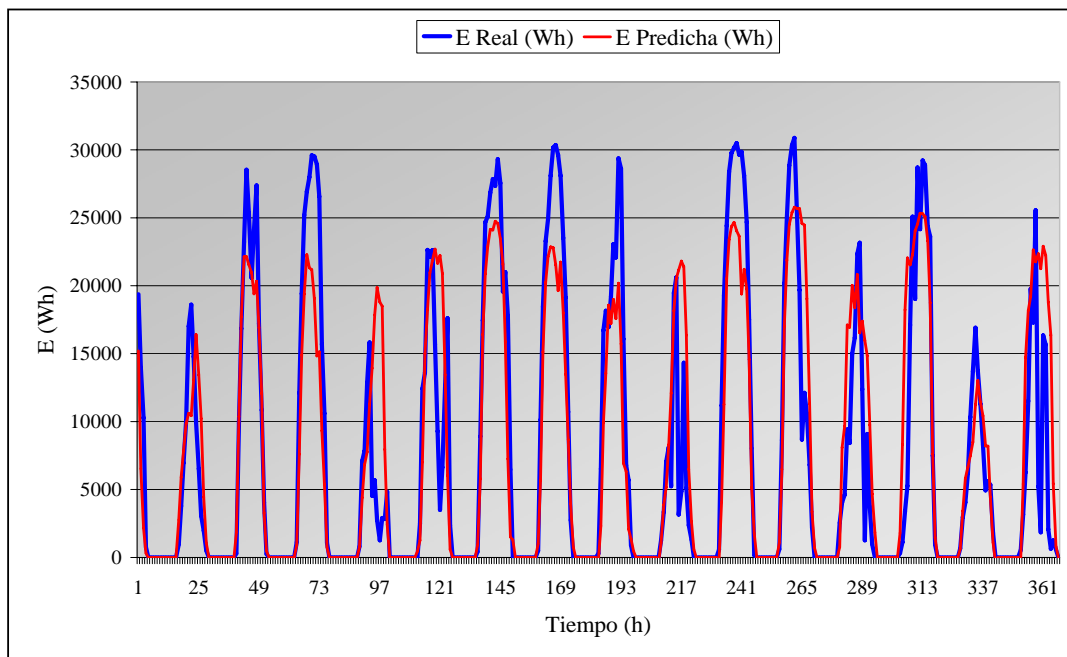


Fig. 3. Gráfico de comparación entre la energía real generada y la prevista con el modelo MLP 2A.

La Fig. 3 representa los valores de la energía prevista y la real generada para todo el grupo de test común con el modelo MLP 2A. El horizonte de predicción va desde 16 horas (primer valor de cada día) hasta 39 horas (último valor de cada día), ya que las predicciones están realizadas a las 9:00 horas del día anterior.

Tabla III. Diferencia entre la energía total real y prevista

	$E_{TReal}$ (Wh)	$E_{TPrevista}$ (Wh)	Error (%)
Propio	12142112	11970986.21	1.41
Común	2849201	2866973.48	0.62

## 6. Conclusiones

El presente artículo muestra los resultados obtenidos mediante una serie de modelos de predicción basados en diferentes técnicas, en la predicción de la producción horaria de energía eléctrica para un parque solar fotovoltaico conectado a la red eléctrica. Este tipo de modelos pueden llegar a ser muy útiles en un futuro próximo, y no sólo para los gestores de grandes parques solares fotovoltaicos, sino también para los propietarios de pequeñas instalaciones fotovoltaicas.

Por tanto, a modo de resumen, podemos decir que la investigación descrita en el artículo, trata de determinar, de la manera más aproximada posible, la energía que va a ser generada por el parque solar fotovoltaico a corto plazo, concretamente para las 24 horas del día siguiente, lo cuál nos proporcionar varias ventajas, pudiendo destacar las siguientes:

- Poder utilizar los valores de predicción para presentar ofertas en el mercado eléctrico.
- Conseguir el máximo beneficio económico por la venta de la energía al reducirse las penalizaciones.

En la actualidad, el equipo de investigación sigue implicado en el desarrollo de este tipo de modelos de predicción, mejorando los modelos descritos en el presente artículo.

## Agradecimientos

Los autores agradecen el apoyo prestado por el Gobierno de La Rioja al financiar este trabajo bajo el proyecto de investigación FOMENTA 2008/9.

## Referencias

- [1] World Wind Energy Association, "World Wind Energy Report 2008", On-line: [http://www.wwindea.org/home/images/stories/worldwindenergyreport2008\\_s.pdf](http://www.wwindea.org/home/images/stories/worldwindenergyreport2008_s.pdf)
- [2] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, E. Feitosa, "A review on the young history of the wind power short-term prediction", *Renewable and Sustainable Energy Reviews*, Vol. 12, No. 6, pp. 1725- 1744, august 2008.
- [3] A.E. Curtright, M.G. Morgan, D.W. Keith, "Expert Assessments of Future Photovoltaic Technologies", *Environmental Science and Technology*, Vol. 42, no. 24, pp. 9031-9038, 2008.
- [4] A. Yona, T. Senjyu, A.Y. Saber, T. Funabashi, H. Sekine, C.-H. Kim, "Application of neural network to 24- hour-ahead generating power forecasting for PV system", in *Proc. IEEE PES General Meeting 2008*, pp. 6.
- [5] J. Dudhia *et al.* "PSU/NCAR Mesoscale modelling system. Tutorial class notes and user's guide: MM5 modelling system version 3", National Center for Atmospheric Research, Camp Springs, Maryland, 2005.
- [6] G.E.P. Box, G.M. Jenkins, "Time series analysis: forecasting and control. 3<sup>rd</sup> ed.", Prentice-Hall, Englewood Cliffs, NJ, 1994.
- [7] C.G. Atkeson, A.W. Moore, S. Schaal, "Locally Weighted Learning", *Artificial Intelligence Review*, Vol. 11, pp. 11-73, 1997.
- [8] J.S.R. Jang, C.T. Sun and E. Mizutani, "Neuro-Fuzzy and Soft Computing", Prentice Hall, Englewood Cliffs, NJ, 1997.