

# Modelos explícitos en generación fotovoltaica

C. R. Sanchez Reinoso<sup>1,4</sup>, S.H. Gallina<sup>2</sup>,  
M.G. Molina<sup>1</sup>, J.A. Solsona<sup>1</sup>

## Resumen

La optimización de los sistemas fotovoltaicos para la generación de energía eléctrica implica la necesidad de disponer datos reales de las diferentes variables involucradas, como así también la determinación de sus correlaciones.

En el ámbito de la energía solar fotovoltaica resulta de interés poder predecir la energía eléctrica generada por los módulos en función de la radiación solar y de los parámetros climáticos. En este trabajo, se propone un método de correlación basado en técnicas de inteligencia artificial, que permite obtener la energía generada para distintas condiciones climáticas durante un año. Además, se propone un modelo que relaciona la corriente de cortocircuito de una celda solar con la radiación global, pero a diferencia de lo usual, se considera el verdadero comportamiento no lineal de la relación entre las variables.

Los resultados del método propuesto empleando datos reales muestran su validez y utilidad en la predicción de energía generada por módulos fotovoltaicos y en avances tendientes a encontrar métodos de medición de radiación solar alternativos con bajo error.

**Palabras clave:** Módulos fotovoltaicos, celda solar, caracterización, predicción de generación, programación genética

## Abstract

In the optimisation of photovoltaic systems for electricity generation real data of the different variables involved are needed as well as determination of their relationships. In the field of photovoltaic solar energy there is interest to predict the energy generation in terms of solar radiation and climatic parameters.

In this paper, we propose a method based on artificial intelligence techniques for obtaining the generated energy under climatic conditions during a year. In addition, we propose a model that relates short-circuit current with radiation, but unlike usual, is considered the true nonlinear behavior of the relationship between variables.

The results of the proposed method using real data show its validity and usefulness in predicting the generated energy by photovoltaic modules and the search for alternative methods of measuring global radiation at low cost and reasonable error.

**Keywords:** Photovoltaic modules, solar cells, characterization, generation prediction, genetic programming.

## 1. Introducción

Al definir un proyecto de instalación para la generación de energía solar fotovoltaica es fun-

<sup>1</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

<sup>2</sup>Facultad de Tecnología y Ciencias Aplicadas (FTyCA), Universidad Nacional de Catamarca, Maximio Victoria 55, 4700, Catamarca.  
E mail: csanchezreinoso@santafe-conicet.gov.ar

damental conocer, además de la disponibilidad del recurso solar, información de parámetros climáticos, para saber cuál es la energía disponible que podría ser aprovechada por la instalación a lo largo del año o en una época dada. Además, para conseguir dicho propósito es necesario una correcta medición de las variables involucradas, y comprender cuales son las más importantes.

Debido a la necesidad de cuantificar la energía solar disponible en una determinada ubicación, se utilizan frecuentemente modelos matemáticos, algunos de ellos complejos [1][2][3]. Actualmente se están utilizando modelos de redes neuronales, los que son capaces de encontrar correlaciones entre datos diversos y que han mostrado ser útiles en la determinación de parámetros de módulos fotovoltaicos [4][5][6], en la estimación de la radiación incidente [7][8][9][10][11] y simulación de sistemas de generación de energía eléctrica [12][13].

Una de las características de los modelos neuronales es que son del tipo caja negra y por lo tanto, no permiten obtener una función explícita de las variables físicas involucradas.

En este trabajo presentamos los primeros resultados de correlaciones entre la radiación global y la corriente de cortocircuito de una celda solar monocristalina teniendo en cuenta parámetros climáticos, discutiendo su posible uso como solarímetro. En una segunda parte se trabaja en obtener la dependencia de la energía eléctrica generada por los módulos de la radiación solar y sus componentes así como de las variables climáticas. En ambos casos la idea es obtener funciones explícitas que permiten un modelamiento alternativo de los fenómenos bajo estudio.

## 2. Mediciones

Sobre un rack ubicado en el predio del Campo de la Universidad Nacional del Litoral, en el paraje El Pozo de la ciudad de Santa Fe, cuya latitud es 31° 42' S, se montaron 4 módulos SOLARTEC de 42 Wp con los siguientes ángulos de inclinación respecto al plano horizontal: 0, 24, 36 y 58 grados. Los paneles se instalaron a nivel del suelo sobre cobertura vegetal, orientados en dirección Norte, sin recibir en ningún momento sombra de árboles o edificios.

Un sistema de adquisición de datos diseñado en nuestro laboratorio del Instituto de Desarrollo Tecnológico para la Industria Química (INTEC) se empleó para medir cada 5 minutos la corriente de corto circuito  $I_{cc}$ , el voltaje de circuito abierto  $V_{oc}$ , la temperatura de los módulos, y tres curvas I-V de cada uno de ellos a las 10, 14 y 16 horas. Estas mediciones se realizaron sin interrupciones durante todo un año, desde junio de 2010 hasta abril de 2011 inclusive. Simultáneamente, con dos solarímetros Kipp & Zonen CM 6 se midió, en el plano horizontal, la radiación solar global, y la difusa, en este último caso se usó un aro provisto por Kipp & Zonen, montado de forma que proyecte sombra sobre el detector, cuya posición se corrige semanalmente de acuerdo al corrimiento del ángulo del sol respecto al horizonte. Se efectuaron las correcciones de la radiación difusa medida, conforme a lo indicado en el manual del solarímetro para el aro de sombra CM 11/121.

La energía generada en Wh por los módulos se calculó utilizando la siguiente ecuación:

$$E = FF \cdot I_{cc} \cdot V_{oc} \cdot t \quad (1)$$

Donde FF es el Factor de Llenado de los módulos calculado diariamente a partir de las curvas I-V medidas;  $I_{cc}$  es la corriente de cortocircuito;  $V_{oc}$  es la tensión a circuito abierto y  $t$  es el intervalo de tiempo entre mediciones, que en este caso es de 5 minutos.

Se obtuvieron datos de temperatura ambiente y humedad relativa provenientes de la estación meteorológica instalada en el predio por el Centro de Investigaciones Meteorológicas perteneciente a la Facultad de Ingeniería y Ciencias Hídricas de la UNL.

## 3. Modelo evolutivo

### a) Programación genética

Una de las principales variantes de las técnicas evolutivas es la Programación genética (PG). Fue propuesta por Cramer y Koza de manera independiente. Ellos sugirieron una estructura de árbol para representar un programa en un genoma [14].

Los individuos en la PG son programas de computadora estructurados jerárquicamente. Dichos individuos se forman mediante conjuntos de términos y funciones, los cuales actúan como primitivas que sirven de base para la construcción de programas. El conjunto de términos se compone de las variables o constantes que sirven como argumentos de las funciones. Los términos son considerados como hojas en la estructura de árbol. El conjunto de funciones está compuesto por los operadores aritméticos, los operadores binarios o funciones de dominio específico, y en el árbol se les conoce como nodos de tipo función [15][16][17].

### b) Modelo en función de datos climáticos

Disponiendo de un conjunto de parámetros del clima y con mediciones de energía generada de los módulos, se buscaron relaciones que permitieran estimar la energía generada por una instalación fotovoltaica en la ciudad de Santa Fe, empleando datos meteorológicos. Como los datos generalmente no se relacionan en forma lineal y presentan típicamente cierta dispersión, una alternativa aplicable son las redes neuronales. Sin embargo, dicha técnica solo permite obtener un modelo de caja negra. Otro estudio que se realiza es sobre el comportamiento de la corriente de cortocircuito en función de la radiación global y las variables climáticas. En las dos partes de este trabajo se pretende obtener un modelo explícito de las variables medidas por lo que se opta por emplear métodos de computación evolutiva.

Se emplea el siguiente algoritmo:

- Inicializar la población.
- Evaluar los programas en la población existente y asignar un valor de aptitud a cada individuo.
- Hasta que la nueva población no sea com-

pletada:

Seleccionar uno o varios individuos en la población aplicando un proceso de selección.

Ejecutar los operadores genéticos en el o los individuos seleccionados de la población.

Insertar a los nuevos individuos en la nueva población.

- Reemplazar la población existente con la nueva población, hasta cumplir el criterio de terminación.

El método de selección usado está basado en el de selección mediante ranking. Por otro lado, para realizar la mutación, se selecciona un nodo al azar y el subárbol es cambiado por uno nuevo generado aleatoriamente.

El operador de cruce se realiza mediante los siguientes pasos:

- Seleccionar dos individuos como padres.
- Seleccionar aleatoriamente un subárbol o segmento de instrucciones.
- Intercambiar los subárboles o segmentos de código entre los dos padres.

## 4. Resultados

### 4.1. Variables relacionadas con la radiación global

En una primera etapa de los experimentos se plantearon dos objetivos. Uno es encontrar la relación existente entre la corriente de cortocircuito y la radiación global. Esto implica estudiar cuales son las variables relevantes del problema en primer lugar, para luego encontrar una función explícita que determine el valor de la radiación global.

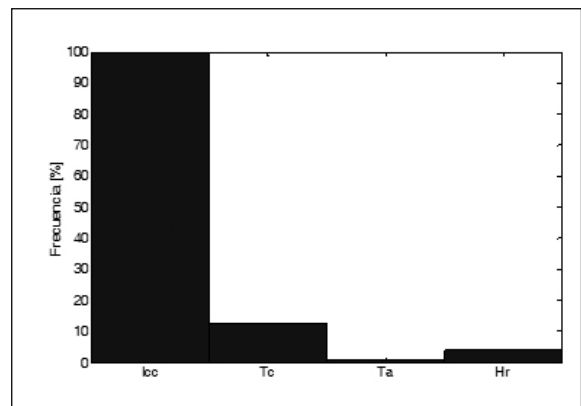
Se realizaron experimentos para entrenar el programa genético con datos de corriente de cortocircuito de la celda solar, temperatura máxima de la celda, temperatura ambiente, humedad relativa, y radiación global.

#### a) Experimentos con 4 meses de datos

La radiación global se planteó como variable predicha y las demás como variables predictoras. El conjunto de datos fue particionado en entrenamiento y prueba. La realización de entrenamientos con los datos permitió obtener un conjunto de soluciones.

Puede verse en la Fig. 1 un análisis estadísti-

co para determinar las variables explicativas más relevantes para el problema planteado. Lo que se muestra es la frecuencia con que las variables explicativas se encuentran en las soluciones encontradas cuya cota máxima de error medio absoluto es de 0.07. Claramente las soluciones que más contribuyen a la disminución del error son la corriente de cortocircuito y la temperatura máxima de la celda, pero en menor medida. En cuanto a la temperatura ambiente y la humedad relativa, en pocas ocasiones mejoran la solución, por lo que no son necesarias si se disponen de las otras variables.



**Figura 1.** Frecuencia de aparición de las variables medidas durante cuatro meses en las mejores soluciones.

La solución seleccionada luego de realizar una serie de experimentos se obtuvo con los siguientes parámetros: Tamaño de población=120, probabilidad de cruce=0.6, probabilidad de mutación=0.031.

Como criterio de error se empleó el error medio absoluto y como criterio de complejidad la cantidad de nodos de la representación en árbol de la solución. Basado en un criterio de doble ordenamiento, se obtuvo una solución que cumplía el requisito de menor error en primer lugar y de menor complejidad en segundo lugar.

La ecuación solución es

$$R_g = 490.11(I_{cc}^2 + 0.59)^{0.18} \quad (2)$$

donde Icc es la corriente de cortocircuito en Amperes; Rg es la radiación global en W/m<sup>2</sup> y se muestran los coeficientes con dos cifras decimales. Se encontraron varias

soluciones que presentaban bajo error, pero se tomó aquella que estaba formada por menos términos.

El coeficiente de correlación entre la variable medida y la calculada por la expresión encontrada es de 0.99, lo cual indica un muy buen ajuste a los datos.

Es importante destacar que también se consideró al inicializar el programa genético, otras variables tales como temperatura ambiente, temperatura de celda y humedad relativa. Durante el proceso evolutivo se encontraron soluciones que prescindían de dichas variables para obtener un buen ajuste de los datos. En el caso de Tc, ésta se encontraba incluida en términos que no aportaban significativamente a la disminución del error, por lo que fueron descartados de la solución. Por lo tanto, se puede afirmar que dichas variables no necesitan ser medidas para obtener una relación de lcc con respecto a la radiación global, para este período de tiempo. También se desprende de este comportamiento que para un buen comportamiento de la función, la solución encontrada es una función no lineal de la corriente de cortocircuito.

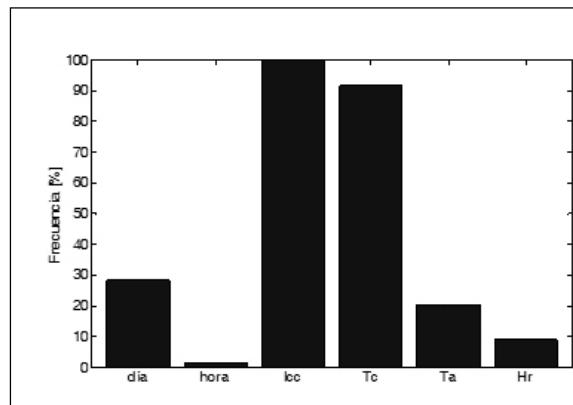
#### b) Experimentos con 1 año de datos

Se realizó el entrenamiento del modelo con datos correspondientes a un año de datos, medidos entre junio de 2010 y abril de 2011. Es claro, que no hay manera que datos de una estación capturen la información del todo el año. Por lo tanto, si bien en la sección precedente se obtuvo una ecuación para cuatro meses del año, en la presente se pretende obtener una expresión que capture el comportamiento a lo largo de un año completo, cubriendo de este modo todas las variaciones anuales que pueda presentar la composición espectral de la radiación solar, considerando que el objetivo es calibrar una celda para ser usada como solarímetro.

Se procedió a realizar un análisis similar al empleado en la subsección anterior. En este caso la participación de las diferentes variables tiene cambios importantes respecto al caso de un grupo de datos de cuatro meses. Esto indica que la solución

debe obtenerse con al menos doce meses de datos.

De acuerdo a la Fig. 2 todas las soluciones consideradas en este análisis incluyen a lcc y un 90% a Tc. Las ecuaciones obtenidas que incluyen a Ta tienen una menor participación de Tc o carecen de ella. La humedad relativa participa en una baja proporción de ecuaciones. En general las expresiones solución que contienen el día son mucho más complejas que las que carecen de ella.



**Figura 2.** Frecuencia de aparición de las variables medidas durante doce meses en las mejores soluciones.

Los resultados muestran que las variables más importantes son lcc y Tc, siendo Tc de valioso aporte en la solución. Es destacable notar que la importancia de Tc es mucho mayor al aumentar la cantidad de datos que uno pretenda modelar. Dado que lo se pretende es una función que represente la relación de las variables para todo el año, debe medirse la temperatura de celda.

Las ecuaciones que incluyen a las demás variables pueden reducir levemente el error, lo que no siempre ocurre. Esto es debido a que el modelado emplea diferentes bloques funcionales no lineales y es beneficioso para la evolución tratar con la menor cantidad de variables posibles. Esto además implica que el modelo evolutivo propuesto realiza naturalmente una selección de variables en el proceso de búsqueda.

La solución seleccionada fue la que presentaba términos formados por las funciones más simples encontradas por el algoritmo. Los parámetros: Tamaño de población=100, probabilidad de cruce=0.4,

probabilidad de mutación=0.0013, permitieron obtener la solución

$$R_g = 1.59I_{cc}T_c + 390I_{cc} - 12.3I_{cc}^3 - 4.17 \quad (3)$$

donde  $I_{cc}$  es la corriente de cortocircuito en A;  $T_c$  es la temperatura de celda en °C;  $R_g$  es la radiación global en W/m<sup>2</sup>. Los coeficientes mostrados están redondeados a dos cifras decimales.

Como criterio de error se empleó el error medio absoluto y como criterio de complejidad la cantidad de nodos de la representación en árbol de la solución. A continuación se tomaron para su comparación la expresión que sea función de todas las variables medidas y presente el menor error obtenido, y otra ecuación que solo considere las variables más relevantes del problema. La solución empleando corriente de cortocircuito y temperatura de celda presenta un R<sup>2</sup> de 0.99 mientras que la expresión que incluye todas las variables medidas tiene un R<sup>2</sup> 0.99. Estos últimos resultados junto a los anteriores muestran que es posible encontrar una expresión de la radiación en función de solo dos variables con fiabilidad para todo un año. Además, la importancia de esto radica en que puede emplearse un dispositivo formado por una celda solar con una termocupla incorporada que constituya un solarímetro de bajo costo que entregue directamente la radiación global, sin realizar ningún cálculo posterior y sin emplear constantes de calibración estacionales y/o que ignoren las no linealidades propias del fenómeno en estudio.

Por lo tanto, la expresión encontrada resulta sumamente útil para la determinación de la radiación global en función de la corriente de cortocircuito y la temperatura de la celda.

#### 4.2. Energía generada en función de variables climáticas

Considerando la obtención de correlaciones de la energía generada por módulos fotovoltaicos con variables climáticas, se intentó encontrar relaciones explícitas basadas en datos reales. La base de datos la formaban la energía generada por día por los módulos, la radiación global en plano

horizontal y sus componentes directa y difusa, la temperatura de módulo, de medio ambiente y la humedad relativa. Mediante computación evolutiva se realizaron experimentos con diferentes parametrizaciones, con datos particionados en entrenamiento y prueba.

##### a) Experimentos con 4 meses de datos

Se tuvo en cuenta en la selección del mejor modelo el compromiso error de prueba-complejidad. Los parámetros empleados en la búsqueda de dicha solución fueron un tamaño de población=300, probabilidad de cruce=0.5, probabilidad de mutación=0.3.

La ecuación obtenida es

$$E_g = T_{mp} + \frac{8.34e6}{12.7R_g + \frac{1.42e7}{R_g} + d^2} - T_{mp}R_{dif}^{0.185} \quad (4)$$

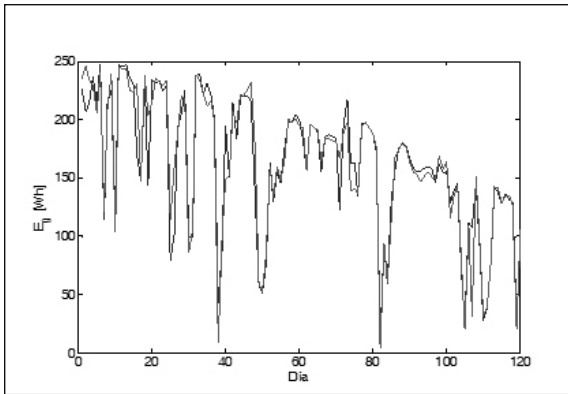
donde  $d$  es el día del año numerado desde enero,  $T_{mp}$  es la temperatura máxima del panel en °C;  $R_{dif}$  es la radiación difusa en W/m<sup>2</sup>;  $R_g$  es la radiación global en W/m<sup>2</sup>;  $E_g$  es la energía generada en Wh. La solución final fue seleccionada por tener menor cantidad de términos que otras de error muy similar. Cabe destacar que dentro de las mejores soluciones encontradas por el algoritmo, en su mayoría incluyen las variables explicativas de la ecuación (2).

Si se intenta relacionar la variable predicha medida con la calculada por el modelo, se consigue un coeficiente R<sup>2</sup> de 0.96, indicando un buen desempeño de la función encontrada. La energía generada para datos de entrenamiento y validación, como así también la solución seleccionada se muestran la Fig. 3

##### b) Experimentos con 1 año de datos

Se realizó el entrenamiento del modelo con datos correspondientes a un año de medición, medidos entre junio de 2010 y abril de 2011.

Si bien ya se obtuvo una ecuación para cuatro meses del año, en esta subsección se pretende obtener una expresión que capture el comportamiento a lo largo de un año completo y analizar dichos resultados. Cuando se emplearon las variables climáticas medidas como variables indepen-



**Figura 3.** Energía generada en función de variables climáticas para los distintos días del año. La cantidad de días corresponde a datos tomados desde enero hasta abril.

dientes de la función energía generada, se encontraron soluciones cuyo mejor desempeño tiene un R2 de 0.75. Estos resultados no son tan alentadores como los obtenidos para un conjunto de datos correspondientes a los primeros cuatro meses del 2011.

Se plantea entonces como variante incluir el ángulo de instalación de los módulos fotovoltaicos. Esto permitió encontrar soluciones mucho más promisorias que sean válidas para todo el año. Con la mejor solución encontrada, si se intenta relacionar la variable predicha medida con la calculada por el modelo, se consigue un coeficiente R2 de 0.9 y coeficiente de correlación de 0.95, indicando un buen desempeño de la función encontrada.

Los parámetros empleados en la búsqueda de dicha solución fueron un tamaño de población=250, probabilidad de cruce=0.4, probabilidad de mutación=0.2 y la ecuación obtenida es

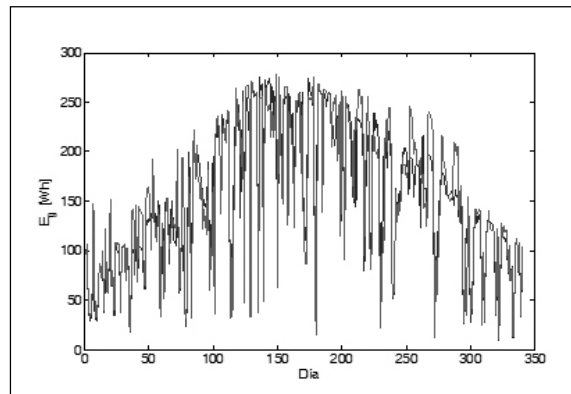
$$E_g = 0.476R_g^{0.012} - 1.84 \cdot \alpha \cdot \sin(-0.00331R_g) - T_a(0.00104R_g \alpha)^{0.443} - R_g^{0.302} \sin(\alpha) \quad (5)$$

donde  $R_g$  es la radiación global en  $W/m^2$ ,  $R_{dif}$  es la radiación difusa en  $W/m^2$ ,  $\alpha$  es el ángulo de montaje de los módulos,  $T_a$  es la temperatura ambiente en  $^{\circ}C$ ;  $E_g$  es la energía generada en Wh, y los coeficientes de las potencias fueron redondeados a tres cifras decimales.

Es de notar que las variables presentes en la solución encontrada son casi todas las

medidas, observándose que no participan en esta solución la temperatura máxima de los módulos, siendo activa la temperatura máxima del medio ambiente, y que la Radiación difusa es mas directa en la representación de la transparencia de la atmosfera que la humedad relativa, que también no participa de la solución.

Un gráfico de la energía generada en función de variables climáticas y el ángulo de montaje (Fig. 4) permite observar un adecuado comportamiento del modelo. En dicho gráfico se distinguen los resultados para los datos de entrenamiento, de validación, y de la solución final. El modelo obtenido permite realizar una buena predicción de la energía total generada por un módulo fotovoltaico en el transcurso de un año.



**Figura 4.** Energía generada en función de variables climáticas para los distintos días del año. La cantidad de días corresponde a datos tomados desde junio de 2010 hasta mayo 2011.

## 5. Conclusiones

Los resultados encontrados utilizando programación genética para encontrar correlaciones entre parámetros de módulos fotovoltaicos y la radiación solar conjuntamente con datos climáticos se muestran muy promisorios. La correlación de la corriente de corto circuito de una celda solar y su temperatura con la radiación global es excelente y reproduce con mínimo error los valores de la radiación global en todo un año. De igual modo es posible encontrar la energía que generaría un módulo fotovoltaico en un año de funcionamiento, esta solución muestra dos aspectos que probablemente requieran de una mayor investigación y es el hecho que la energía no tenga una dependencia clara de la temperatura del módulo y que sea una función explícita

de la radiación difusa. La metodología propuesta realiza una selección automática del modelo mediante un doble criterio dado por el error y la complejidad. En trabajos futuros se incorporará una última etapa de selección manual de la solución basada en criterios físicos.

## Referencias

- [1] W. F. Philips (1984), "Harmonic analysis of climatic data," *Solar Energy*, vol. 32, pp. 319-328.
- [2] A. C. Herrero (1993), "Harmonic analysis of monthly solar radiation data in Spain," *Ambient Energy*, vol. 14, pp. 35-40.
- [3] A. Dorvlo y B. Ampratwum (1999), "Modeling of weather data for Oman," *Renewable Energy*, vol. 17, pp. 421-428.
- [4] L. S. De Bernardez, R. H. Buitrago y N. O. Garcia (2008), "Aplicación de redes neurales para el cálculo de la energía generada por paneles fotovoltaicos a partir de datos climáticos," in *Proc. 2008 XIV Congreso Ibérico y IX Congreso Iberoamericano de Energía Solar*, pp. 973-978.
- [5] E. Karatepe, M. Boztepe y M. Colak, "Estimation of equivalent circuit parameters of PV module using neural network," in *IJCI Proc. 2003 International XII Turkish Symposium on Artificial Intelligence and Neural Networks*, pp. 76-78.
- [6] G. Tamizh, L. Ji, Y. Tang, L. Petacci y C. Osterwald (2003), "Photovoltaic module thermal/wind performance," in *Proc. NCPV and Solar Program Review Meeting, NREL 520-33586*.
- [7] K. H. Elminir, J. Alamjam, U. R. Ali y V. Benda, "Estimation of Solar Radiations Incident on a Photovoltaic Solar Module using Neural networks," in *Proc. 2001 XXVI ASR Instruments and Control*.
- [8] F. S. Tymvios et al. (2005), "Comparative study of Angström's and artificial neural networks methodologies in estimating global solar radiation," *Solar Energy*, vol. 78, pp. 752-762.
- [9] G. López, F. J. Batles y J. Tovar (2005), "Selection of input parameters to model direct solar irradiance by using artificial neural networks," *Energy*, vol. 30, pp. 1675-1684.
- [10] A. I. Serrano, P. S. Lucio, A. M. Silva y S. M. Leite (2006), "Meteorological Time Series Reconstruction via Artificial Neural Network," *Geophysical Research*, vol. 8, pp. 879.
- [11] A. Mellit, M. Benghanem, A. Hadj y A. Guessoum (2005), "A simplified model for generating sequences of global solar radiation data for isolated sites: Using artificial neural network and a library of Markov transition matrices approach," *Solar Energy*, vol. 79, pp. 469-482.
- [12] C. R. Sanchez Reinoso, D. H. Milone y R. H. Buitrago (2011), "Efficiency study of different photovoltaic plant efficiency connection schemes under dynamic shading," *International Journal of Hydrogen Energy*, vol. 35, pp. 5838-5843.
- [13] C. R. Sanchez Reinoso, D. H. Milone y R. H. Buitrago, "Desarrollo de un modelo para estudio de centrales fotovoltaicas bajo diferentes configuraciones," in *Proc. 2009 Eighth Latin-American Congress on Electricity Generation and Transmission*, pp. 1-8.
- [14] Koza J. R (1992), *Genetic Programming. On the Programming of Computers by Means of Natural Selection*. Cambridge, USA: MIT Press, 1992, p. 819
- [15] W. Banzhaf, P. Nordin, R. Keller, F. Francone, *Genetic programming an introduction. In On the Automatic Evolution of Computer Programs and Its Applications*. San Francisco, USA: Morgan Kaufmann Publishers, 1998, p. 479
- [16] J. R. Koza, *Genetic Programming II. Automatic Discovery of Reusable Programs*. Cambridge, USA: MIT Press, 1998, p. 768
- [17] J. R. Koza, *Genetic Programming III. Darwinian Invention and Problem Solving*. San Francisco, USA: Morgan Kaufmann Publishers, 1999, p. 1154

