

Pronóstico de generación ideal a escala mensual usando ANFIS

Ideal generation forecast in monthly scale using ANFIS

Julian Moreno Cadavid¹, MSc., Yony Fernando Ceballos²

1. Universidad Nacional de Colombia – Sede Medellín. 2. Universidad de Antioquia
jmoreno1@unalmed.edu.co; fceball@udea.edu.co

Recibido para revisión: 21 de Abril de 2008, Aceptado: 28 de Noviembre de 2008, Versión final: 15 de Diciembre de 2008

Resumen—En este artículo se propone el uso de un modelo ANFIS para el pronóstico de la generación ideal mensual de un agente con una planta hidráulica dentro del mercado eléctrico colombiano. El modelo propuesto considera factores de diversa índole como el embalse de la planta, el caudal medio de los ríos que lo alimentan, el clima esperado representado por la anomalía de la SST en la zona niño 3.4 y la magnitud de los contratos de venta que tiene el agente. El desempeño de dicho modelo se mide con los datos reales del período 2002-2006 y es comparado contra un modelo de regresión lineal múltiple. Los resultados obtenidos evidencian una disminución considerable del error porcentual medio, lo que da luces sobre su validez y posible aplicación para otros agentes.

Palabras Clave—Aplicaciones de inteligencia artificial, ANFIS, Regresión lineal múltiple, Mercado eléctrico, Generación ideal.

Abstract—In this paper the use of an ANFIS model is proposed to forecast the monthly ideal generation of an agent with a hydraulic plant within the Colombian electricity market. The proposed model considers factors of different kind as the plant's dam, the mean stream flow of the rivers which feed it, the expected weather represented by the SST anomaly in Niño 3.4 zone, and the magnitude of sale contracts that the agent has. The fitness of such model is measured with real data from period 2002-2006 and it is compared against a multiple linear regression model. The obtained results show a considerable decrease of the mean percentage error, which is an evidence of its validity and possible application in other agents.

Keywords—Artificial intelligence applications, ANFIS, Multiple linear regression, Power market, Ideal generation.

I. INTRODUCCIÓN

Dentro del contexto del mercado eléctrico Colombiano, así como de otros con un esquema similar, la generación ideal de una planta se refiere a la generación que se asignaría dentro del despacho a dicha planta para un período determinado considerando una condición “ideal” de la red. En este caso “ideal” se refiere a que no hay restricciones en los límites de importación entre las áreas eléctricas, ni existen problemas de estabilidad que requieran de generaciones obligatorias. En este sentido, la generación ideal (más que la real) denota la intención de generar de un agente dadas sus condiciones propias y las del mercado en un momento dado.

El pronóstico de dicha generación es de gran interés tanto para los agentes del mercado como para las entidades encargadas de la planeación y control. Para el caso de los primeros, una estimación de su generación y la de sus competidores les puede permitir definir estrategias para maximizar sus ganancias, gestionar sus recursos (principalmente los embalses) y planificar sus operaciones de mantenimiento. Para el caso de los segundos tal estimación puede llegar a ser útil en algunas de sus funciones como determinar planes para garantizar la estabilidad del sistema, aprobar planes de mantenimiento, entre otras.

En el trabajo presentado en este artículo se explora la posibilidad de emplear un modelo para realizar este pronóstico a escala mensual considerando diversos factores económicos e hidroclimáticos, siendo los de esta última categoría tanto del agente como del sector en general. El modelo estudiado corresponde a un ANFIS (sistema de inferencia neuro difuso adaptativo), proveniente de la Inteligencia Artificial del cual

algunas aplicaciones en series de tiempo, y más específicamente en otros problemas del sector eléctrico se pueden encontrar en [1-5]

El resto de este artículo está organizado de la siguiente manera: en la sección siguiente se hace una breve descripción de las variables consideradas y su relación con la generación ideal; en la sección III se presenta el modelo ANFIS empleado, describiendo tanto su arquitectura como su calibración; en la sección IV se muestran los resultados obtenidos y se contrastan contra los de un modelo ampliamente utilizado como es la regresión lineal múltiple. Finalmente, en la sección V se presentan las conclusiones fruto de este trabajo y se esboza el trabajo futuro que debería realizarse.

II. ANÁLISIS DE VARIABLES

Para enfrentarse a la tarea de pronosticar la generación ideal de un agente dentro de un mercado eléctrico con las características como el colombiano es posible recurrir al análisis de los factores que bajo una perspectiva física y económica pueden afectar esta variable. Dentro de la primera categoría se encuentran aquellos factores que se relacionan directamente con la capacidad de generación tanto presente como futura de la planta correspondiente. Entre estos se encuentran el embalse ofertable, el caudal medio y el clima esperado dentro de un horizonte de mediano plazo. Dentro de la segunda categoría se encuentran los factores que inducen a los agentes a tomar estrategias bien sea de maximización de utilidades o de minimización de riesgo. Entre dichos factores se encuentran los contratos firmados y el precio esperado de la bolsa, aunque este último por su complejidad propia no fue considerado en este trabajo.

Todos estos factores, más otros fuera del alcance de esta investigación, se juntan para determinar de alguna manera cual debe ser la generación ideal futura de un agente de acuerdo a sus preferencias. Así por ejemplo, para un agente adverso al riesgo la generación ideal del mes siguiente será baja si su nivel de embalse es moderado pero el caudal esperado y la expectativa de condiciones frías es alta. Esto siempre y cuando la magnitud de los contratos de venta firmados para ese período no sea alta, pues esto significaría una alta exposición a bolsa que no está dispuesto a enfrentar. Como puede verse en este ejemplo sencillo las decisiones a tomar no son sencillas e involucran muchas opciones, con el agravante adicional que difícilmente podrán ser cuantificadas de manera exacta. Estas características hacen suponer que un modelo que contemple el manejo de variables y reglas difusas puede ser válido para este problema. La descripción del modelo que cumple con estas particularidades se presenta en la siguiente sección, por lo pronto se hace a continuación una breve explicación de las variables consideradas.

A. Embalse en $t-1$

Se refiere al embalse ofertable correspondiente a la planta analizada durante el período anterior (en este caso durante el mes anterior) medido en porcentaje respecto a su capacidad total.

B. Caudal medio

El caudal medio dentro del contexto de este trabajo se define como la sumatoria del caudal esperado medido en m^3/seg de los ríos que alimentan a la planta correspondiente. Esto independientemente de que el embalse respectivo tenga o no capacidad de regulación. El cálculo de este factor se realiza con un promedio móvil de los últimos n años, pero para efectos prácticos se puede utilizar simplemente el que reporta la fuente consultada.

C. Contratos

Este factor se refiere a la magnitud de los contratos de venta (referidos en MWh-mes) que el agente tiene firmados bien sea con comercializadores, con otros generadores, o en el caso de la entrada del MOR con el representante de la demanda.

D. Anomalía 3 meses

El valor de este factor se refiere a la anomalía que espera la NOAA dentro de 3 meses para la zona niño 3.4. Esta anomalía se mide como la diferencia entre la temperatura esperada y la temperatura media de los últimos años para el mes correspondiente. Su importancia, más allá de la exactitud de los modelos de pronóstico que emplea esta entidad, es la expectativa que crea en los agentes siendo la NOAA una de las fuentes oficiales más consultadas.

E. Anomalía 6 meses

Este factor se interpreta de la misma manera que el anterior, con la diferencia que se trata del valor esperado de la anomalía 6 meses hacia adelante.

Para conseguir las series de tiempo de las variables analizadas se acudió a la herramienta Neón [6] para el caso del embalse ofertable, el caudal medio y los contratos, mientras que para las climáticas se acudió a la NOAA [7]. La relación de estas cinco variables con la variable de interés (consultada también desde Neón) puede observarse de manera cuantitativa en las Tablas 1 a 3, y de manera gráfica en la Figura 1. El agente escogido para este estudio fue Chivor, dado que cuenta con una sola planta de generación.

Las letras A-E que aparecen en la Tabla 1 corresponden a la nomenclatura utilizada en la descripción de las variables, mientras que la F corresponde a la generación ideal (medida al igual que los contratos en MWh-mes).

Tabla 1. Coeficiente de correlación

A	A	B	C	D	E	F
A	1,00	- 0,06	0,30	0,65	0,58	0,16
B	- 0,06	1,00	0,54	0,06	- 0,21	0,67
C	0,30	0,54	1,00	0,14	0,21	0,58
D	0,65	0,06	0,14	1,00	0,69	0,25
E	0,58	- 0,21	0,21	0,69	1,00	- 0,01
F	0,16	0,67	0,58	0,25	- 0,01	1,00

Tabla 2. Valor p de las variables analizadas versus generacion ideal

Multiple Regression Analysis				
Dependent variable: GeneracionIdeal				
Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	120567,0	59855,8	2,01429	0,0490
Anomalia3meses	-81455,4	52761,9	-1,54383	0,1285
Anomalia6meses	116710,0	53041,6	2,20035	0,0321
CaudalMedio	0,216302	0,0684922	3,15806	0,0026
Contratos	1,03055	0,361272	2,85254	0,0061
EmbalseTmesos1	-8549,54	68605,3	-0,124619	0,9013

Tabla 3. Estadísticos de la regresión lineal múltiple
Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	7,62952E11	5	1,5259E11	14,16	0,0000
Residual	5,81752E11	54	1,07732E10		
Total (Corr.)	1,3447E12	59			

R-squared = 56,7375 percent
 R-squared (adjusted for d.f.) = 52,7318 percent
 Standard Error of Est. = 103794,0
 Mean absolute error = 78542,5
 Durbin-Watson statistic = 1,19573

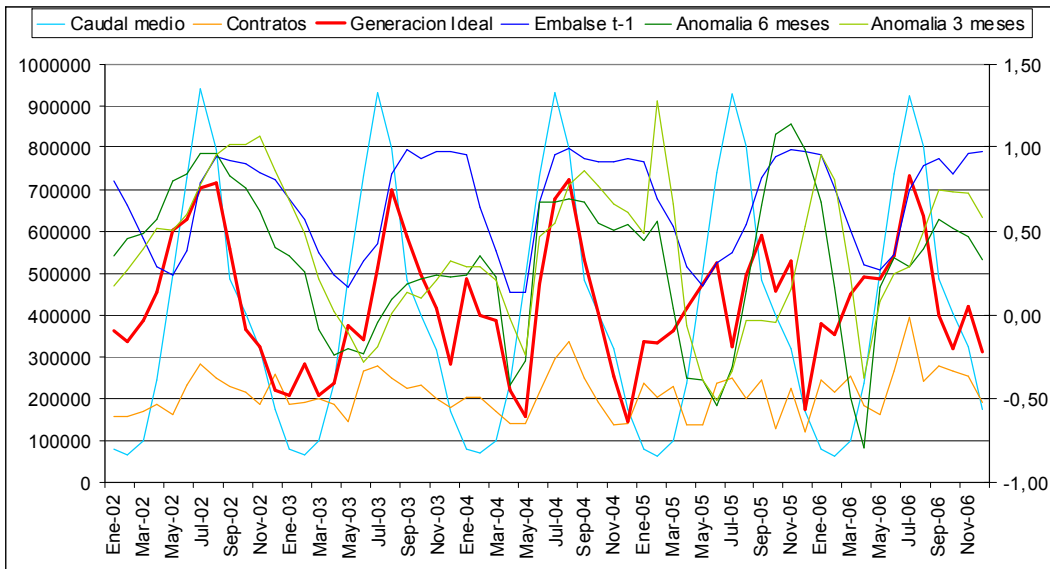


Figura 1. Evolución de las variables consideradas durante el periodo 2002-2006

De la Figura 1 cabe señalar que el eje de las ordenadas y del lado izquierdo corresponde con la escala del caudal medio, los contratos y la generación ideal; mientras que el de la derecha corresponde a la escala del resto de variables.

III. MODELO ANFIS

El modelo de inferencia neurodifuso adaptativo o ANFIS (por sus siglas en inglés) propuesto por Jang & Sun [8] busca combinar las características de la lógica difusa y las redes neuronales. Este modelo como se puede apreciar en la Figura 2 está compuesto de cinco capas donde cada una corresponde

a una parte diferente del sistema de inferencia. Si bien dentro de este artículo no se ahonda en sus fundamentos matemáticos puesto que el lector puede encontrar literatura asociada (alguna de la cual se encuentra entre las referencias de este artículo), cabe mencionar a manera de resumen que un ANFIS es un sistema de reglas SI <antecedentes> ENTONCES <consecuentes> donde los antecedentes son conjuntos difusos y los consecuentes son funciones lineales. La forma de los conjuntos y los parámetros de las funciones son calibradas por el modelo minimizando una medida del error. En otras palabras un ANFIS hace uso de una regla de aprendizaje para optimizar los parámetros de un sistema Takagi-Sugeno de primer orden [10,11].

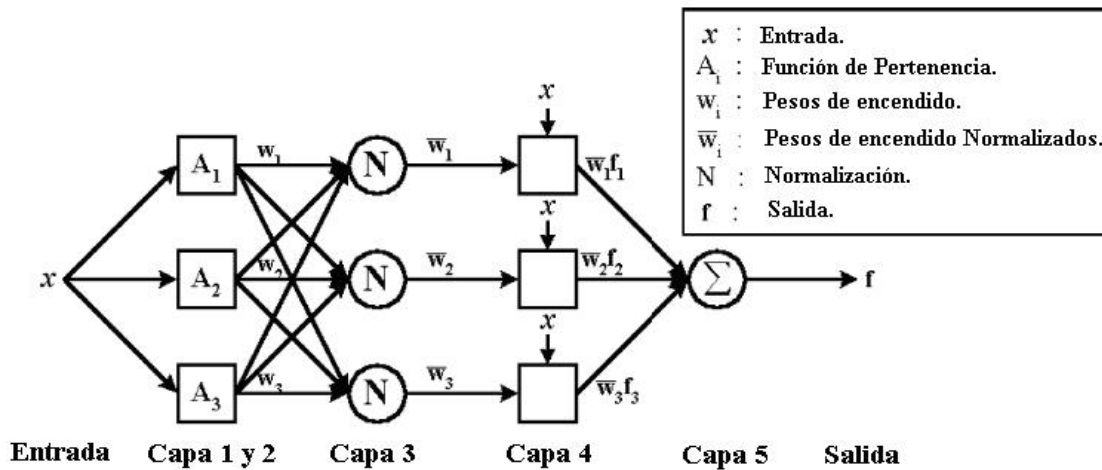


Figura 2. Arquitectura general del modelo ANFIS

Para generar los conjuntos difusos iniciales se puede elegir una forma arbitraria, una distribución uniforme dentro del universo del discurso, o utilizar algún método de agrupamiento. En cualquiera de los tres casos se deben emplear alguna de las formas estándar de representación: triangular, trapezoidal, campana de Gauss, etc. Para este trabajo se utilizó como método

de agrupamiento el clustering sustractivo [9], empleando campanas de Gauss. De los resultados de la calibración se pueden observar los conjuntos de pertenencia obtenidos para dos de las variables de entrada en las Figuras 3 y 4, mientras que la evolución de la medida del error se puede observar en la Figura 5.

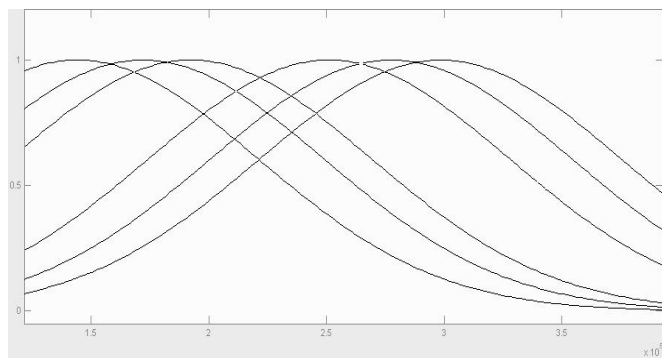


Figura 3. Funciones de pertenencia para la variable contratos

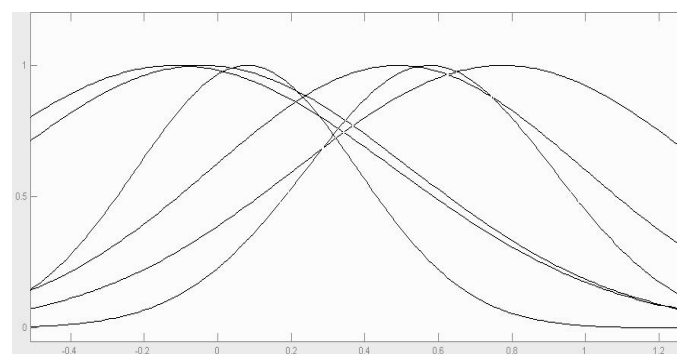


Figura 4. Funciones de pertenencia para la variable anomalía 3 meses

En la Figura 5 se puede observar que el algoritmo de calibración (en este caso back propagation) empieza a volverse asintótico respecto al error a cerca de la iteración 5000, y esto precisamente se usó como criterio de parada para evitar sobreentrenamientos.

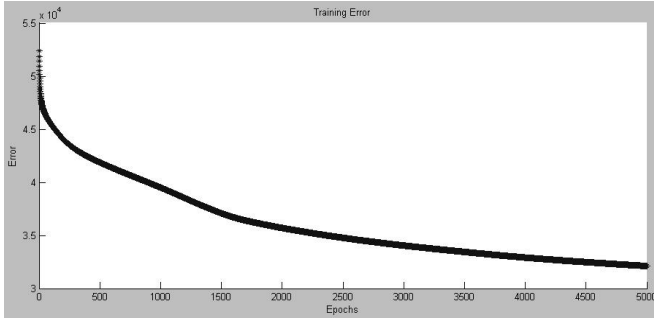


Figura 5. Evolución del error para 5000 iteraciones

IV. RESULTADOS

Para medir el desempeño de un modelo es importante tener un punto de comparación y en este caso puede tratarse de una regresión lineal múltiple que contemple las mismas variables independientes. Ya en la sección II, al hacer el análisis de las variables consideradas se mencionaron los estadísticos de esta regresión. En la Tabla III se muestra por ejemplo que el valor P del modelo de regresión (<0.0001) indica que este es estadísticamente significativo a la hora de explicar la variable de salida. Así mismo un R² de 56.74 y un R² ajustado de 52.73 muestran la bondad de dicho modelo. En la Tabla II se observa que el valor P de las variables Anomalía 3 meses y Embalse t-1 indica que estas variables no son relevantes en la regresión, por lo que ensayó una versión simplificada que no considera

estas variables. En este caso se obtuvo un R² de 54.8 y un R² ajustado de 52.38, conservando un valor P del modelo de regresión menor a 0.0001.

Otro estadístico de interés que no se mencionó en la sección II es el error porcentual medio que se calcula como se presenta a continuación.

$$MAPE = \frac{\sum_{i=1}^P \left| \frac{s_i - t_i}{t_i} \right|}{P} \tag{1}$$

siendo P el número de muestras

En la regresión el valor de este estadístico fue de 23.6% cuando se consideraron todas las variables y de 24.9% cuando se eliminan las estadísticamente irrelevantes. En ambos casos estos valores para efectos de pronóstico pueden ser demasiado altos. Otra forma de visualizar el rendimiento de los modelos probados es por medio de la gráfica de sus salidas versus la variable dependiente (en este caso generación ideal de Chivor). En la Figura 6 se ilustra esta comparación para los dos modelos de regresión analizados y puede observarse que la diferencia entre ambos no es muy marcada, pero si lo es contra la generación ideal. Haciendo un análisis meramente cualitativo (pues los estadísticos expuestos anteriormente ya dan cuenta de su rendimiento cuantitativo) puede decirse que los modelos de regresión empleados intentan seguir la tendencia y estacionalidad de la serie pero se quedan cortos a la hora de seguir puntos extremos sean picos o valles (mírese agosto de 2003 y diciembre de 2004 por ejemplo) y detectar cambios abruptos (mírese enero de 2004 y julio de 2005 por ejemplo).

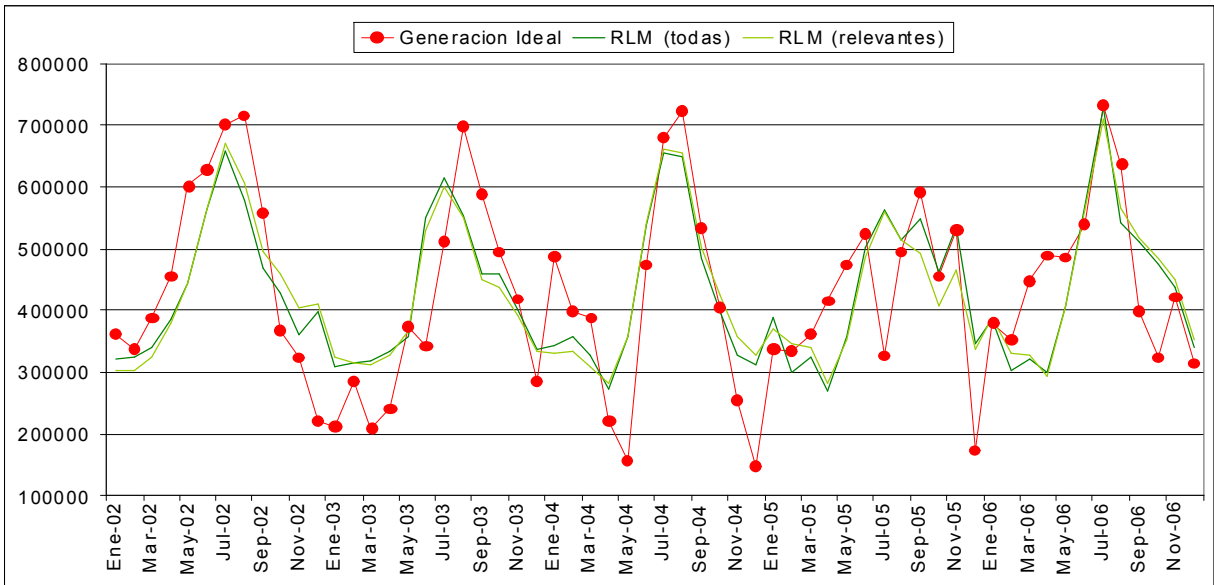


Figura 6. Salida del modelo de regresión lineal múltiple (RLM) versus Dato real

Por su parte en la Figura 7 se muestra la misma comparación gráfica pero esta vez respecto al modelo ANFIS. En este caso se puede observar que el modelo sigue de manera apropiada la evolución de la serie y salvo en 8 ocasiones (recordando que el número total de muestras es 60) la desviación del modelo no sobrepasa el 20% y de estas excepciones solamente en una (mayo de 2004) se sobrepasa el 50%. Este último caso debido

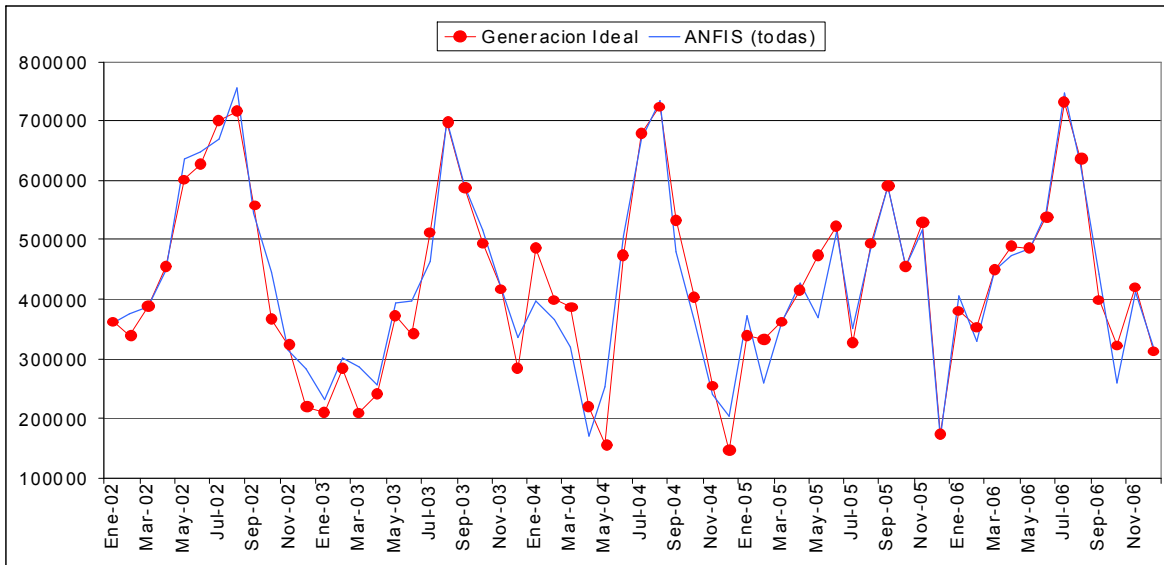


Figura 7. Salida del modelo ANFIS versus Dato real

V. CONCLUSIONES Y TRABAJO FUTURO

Siendo el pronóstico de la generación ideal dentro de un mercado eléctrico una tarea de interés no solo para los agentes involucrados si no para las entidades de planeación, en este artículo se presenta como propuesta un modelo ANFIS cuyas variables de entrada corresponden a factores hidrológicos, climáticos y económicos. Luego de calibrar el modelo con los datos mensuales reales de 6 años de un agente específico, el modelo propuesto mostró un error porcentual medio inferior al 9%, superando claramente a modelos de regresión lineal múltiple que consideran las mismas variables. Este resultado da cuenta de la no linealidad de las relaciones entre las variables de entrada y la de salida, dándole validez a la hipótesis de la pertinencia de modelos que, como el ANFIS, buscan representar el conocimiento de manera similar a como lo hacen los expertos humanos.

A pesar del rendimiento obtenido es importante aclarar que todavía quedan algunos frentes que se pueden abordar como trabajo futuro, esto con el fin de brindarle mayor robustez y mejorar aún más su rendimiento. Entre estos frentes se encuentran:

- Probar el modelo con las series históricas de otros agentes generadores, realizando la calibración respectiva.
- Determinar el horizonte u horizontes óptimos que deberían

principalmente a la magnitud de los valores para esta fecha particular.

Es necesario mencionar también que con el modelo ANFIS el error porcentual medio obtenido fue de 8.93% que corresponde a casi una tercera parte del error obtenido por la regresión lineal múltiple que contempla las mismas variables de entrada.

considerarse para la variable de entrada anomalía de la SST para la zona Niño 3.4

- Como alternativa al anterior, podría integrarse la expectativa climática no como un valor esperado para un horizonte definido, si no como un promedio de un rango establecido. Un ejemplo de esto sería utilizar una variable que fuera el promedio de los pronósticos de anomalía de los próximos n meses.
- Incorporar el elemento autoregresivo de la serie considerando como variable explicativa adicional la misma generación ideal para uno o varios rezagos determinados.

REFERENCIAS

- [1] Rodríguez, C. P. and Anders, G. J., 2004. Energy Price Forecasting in the Ontario Competitive Power System Market. IEEE Transactions on Power Systems, Vol. 19, No. 1, pp. 366 – 374.
- [2] Hong, Y.Y. and Hsiao, C.Y., 2002. Locational Marginal Price Forecasting in Deregulated Electric Markets Using Artificial Intelligence. IEE Proc.-Gener. Transm. Distrib., Vol. 149, No. 5.
- [3] Zapata, E., 2006. Modelado de Series Temporales Usando Sistemas Neuro Difusos ANFIS. Tesis de Maestría, Universidad nacional de Colombia – Sede Medellín.
- [4] Mora, J. J., Pérez, L. P. y Pérez, S. M., 2006. Utilización de redes ANFIS y señales de corriente para localización de la zona de falla en sistemas de distribución de energía eléctrica. Ingeniería e Investigación v.26 n.3 Bogotá.
- [5] Calderón, J. A., 2007. Algoritmo para la Detección y Clasificación de fallas en líneas de transmisión usando ANFIS. Revista Avances en Sistemas e

Informática, Vol 4 No 1, pp. 101 – 108.

- [6] Neón. Servicio de información sobre el Mercado de Energía Mayorista Colombiano. Página Web: <http://sv04.xm.com.co/neonweb/>
- [7] NOAA - National Oceanic and Atmospheric Administration. Página Web: <http://www.noaa.gov/>
- [8] Jang, J.-S.R. and Sun, C.-T., 1997. Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Prentice-Hall, New Jersey.
- [9] Taheri, J. and A. Zomaya, Y., 2006. Handbook of Nature-Inspired and Innovative Computing Integrating Classical Models with Emerging Technologies. Ed. Springer.
- [10] Hines, J. W., 1997. MATLAB supplement to fuzzy and neural approaches in engineering. New York, Wiley.
- [11] Jang, J.-S.R. and Sun, C.-T., 1999. Neuro-fuzzy modeling and control. Proceedings of IEEE 83(3), pp. 378–406.

Julian Moreno Cadavid. Profesor auxiliar, Universidad Nacional de Colombia – Sede Medellín Ingeniero de sistemas e informática (2004),. Magister en Ingeniería de sistemas (2007). Integrante de GIDIA: Grupo de Investigación y Desarrollo en Inteligencia Artificial, Categoría A de Colciencias. El área de énfasis de su investigación es la Inteligencia Artificial aplicada a diversos problemas de la industria. Más específicamente ha trabajado en sistemas híbridos que integran Sistemas Expertos, Sistemas Neuro-Difusos, Sistemas Multi-Agente y otras herramientas para el apoyo a la toma de decisiones en áreas como los Mercados de Energía, la Detección de Fallas en Líneas de Transmisión, así como en diversas aplicaciones de simulación y pronóstico.

Yony Fernando Ceballos. Profesor de tiempo completo de la Universidad de Antioquia. Ingeniero de sistemas e informática de la universidad Nacional de Colombia. M.Sc en Ingeniería de sistemas de la Universidad Nacional de Colombia. Investigador en proyectos relacionados con el área de Dinámica de Sistemas y pertenece al grupo Ingeniería y Software de la Universidad de Antioquia (Categoría B - Colciencias) Ha trabajado en la actualidad en la Universidad Nacional de Colombia en diferentes proyectos relacionados con el área de Dinámica de sistemas, así como diversas aplicaciones en el sector eléctrico.

Universidad Nacional de Colombia Sede Medellín

Facultad de Minas



Reseña Histórica

La Escuela Nacional de Minas fue fundada el 11 de abril de 1887, bajo la dirección del general Pedro Nel Ospina como rector y como Vice-rector Luís Tisnés, aunque el general Pedro Nel Ospina no se posesiono, elaboro con ayuda de su hermano Tulio los estatutos y reglamentos de la escuela, los cuales fueron una adaptación de los estatutos y reglamentos de la Escuela de Minas de California (Berkeley) los cuales fueron cambiando de acuerdo a las necesidades de cada década, en ellos se fomento una filosofía con valores cívicos, éticos y de orden por medio del estímulo y el ejemplo que comprometían el comportamiento del estudiante no solo dentro de la escuela sino fuera de ella, a demás se introdujeron hábitos de sobriedad, de economía y principios morales de honradez, honestidad y respeto.



En sus inicios contó con 22 alumnos matriculados, y luego de tres meses fue cerrada por la poca cantidad de estudiantes, fue reabierta un año después, el 2 de enero de 1888, bajo la rectoría de Tulio Ospina V, esta vez contó con 27 alumnos matriculados y con un plan de estudios de 4 años de un mejor control de los programas curriculares y adaptarlos a nuevas condiciones adelantándose a las necesidades futuras de la educación y asegurando así un buen desempeño de los futuros profesionales.

En 1906 la Escuela Nacional de Minas se anexo a la universidad de Antioquia, a la que perteneció durante cinco años más, en 1911 paso a ser de nuevo una entidad independiente.

En 1940 la institución fue incorporada a la Universidad Nacional y continuó con el nombre de Escuela Nacional de Minas, ese mismo año comenzó la construcción de la actual sede, la cual fue inaugurada el 19 de diciembre de 1944, en el marco del primer Congreso Nacional de Ingenieros.

Entre 1941 y 1950 se crean las carreras de ingeniería geológica y petróleos y arquitectura, la cual se separo de la facultad de Minas en 1954, en 1960 se crea la carrera de ingeniería administrativa, luego se crearon los programas de ingeniería industrial, ingeniería mecánica e ingeniería química y se separaron los programas de ingeniería geológica y petróleos en dos programas diferentes, actualmente la Facultad de Minas Administra 11 programas de pregrado en ingeniería, 17 de posgrado y cuatro doctorados.

La Facultad a lo largo de su existencia ha sido motora del desarrollo de la ciudad, del departamento y del país, a través de sus 12.000 egresados quienes han constituido la mayor parte del personal dirigente y técnico en las explotaciones mineras, las construcciones de distinto tipo, la infraestructura vial, los desarrollos hidroeléctricos, las obras de abastecimiento de agua, las obras sanitarias y la industria, así como en los planes de desarrollo físico, económico y social.