

## **Técnicas de Inteligencia Computacional Aplicadas a Modelos de Estimación de Coagulante en el Proceso de Potabilización de Agua**

### **Computational Intelligence Techniques Applied to Coagulant Estimation Models in Water Purification Process**

*Carlos Alberto Villarreal Campos<sup>1\*</sup>, Eduardo Caicedo Bravo<sup>2</sup>*

<sup>1</sup> Departamento Control Procesos y Calidad en Planta. ACUAVALLE S.A. – E.S.P. Calle 56Norte No. 3N-19. Cali, Colombia.

<sup>2</sup> Escuela de Ingeniería Eléctrica y Electrónica. Universidad del Valle, Ciudad Universitaria Meléndez, A.A. 25360. Cali, Colombia.

(Recibido el 01 de mayo de 2011. Aceptado el 26 de junio de 2013)

#### **Resumen**

En este artículo se presentan cuatro modelos para la estimación de la dosificación de coagulante en el proceso de potabilización de agua, los cuales utilizan para su implementación, técnicas de inteligencia computacional que incluyen redes neuronales artificiales, lógica difusa tipo *Sugeno*, lógica difusa tipo *Mamdani* y estructuras neuro-difusas tipo ANFIS. La metodología descrita en la obtención de los modelos se basa en la extracción de la experiencia almacenada en los datos de operación de la planta (para el caso de los modelos basados en redes neuronales, lógica difusa Sugeno y estructuras ANFIS) y en el conocimiento expresado en forma lingüística por un operador experimentado de la planta (para el caso de la lógica difusa *Mamdani*). Los resultados obtenidos en las diferentes simulaciones realizadas demuestran que algunos de los modelos implementados pueden ser incluidos en sistemas de control que permiten superar algunas de las limitaciones que actualmente presentan los métodos de dosificación utilizados más ampliamente en las plantas de potabilización; la prueba de jarras y el *streaming current detector*.

**Palabras clave:** Potabilización del agua, dosificación de coagulante, redes neuronales artificiales, lógica difusa Sugeno, lógica difusa Mamdani, ANFIS

---

\* Autor de correspondencia: teléfono: +57+2+6653567 ext. 1266, fax: +57+2+2224261, correo electrónico: cvillarreal@acuavalle.gov.co (C. Villarreal)

## Abstract

Four coagulant estimation models in water purification process are presented. For its developing computational intelligence techniques are used, which include neural networks, fuzzy logic (Sugeno and Mamdani type) and ANFIS structures. The methodology described is based on the operator's experience extraction from operational historical data (in neural network, fuzzy logic Sugeno type and ANFIS case), and the linguistic information given by an experimented plant operator (in fuzzy logic Mamdani type case). According to the reached results, by applying some of this models in a control system strategy, could be possible overcome some of the current limitations found in the most widely used techniques for coagulant dosing in water purification plants; jar tests and streaming current detector .

**Keywords:** Water purification, coagulant dosing, artificial neural networks, fuzzy logic sugeno type, fuzzy logic Mamdani type, ANFIS

## Introducción

La obtención de agua potable, en su forma tradicional, implica el uso completo o parcial de cuatro sub-procesos denominados coagulación, floculación, sedimentación y filtración. La calidad y los costos del agua obtenida al final del proceso dependen en gran parte de la dosificación del coagulante. Un exceso en la dosificación eleva los costos de producción, mientras que una dosificación por debajo de la requerida puede llevar a la obtención de agua que no cumpla con los parámetros de calidad exigidos por la legislación respectiva y a una operación menos eficiente de la planta. Los métodos de dosificación más utilizados, como son la prueba de jarras y el *streaming current detector* (SCD), presentan limitaciones en su aplicación. Para el caso de la prueba de jarras se encuentran como limitaciones principales la alta dependencia de la experiencia del operador para su correcto funcionamiento y baja capacidad de respuesta ante cambios bruscos en las características del agua que ingresa a la planta. Por su parte, la dificultad en la parametrización del instrumento, la dificultad en la obtención del punto óptimo de funcionamiento y la necesidad de un mantenimiento frecuente y especializado del equipo se pueden mencionar como las desventajas principales que se desprenden del uso de un *streaming current detector* para la dosificación de coagulante en el proceso de potabilización

de agua. En este artículo se presentan cuatro modelos basados en inteligencia computacional, los cuales al ser incluidos en sistemas de control, podrían estar en la capacidad de superar las limitaciones antes mencionadas. En la primera parte del presente artículo se presentará una descripción de trabajos previos relacionados con modelos basados en inteligencia computacional para solucionar el problema de la dosificación de coagulante en el proceso de potabilización de agua. Posteriormente, se presentaran los modelos neuronales, difusos y neuro-difusos propuestos en el presente artículo, se revisará la experimentación desarrollada para la obtención de los modelos previamente propuestos y se analizaran los resultados obtenidos. Finalmente, se presentaran las conclusiones del trabajo realizado.

## Estado del arte

En trabajos como [1,2] para mencionar solo algunos, se ha resaltado la dependencia existente entre la dosificación de coagulante requerida para la correcta potabilización del agua y variables tan diversas como la *turbiedad*, *el pH*, *la temperatura*, *la alcalinidad*, *la conductividad* y *el color* del agua cruda. Además de la dependencia de esta gran cantidad de parámetros, la relación de la rata de dosificación de coagulante con cada uno de ellos es altamente no lineal lo cual, a finales del siglo pasado y principios del

siglo XXI popularizó el uso de las herramientas de inteligencia computacional para resolver el problema de la dosificación de coagulante en el proceso de potabilización del agua. La técnica más utilizada ha sido las redes neuronales, pero también se reportan trabajos que utilizan lógica difusa y modelos neuro-difusos tipo ANFIS.

### Redes neuronales

Uno de los primeros trabajos, que emplea redes neuronales para la estimación de dosificación de coagulante, data de 1995 y fue desarrollado con datos operativos de la planta de potabilización de Wyong en Australia [1]. Siguiendo esta misma línea, se construyeron modelos de estimación de coagulante como los reportados en [2-8]. Dichos trabajos desarrollados entre 1999 y 2009 si bien perseguían el mismo objetivo, predecir la dosificación de coagulante, presentaban ciertas particularidades en su implementación y aplicación. Por ejemplo, en [2] el modelo obtenido hacía parte de un sistema de dosificación automático de coagulante para la planta de potabilización de Viry –Chatillon (Francia), a diferencia del modelo implementado en [1] el cual solo se utilizó como herramienta de apoyo para la toma de decisiones por parte del operador de planta. En [3] el modelo de estimación implementado, que hacía parte de un amplio estudio realizado por la Universidad de Newcastle (Inglaterra), era un modelo directo, contrario a los modelos inversos planteados en trabajos como [1, 2, 4-8]. En el caso del modelo directo, este no determina directamente la dosificación de coagulante como sucede en los modelos inversos, sino que predice un parámetro de salida del proceso, el cual es utilizado para minimizar funciones de costo previamente definidas. En [4] el modelo inverso de predicción de coagulante, es utilizado en conjunto con otros modelos directos para la implementación de dos herramientas de simulación, las cuales además de permitir la estimación de coagulante, hacían posible el entrenamiento de los operadores de planta, mediante la implementación de modelos capaces de predecir los parámetros de calidad del agua

tratada basándose en los parámetros de calidad del agua cruda y la dosificación de coagulante aplicada. En [5] se desarrollan 17 modelos, varios de ellos del tipo auto-predictivo, que basan su estimación únicamente en las dosificaciones pasadas, y comparan su desempeño respecto a los otros modelos que basan su predicción en los parámetros de calidad del agua cruda y que pueden incluir o no dosificaciones pasadas. El trabajo reseñado en [6] presenta como mayor novedad respecto a los estudios anteriores la inclusión de la turbiedad del agua sedimentada dentro de las variables de entrada del modelo, lo cual en teoría podría permitir la utilización del mismo para reducir costos de operación. Finalmente en [7, 8] se presenta la implementación de modelos predictivos basados en redes neuronales para la estimación de coagulante, desarrollados a partir de datos de operación de plantas de potabilización de Marruecos y China respectivamente.

### Lógica difusa

En [9], publicado en 1997, para resolver el problema de la dosificación de coagulante en la planta de potabilización de Seúl (Corea del Sur) se consideraron dos condiciones de operación denominadas normales (turbiedades bajas) y anormales (turbiedades superiores a 30 NTU). Para la condición anormal, se desarrolló un modelo basado en redes neuronales el cual determinaba la dosificación de coagulante a emplear. Para cumplir esta función en la condición normal se desarrolló un modelo de inferencia difuso. Para la determinación de los parámetros que definían las funciones de membresía (centro y ancho izquierdo y derecho del triángulo) se emplearon algoritmos genéticos. En [10] se presenta una aplicación de control difuso para la dosificación de coagulante en la planta *Wiesner* que provee cerca del 70% del agua consumida en la ciudad de Bogotá (Colombia). El modelo se construyó con base en la información entregada por expertos en operación de la planta y se obtuvieron un total de 175 reglas como base del conocimiento del modelo desarrollado. Finalmente en 2009, se presenta un sistema de inferencia difusa el cual

no solo incluye variables del agua cruda como entradas del modelo, sino también la turbiedad del agua tratada y la señal proveniente de un analizador de dispersión fotométrico el cual mide el tamaño del *floc* después de la coagulación [11]. Las reglas del sistema de inferencia difuso desarrollado fueron obtenidas a partir de pruebas realizadas en planta y a la información aportada por el operador de la planta de Harbin (China).

### Estructuras ANFIS

En [12] se describe el uso de la arquitectura ANFIS para superar la limitación que tienen los modelos de lógica difusa en cuanto a la definición de las reglas *if-then* para el caso de la estimación de dosificación de coagulante, debido a la poca regularidad que se puede encontrar en las variables que intervienen en la dosificación y la problemática del crecimiento exponencial de las reglas *if-then* con el número de entradas utilizadas en el modelo, mediante el uso de los métodos *Conditional Fuzzy C-means* y *Fuzzy Equalization*. De igual manera, en [5] se realiza una comparación entre los resultados obtenidos por 17 modelos basados en redes neuronales y 17 basados en estructuras ANFIS.

### Estimación de Coagulante basada en inteligencia computacional

Del estado del arte, es claro que existen varias técnicas de inteligencia computacional disponibles para la implementación de modelos de estimación de coagulante para la potabilización del agua. Definir cuál de las técnicas señaladas es la más apropiada para el problema de dosificación de coagulante, dependerá de factores como el objetivo del modelo a implementar (herramienta de apoyo para la toma de decisiones por parte del operador, capacitación de personal, automatización de la dosificación) la información disponible para la implementación del modelo, precisión de los modelos obtenidos y herramientas disponibles para su aplicación.

El principal propósito del presente artículo es la implementación de cuatro modelos de estimación de coagulante utilizando cuatro técnicas de inteligencia computacional diferentes (redes neuronales, lógica difusa Sugeno, lógica difusa Mamdani y ANFIS) evaluar su desempeño y establecer el rango de aplicaciones de cada modelo obtenido.

### Experimentación y resultados

De los cuatro modelos a ser obtenidos, tres de ellos (redes neuronales, difuso Sugeno y ANFIS) serán desarrollados a partir de los resultados de las pruebas de jarras efectuadas diariamente por los operadores de la planta de potabilización del municipio de Roldanillo (Valle del Cauca) durante un periodo de 2 años y 7 meses. Las ocho variables medidas en el agua cruda al momento de realizar la prueba de jarras incluyen *turbiedad, color, pH, alcalinidad, hierro, dureza, temperatura y conductividad*. A continuación se describirán los pasos necesarios para la obtención de cada uno de estos tres modelos.

### Detección y eliminación de outliers

Un *outlier* puede definirse como: “una observación en un conjunto de datos la cual parece ser inconsistente con el conjunto de datos restantes” [13]. La no detección de *outliers* y su inclusión en el proceso de obtención de modelos pueden llevar a resultados incorrectos o sesgos en la predicción final de los parámetros estimados por el modelo. La detección de *outliers* en sí, es un campo bastante amplio, el cual ha sido tema de una gran variedad de trabajos, donde se ha abordado la búsqueda de soluciones mediante aproximaciones estadísticas, minería de datos, *clustering* y redes neuronales, entre otros. En el presente trabajo se empleó la metodología descrita en [14], que combina métodos estadísticos y redes neuronales para la detección y eliminación de *outliers*.

## Selección de las variables que servirán de entrada al modelo

A continuación fue necesario determinar si entre las 8 variables con que inicialmente se cuenta para la implementación del modelo existe información redundante, o en otras palabras, si existen variables altamente correlacionadas de manera tal que no se requieran incluir las 8 variables como entradas al modelo, simplificando así su implementación. Para tal fin, se utilizó el método descrito en [15] el cual se basa en la interpretación de los planos de componentes de los mapas auto-organizados de Kohonen (SOMs, por sus siglas en inglés) hallados para cada variable. Mediante la comparación de los planos de componentes fue posible descartar, el *color*, la *alcalinidad* y la *dureza*, como entradas al modelo por presentar estas variables alta correlación con la *turbiedad* y la *conductividad*.

La siguiente consideración, relacionada con la selección de las variables que servirá de entrada al modelo, está relacionada con determinar, cuáles de las variables restantes, *turbiedad*, *pH*, *conductividad*, *hierro* y *temperatura*, afectan realmente la salida del modelo, es decir la dosificación de coagulante. Para tal fin se empleó una metodología similar a la propuesta en [16] consistente en un proceso de selección tipo *backward*, en el cual se inicia con un modelo que incluya todas las variables candidatas y se van eliminando del modelo aquellas que al ser eliminadas no afectan de forma significativa el desempeño del modelo de estimación. De la aplicación de las metodologías mencionadas se obtiene finalmente que la *turbiedad*, el *pH* y la *conductividad* son las variables que se deben utilizar como entradas a los modelos a implementar.

## Implementación de los modelos

Para la implementación del modelo de redes neuronales, se empleó una arquitectura perceptron multicapa (*MLP*, *Multi-Layer Perceptron*), aunque también se dispone de desarrollos con redes de base radial (*RBF*, *Radial*

*Basis Function*). Una discusión sobre las ventajas de cada una de estas arquitecturas aplicada a problemas de control puede encontrarse en [17]. Para la implementación del modelo difuso tipo *Sugeno*, se empleó la técnica de *subtractive clustering* la cual presenta como principal ventaja sobre otras técnicas de *clustering* que no requiere el conocimiento *a priori* del número de *clusters* y sus ubicaciones. Información adicional sobre las diferentes técnicas de *clustering* disponibles para la implementación de modelos difusos y sus características se pueden encontrar en [18-21]. Para la implementación del tercer modelo, se hizo uso de una arquitectura ANFIS para ajustar el modelo Sugeno desarrollado previamente mediante *subtractive clustering*. Con la utilización de la estructura ANFIS, se modifican los parámetros  $\sigma$  (ancho) y  $c$  (centro) de las funciones gaussianas que definen las funciones de membresía del sistema difuso.

## Índices de desempeño

Para evaluar y comparar los resultados de las estimaciones efectuadas por los diferentes modelos implementados, se emplearon tres índices de desempeño. El primer índice es el error cuadrático medio (*RMSE*, por sus siglas en inglés)  $RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^n (o_p - o_d)^2}$  en donde,  $n$  es el número de datos,  $o_p$  es el valor estimado por el modelo de estimación y  $o_d$  el valor real. Otro índice empleado fue el coeficiente de correlación ( $R^2$ ) entre los valores reales y los datos estimados por el modelo de estimación. Finalmente, el error medio absoluto (*MAE*, por sus siglas en inglés)  $MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i|$ , donde  $f_i$  corresponde a la estimación del modelo,  $y_i$  al valor real y  $n$  al número de datos evaluados, fue el tercer índice de desempeño utilizado.

## Modelo de redes neuronales

Para la obtención del modelo de redes neuronales, se realizaron varias simulaciones modificando el número de neuronas en la capa oculta (se trabajó con una sola capa oculta, ya que capas

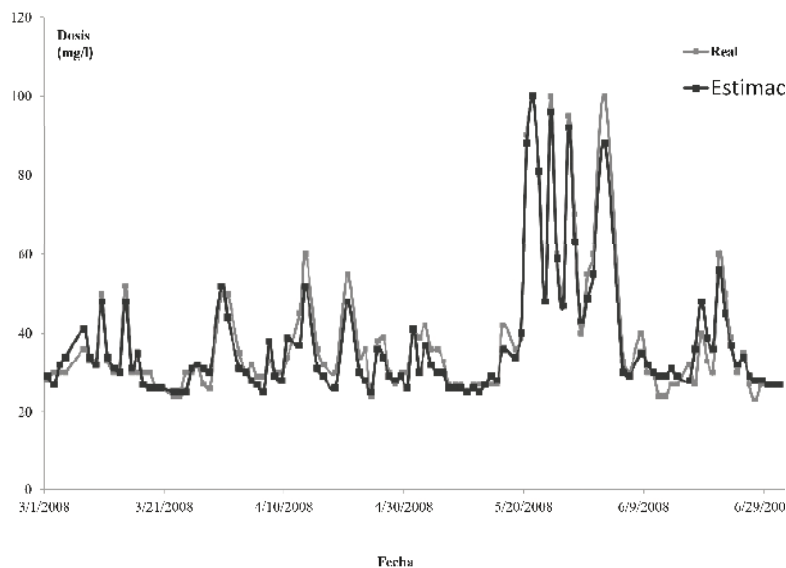


adicionales no mejoraban significativamente el desempeño del modelo), se encontró que los mejores resultados se obtuvieron con 7 neuronas en dicha capa, tal como se observa en la tabla 1.

En la figura 1, se observa una comparación entre las dosificaciones realizadas por medio de prueba de jarras y las estimadas por el modelo de redes neuronales.

**Tabla 1** Desempeño de los modelos de redes neuronales obtenidos

No. Neuronas	$R^2$	RMSE	MAE (mg/l)
9	0,9128	0,1690	3,4445
8	0,9322	0,1521	3,2093
7	0,9393	0,1419	2,9642
6	0,9382	0,1433	2,9860
5	0,9246	0,1575	3,1639
4	0,9228	0,1594	3,3181



**Figura 1** Prueba de Jarras (Real) v.s. Modelo Neuronal (Estimado)

### Modelo difuso tipo sugeno

Al utilizar la metodología de *subtractive clustering* para la obtención del modelo difuso, el parámetro más importante a definir es el radio del clúster, el cual determina el rango de influencia de los centros de los clústeres y cuya selección afecta tanto los índices de desempeño del modelo difuso como su número de reglas. Un

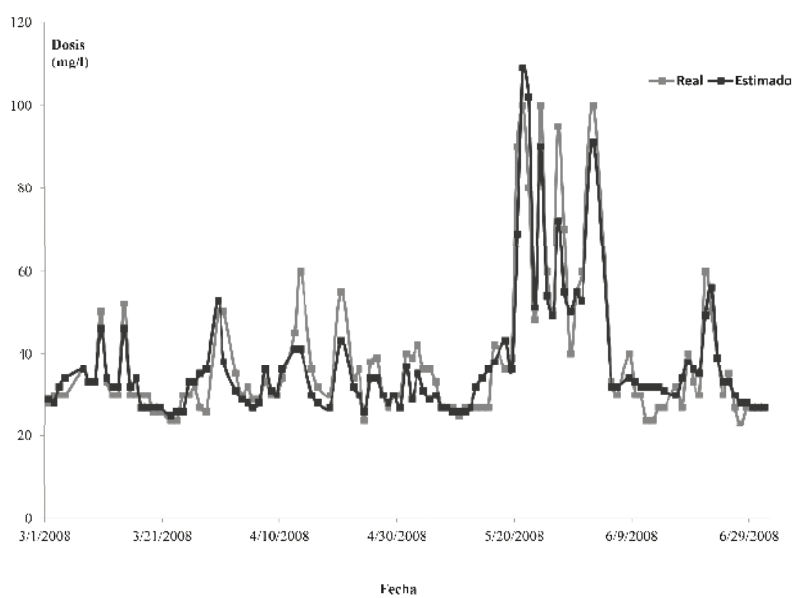
radio muy grande resultaría en pocos clústeres y por consiguiente en un modelo muy burdo. Un radio muy pequeño, por su parte, implicaría un número excesivo de clústeres y en un modelo con capacidad limitada de generalización (*overfitting*) [19]. Tomando como base los mismos datos empleados para el desarrollo del modelo de redes neuronales, se puede observar en la tabla 2 los índices de desempeño y el número de

reglas obtenidas para distintas selecciones en el radio del clúster. En la figura 2, se comparan las dosificaciones obtenidas por medio de pruebas de

jarras y las estimaciones realizadas por el modelo difuso tipo Sugeno.

**Tabla 2** Desempeño del modelo Sugeno v.s. radio del clúster

<i>Radio</i>	<i>R<sup>2</sup></i>	<i>RMSE</i>	<i>MAE (mg/l)</i>	<i>No. de Reglas</i>
<b>0,5</b>	0,8533	0,2158	4,4150	1
<b>0,2</b>	0,8912	0,1880	3,6800	4
<b>0,15</b>	0,8962	0,1839	3,6048	7
<b>0,1</b>	0,9043	0,1771	3,4488	11
<b>0,09</b>	0,9046	0,1763	3,3658	14



**Figura 2** Prueba de Jarras (Real) v.s. Modelo Sugeno r=0.15 (Estimado)

### Modelo neuro-difuso tipo ANFIS

Haciendo uso de una arquitectura ANFIS se realiza el ajuste del modelo Sugeno desarrollado en la sección anterior. Con la utilización de la estructura ANFIS, se modifican los parámetros  $\sigma$  (ancho) y  $c$  (centro) de las funciones gaussianas que definen las funciones de membresía del sistema difuso. Cabe resaltar, como el sistema

difuso Sugeno hallado por *subtractive clustering* deja fijo el parámetro  $\sigma$  en las funciones de membresía de cada variable (de hecho para la *turbiedad* y el *pH* el parámetro es el mismo) mientras que al implementar el modelo ANFIS dicho parámetro es diferente en cada una de las funciones de membresía, tal como se observa en las tablas 3 y 4.

**Tabla 3** Parámetros de la funciones de membresía modelo Sugeno

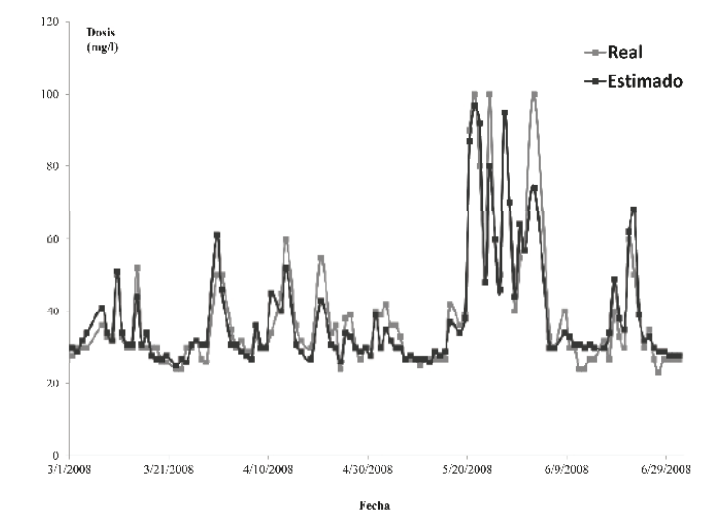
<i>Función de Membrecía</i>	<i>Turbiedad (<math>\sigma/c</math>)</i>	<i>pH (<math>\sigma/c</math>)</i>	<i>Conductividad (<math>\sigma/c</math>)</i>
1	0,0530/0,0104	0,0530/0,6552	0,0510/0,8556
2	0,0530/0,0048	0,0530/0,5172	0,0510/0,8770
3	0,0530/0,0226	0,0530/0,6207	0,0510/0,7487
4	0,0530/0,0045	0,0530/0,6897	0,0510/0,9519
5	0,0530/0,0102	0,0530/0,2414	0,0510/0,5348
6	0,0530/0,0096	0,0530/0,8276	0,0510/0,7968
7	0,0530/0,0022	0,0530/0,5172	0,0510/1,0000

**Tabla 4** Parámetros de la funciones de membresía modelo ANFIS

<i>Función de Membrecía</i>	<i>Turbiedad (<math>\sigma/c</math>)</i>	<i>pH (<math>\sigma/c</math>)</i>	<i>Conductividad (<math>\sigma/c</math>)</i>
1	0,0769/0,0183	0,0542/0,6536	0,0649/0,8578
2	0,0494/0,0042	0,0514/0,5175	0,0755/0,8701
3	0,0385/0,0166	0,0410/0,6294	0,0569/0,7498
4	0,0473/0,0012	0,0548/0,6901	0,0623/0,9461
5	0,0459/0,0109	0,0486/0,2401	0,0541/0,5334
6	0,0530/0,0096	0,0546/0,8264	0,0570/0,7936
7	0,0531/0,0026	0,0548/0,5176	0,0629/0,9953

En la figura 3, se observa una comparación entre las dosificaciones realizadas por medio de prueba

de jarras y las estimadas por el modelo neuro-difuso ANFIS.



**Figura 3** Prueba de Jarras (Real) v.s. Modelo ANFIS (Estimado)



### Modelo difuso tipo mamdani

A diferencia de los modelos anteriores que fueron obtenidos a partir de los datos históricos de operación de la planta, la obtención del modelo para estimación de la dosificación de coagulante tipo Mamdani fue a través de una entrevista realizada a uno de los operadores de la planta del municipio de Roldanillo (Valle del Cauca). Aunque inicialmente fue planteado al operador, la posibilidad de utilizar las mismas tres variables de entrada para la construcción

de las reglas (*turbiedad, pH y conductividad*), se trabajó con la *turbiedad* y el *pH*, ya que para el operador no era claro como cuantificar la influencia de la conductividad en el resultado final de la dosificación de coagulante. Dado que se seleccionaron 7 variables lingüísticas para la *turbiedad* y 3 para el *pH*, se generaron un total de 21 reglas del tipo *if-then*. La comparación entre las estimaciones realizadas por este modelo Mamdani con la obtenida por la pruebas de jarras se puede observar en la figura 4.

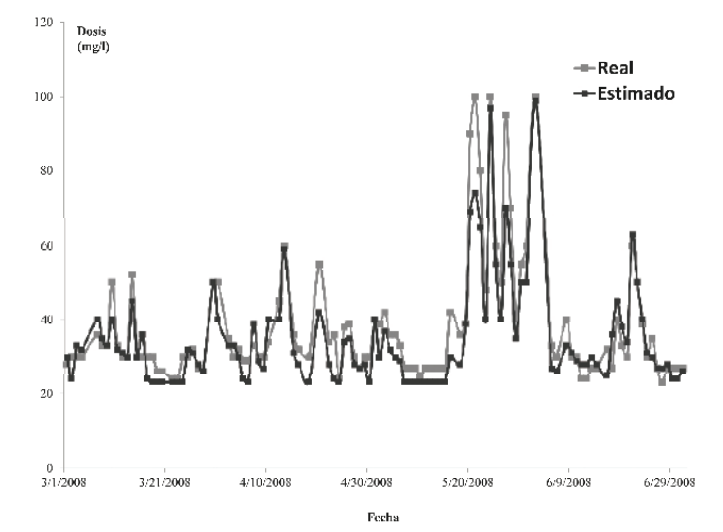


Figura 4 Prueba de Jarras (Real) v.s. Modelo Mamdani (Estimado)

### Discusión de los resultados obtenidos

En la tabla 5 se muestran los índices de desempeño obtenidos para el modelo de redes neuronales con

7 neuronas en la capa oculta, el modelo *Sugeno* de 7 reglas, el modelo ANFIS y el modelo *Mamdani*.

Tabla 5 Resumen índices de desempeño de los modelos obtenidos

Modelo	R <sup>2</sup>	RMSE	MAE (mg/l)
Redes Neuronales	0,9393	0,1419	2,9642
Sugeno	0,8962	0,1839	3,6048
ANFIS	0,9232	0,1592	3,0501
Mamdani	0,8615	0,2298	4,7160

De acuerdo con lo observado en la tabla 5, los mejores índices de desempeño son alcanzados con el modelo de estimación basado en redes neuronales. La sencillez de este modelo hace que sea posible su utilización en esquemas de control, mediante su implementación en equipos industriales tipo controladores lógicos programables, que cuenten con capacidad de ejecutar operaciones básicas en coma flotante e incluyan la función exponencial lo cual es indispensable para la implementación de las funciones de activación de las redes neuronales. El modelo de inferencia tipo *Sugeno* obtenido mediante *subtractive clustering* muestra un desempeño inferior al obtenido con redes neuronales. Inclusive, tal como se muestra en la tabla 2, disminuir el radio de clúster (lo cual aumenta el número de reglas) no introduce mejoras significativas al desempeño del modelo y por el contrario incrementaría la complejidad de su implementación. Sin embargo, mediante la utilización de la estructura ANFIS fue posible afinar las reglas del modelo *Sugeno* hallado previamente, llegándose a obtener índices de desempeño comparables a los encontrados con el uso de redes neuronales. La implementación de este modelo de inferencia también es posible en un controlador lógico programable, siempre y cuando se cuente con la posibilidad de efectuar operaciones básicas en coma flotante y funciones tipo exponencial, lo cual es necesario para la implementación de las funciones de membresía del modelo. El modelo de inferencia tipo Mamdani fue el que presentó los índices de desempeño más bajos y además al estar construido a partir de información lingüística, su implementación en un controlador lógico programable no es tan directa como en el caso de las redes neuronales y en el modelo ANFIS, y se requeriría de un procesamiento previo del modelo antes de su implementación.

### Conclusiones

Se comparan cuatro modelos obtenidos mediante técnicas de inteligencia computacional que permiten la estimación de la dosificación

de coagulante en el proceso de potabilización del agua. Los tres primeros modelos (redes neuronales, Sugeno y ANFIS) se obtuvieron a partir de información numérica correspondiente a las pruebas de jarras realizadas por los operadores de planta y el último a partir de información lingüística proporcionada por un operador (Mamdani). De acuerdo con lo observado, el modelo que presenta los mejores índices de desempeño es el basado en redes neuronales, mientras que de los tres modelos que se implementaron usando inferencia difusa el que mejor resultado entregó fue el que empleó la estructura ANFIS, la cual comparte con las redes neuronales el mismo tipo de entrenamiento. Por su parte el Modelo *Mamdani* el cual se obtuvo a partir de información lingüística proporcionada por el operador de planta fue el que peores índices de desempeño presentó, sin embargo, presenta la ventaja de que fue construido con base en reglas fácilmente entendibles por el ser humano. Los bajos MAE obtenidos para los modelos basados en redes neuronales y en la estructura ANFIS, hacen factible su inclusión en esquemas de control que permitan la dosificación de coagulante en línea, en un proceso real de potabilización de agua. Por otra parte, si no se cuenta con la infraestructura requerida para la implementación de un esquema de control que haga uso de estos modelos (sensores, controlador lógico programable, bombas dosificadoras, etc.) los modelos aquí presentados pueden ser implementados, inclusive en hojas de cálculo, constituyéndose de esta manera, en una herramienta de apoyo para el operador de planta.

### Agradecimientos

Los autores del presente artículo expresan su agradecimiento al Programa de Posgrado de la Escuela de Ingeniería Eléctrica y Electrónica de la Universidad del Valle y a la empresa ACUAVALLE S.A. –E.S.P. por el apoyo prestado durante el desarrollo de la presente investigación.

## Referencias

1. A. Mirsepassi, B. Carthers, H. Dharmappa. "Application of Artificial Neural Networks to the Real Time Operation of Water Treatment Plants". *Proceedings of the IEEE International Conference on Neural Networks*. Vol. 1. 1995. pp. 516-521.
2. N. Valentin, T. Denoux, F. Fotoohi. "An Hybrid Neural Network Based System for Optimization of Coagulant Dosing in a Water Treatment Plant". *International Joint Conference on Neural Networks*. Vol. 5. 1999. pp. 3380-3385.
3. B. Lennox, G. Montague, A. Frith, C. Gent, V. Bevan. "Industrial application of neural networks – An investigation". *Journal of Process Control*. Vol 11. 2001. pp. 497-507.
4. H. Maeir, N. Morgan, W. Chow. "Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters". *Environmental Modelling & Software*. Vol. 19. 2004. pp. 485-494.
5. G. Wu, S. Lo. "Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network-based fuzzy inference system". *Engineering Applications of Artificial Intelligence*. Vol. 21. 2008. pp. 1189-1195.
6. X. Yi, Q. Wei, C. Li. *The BP network model for the control of the coagulation and the filter backwashing system*. The Second International Conference on Bioinformatics and Biomedical Engineering. Shangai, China. 2008. pp. 1518-1521.
7. B. Lamrini, A. Benhammou, A. Karama, M. Le Lann. "A Neural Network System for modeling of coagulant dosage used in drinking water treatment". *Adaptive and Natural Computing Algorithms*. Part I. 2005. pp. 96-99.
8. Z. Song, Y. Zhao, X. Song, Ch. Liu. *Research on Prediction Model of Optimal Coagulant Dosage in Water Purifying Plant Based on Neural Networks*. International Colloquium on Computing, Communication, Control and Management ISECS. Sanya, China. 2009. pp. 258-261.
9. E. Nham, T. Ham, K. Woo, C. Kim, J. Ryu. *Optimization of coagulant dosing process in water purification system*. Proceedings of the 36<sup>th</sup> SICE Annual Conference, International Session Papers. Tokushima, China. 1997. pp. 1105-1109.
10. M. Duque, E. Giraldo, A. Martínez, A. Gauthier, J. Villa. "El Uso de la Lógica Difusa para la Potabilización del Agua". *Revista de Ingeniería de la Universidad de los Andes*. No. 12. 2000. pp. 31-35.
11. H. Bai. *A Fuzzy Logic Based Coagulant Real Time Control Squeume for Water Purification Systems*. Proceedings of the IEEE International Conference on Mechatronics and Automation. Changchun, China. 2009. pp. 2123-2127.
12. M. Chun, K. Kwak, J. Ryu. "Application of ANFIS for Coagulant Dosing Process in a Water Purification Plant". *Fuzzy Systems Conference Proceedings*. Vol 3. 1999. pp. 1743-1748.
13. I. Ben. *Outlier Detection*. Ed. Kluwer Academic Publishers Dordrecht, Netherlands Available on: Consultado en Julio de 2010. [www.eng.tau.ac.il/~bengal/outlier.pdf](http://www.eng.tau.ac.il/~bengal/outlier.pdf) Accessed: july 2010.
14. D. Joo, D. Choi, H. Park. "The effects of data preprocessing in the determination of coagulant dosing rate". *Wat. Res*. Vol. 34. 2000. pp. 3295-3302.
15. M. Santos, A. Monteiro, J. Simeao. *Visualization of geospatial data by component planes and U- Matrix*. Disponible en: [www.geoinfo.info/geoinfo2004/papers/6419.pdf](http://www.geoinfo.info/geoinfo2004/papers/6419.pdf). Consultado en Julio 2010.
16. S. Chiu. "Selecting Input Variables for Fuzzy Models". *Journal of Intelligent & Fuzzy Systems*. Vol. 4. 1996. pp. 243-256.
17. G. Bloch, T. Denouex. "Neural Networks for Process Control and Optimization: Two Industrial Applications". *ISA Transactions*. Vol 42. 2003. pp. 39-51.
18. D. Kaur, D. Baumgartner. *A comparative Analysis of Neuro-Fuzzy and Grammatical Evolution models for simulating Field-Effect Transistors*. World Congress on Computer Science and Information Engineering. Toledo, USA. Vol. 5. 2009. pp. 179-183.
19. S. Chiu. *Extracting Fuzzy Rules from Data for Function Approximation and Pattern Classification*. Disponible en: <http://chius.homestead.com/files/ExtractRulesFromData.pdf>. Consultado en Julio 2010.
20. U. Kayman, M. Setnes. *Extended Fuzzy Clustering Algorithms*. Disponible en: <http://publishing.eur.nl/ir/repub/asset/57/erimrs20001123094510.pdf>. Consultado en Julio 2010.
21. W. Zhang, H. Hu, W. Liu. *Rules Extraction of Interval Type-2 Fuzzy Logic System based on Fuzzy c-Means Clustering*. 4<sup>th</sup> International Conference on Fuzzy Systems and Knowledge Discovery. Haikou, China. Vol. 2. 2007. pp. 256-260.