

## Aproximación al Modelo de Estimación Para El Uso De Agua Del Rio Bogotá, Basado En El Análisis De Vertimientos En Aguas Superficiales

### Approach to Model Estimate for Bogotá River Water Use, Based on Analysis of Discharge into Surface Waters

### Abordagem Para Modelar Estimativa Para Bogotá River Water Use, Com Base Na Análise Dos Efluentes Em Águas Superficiais

Wilson Javier Peña Caro<sup>1</sup>

Daissy Díaz<sup>2</sup>

Paulo Gaona García<sup>3</sup>

Carlos Enrique Montenegro Marín<sup>4</sup>

Mario Castro<sup>5</sup>

#### Resumen

En este documento se describe el desarrollo de una metodología para el análisis de las variables en las fuentes de agua en particular el río Bogotá, con el fin de evaluar, probar y predecir la tasa de contaminación de las aguas, se plantean los aspectos necesarios para el procesamiento de datos para que pueda ser extraído de la base de datos que maneja la CAR. Adicionalmente se hace una revisión de algunos trabajos previos y de la importancia de determinar la forma en que se miden los índices de calidad del agua que se propone en este ámbito de trabajo, el enfoque de las variables más importantes que deban ser considerados para esta investigación, de acuerdo con los procedimientos de análisis de datos estándar.

#### Palabras Clave:

Gestión del conocimiento, Calidad del agua, Análisis de Datos, Río Bogotá.

#### Abstract

This document describes the development of a methodology for analysis of the variables in the water sources particularly the river Bogotá, in order to evaluate, test and predict the rate of water pollution, it is proposed raised the necessary aspects for

<sup>1</sup> Universidad Distrital Francisco José de Caldas Bogotá- Colombia [wjpenac@correo.udistrital.edu.co](mailto:wjpenac@correo.udistrital.edu.co)

<sup>2</sup> Universidad Distrital Francisco José de Caldas Bogotá- Colombia [daissymdiaz@gmail.com](mailto:daissymdiaz@gmail.com)

<sup>3</sup> Universidad Distrital Francisco José de Caldas Bogotá- Colombia [pagaonag@udistrital.edu.co](mailto:pagaonag@udistrital.edu.co)

<sup>4</sup> Universidad Distrital Francisco José de Caldas Bogotá- Colombia [cemontenegrom@udistrital.edu.co](mailto:cemontenegrom@udistrital.edu.co)

<sup>5</sup> Universidad Cooperativa de Colombia Bogotá- Colombia [mario.castrof@campus.ucc.edu.co](mailto:mario.castrof@campus.ucc.edu.co)

	<p>processing data so that it can be extracted from the database that has the CAR. a review of some previous work and the importance of determining how they will be measured indices water quality as proposed in this work scope is done, the approach of the most important variables are to be considered for this research, according to procedures standard data analysis.</p> <p><b>Keywords:</b> Knowledge management, water quality, data analysis, Bogotá River.</p> <p><b>Resumo</b> Este documento descreve o desenvolvimento de uma metodologia para a análise de variáveis em fontes de água, incluindo o rio Bogotá, a fim de avaliar, testar e prever a taxa de poluição da água descrito, levantou os aspectos necessários para processamento de dados de modo que pode ser extraído da base de dados que gere a CAR. Além disso, uma revisão de alguns trabalhos anteriores e da importância de determinar como índices de qualidade de água é proposto nesta área de trabalho são medidos é o foco das variáveis mais importantes que devem ser considerados para esta pesquisa de acordo com procedimentos padrão de análise de dados.</p> <p><b>Palavras-chave:</b> Gestão do conhecimento, qualidade da água, análise de dados, Río Bogotá.</p>
--	---

## INTRODUCCIÓN

Las fuentes hídricas son uno de los factores geológicos más importantes y que se han estudiado a lo largo del tiempo con el fin de poder predecir su comportamiento y poder evaluar sus características ya que se constituyen como un recurso primordial para la vida humana y de todos los organismos vivos que habitan el planeta. De esta manera se hace necesario realizar monitoreo de las principales características que las fuentes hídricas pueden presentar, sin embargo, el “monitoreo no es sólo hacer mediciones: se reconoce cada vez más que los datos deben estar disponibles. La comunicación de los resultados implica una retroalimentación, y además, permite que la información recopilada sea utilizada en las decisiones de gestión”(Castro et al. 2014).

Para el caso de Colombia y especialmente para Bogotá y Cundinamarca el análisis de los datos disponibles del Río Bogotá se hace necesario principalmente porque se requiere evaluar la calidad del agua y los factores que pueden influir en mayor o menor medida sobre este cuerpo de agua. Esta evaluación es importante para la preservación de los recursos naturales y la productividad de la zona ya que del agua depende gran parte de las actividades industriales, económicas y sociales de cualquier país.

Con el desarrollo de las tecnologías de la información se han implementado diversos mecanismos de recopilación de datos geoespaciales como los GIS (Geospatial information systems) que manejan grandes bases de datos para almacenar las medidas que son

tomadas en terreno y en laboratorio de las variables presentes en las fuentes hídricas (Karimipour 2005). Debido a la gran cantidad de datos disponibles es difícil analizarlos y presentarlos de forma que puedan inferirse resultados o criterios que indiquen relaciones entre variables de tal manera que se hace necesario implementar métodos numéricos, estadísticos y computacionales que permitan extraer información de los datos y mediante tratamiento computacional presentarlos para que sean entendibles (Neighborhood Analysis in Water Pollution Estimation 2015). Es aquí en donde las herramientas como la minería de datos, las redes neuronales, la lógica difusa, entre otras cobran una gran relevancia en el estudio de los cuerpos de agua y de cualquier otro factor geológico que se quiera estudiar y que maneje enormes cantidades de datos.

## **El Problema De La Calidad del Agua En El Rio Bogotá**

### **¿De Dónde Salen los Datos y Como se Obtienen?**

Actualmente la CAR mide muchas variables presentes en el Rio Bogotá con personal en terreno y con pruebas de laboratorio que permiten obtener una base de datos amplia de las diferentes características del rio a lo largo del tiempo, desde el 2008 se viene construyendo una base de datos a partir de las variables que se miden sobre el rio, pero por la gran cantidad de datos obtenidos es difícil inferir una relación entre las variables y aun es más difícil sacar conclusiones a cerca del comportamiento del rio y la calidad del agua que se encuentra en las cuencas alta, media y baja .

El monitoreo directo es el mejor método para evaluar el impacto ambiental de los generadores de vertimientos, en ausencia de suficiente información de este tipo, una alternativa práctica, viable y menos costosa es la estimación teórica a partir de la aplicación de factores de vertimiento, denominada evaluación rápida o inventario rápido de fuentes de contaminación del agua (Orjuela et al. 2010) . Una herramienta importante para el análisis de la estimación teórica es el cálculo de los indicadores ambientales (Castro et al. 2014) que para este caso se denominan índices de calidad del agua los cuales toman un papel muy importante en el análisis de los datos ya que reducen el número de mediciones y de parámetros que normalmente se requieran para hacer una representación exacta de una situación y simplifican el proceso de comunicación de los resultados de la medición (Oecd 2004).

### **Pronostico del Problema**

Hoy en día se toman medidas de alrededor de 28 variables presentes en el Rio Bogotá sin saber que variables influyen de forma directa en la calidad del agua, además toda esta información no se está aprovechando porque no se está analizando ni tratando para poder inferir situaciones de riesgo o control sobre el rio.

El tiempo en que los profesionales realizan las pruebas y la toma de datos prácticamente se está perdiendo porque no se tienen conclusiones de esta labor y no se está aportando al control de rio, no se sabe si las variables medidas tienen alguna relación entre si y mucho menos si es necesario tomarlas todas o peor aún si son suficientes para realizar un diagnóstico.

De persistir esta situación la CAR estará realizando un trabajo infructuoso y seguramente estos datos se perderán por no ser útiles en la toma de decisiones, ya que son datos brutos que no muestran una relación directa con el comportamiento del río.

### **¿Qué es lo Que se Propone?**

Para poder obtener información y conocimiento a partir de los datos medidos se hace necesario desarrollar una herramienta computacional que permita tratar y procesar los datos y de esta manera poder sacar conclusiones para desarrollar un modelo de estimación para el uso de agua, basado en el análisis de vertimientos en aguas superficiales, que será una herramienta en el diagnóstico y toma de decisiones en cuanto a acciones de control sobre el río Bogotá.

### **Trabajos Previos**

A nivel mundial se han desarrollado múltiples métodos de análisis de datos que permiten obtener información de las variables presentes en los ríos ya que su estudio es de vital importancia en la preservación de los recursos naturales (Hammond et al. 1995).

Una serie de ítems tienen que evaluarse en este sentido, incluyendo la industria, la agricultura y las políticas. Una de las cuestiones clave es la descarga de efluentes en los ríos. Los reglamentos de afluentes deben cuidar principalmente tres grupos de elementos:

- a) La asignación de la carga de residuos
- b) La medición final de la tubería.
- c) La selección de parámetros.

La primera de las muchas políticas para combatir este problema fue la Ley de Agua Limpia de 1972 (Dutta and Chaki 2012) (Dutta and Chaki 2012). Los grupos de elementos incluían muchos atributos. Estos estudios marcaron el inicio de la integración de las variables del medio ambiente con el monitoreo geográfico y la previsión. La mejora tecnológica en la última década ha ayudado a recopilar gran cantidad de datos geoespaciales. Este creciente volumen de datos necesita ser transformado en conocimiento esencial mediante el desarrollo de nuevas tecnologías capaces de analizar y modelar los datos de forma más rápida y precisa. No se pueden implementar análisis convencionales como métodos estadísticos simples para explorar las perspectivas de evolución y la relación oculta entre y dentro de los datos espaciales digitales, ya que estos enfoques dependen de que el conjunto de datos sea confiable, que está lejos de ser la clase de datos que se toman de los afluentes hídricos por su carácter difuso.

Los métodos numéricos tradicionales de medición de la calidad del agua no son fáciles de usar y lo que se necesita son técnicas de aprendizaje inteligentes. En (Chau 2006), los autores han demostrado diferentes categorías de técnicas de análisis como Redes Neuronales Artificiales (RNA), Sistemas Difusos y diferentes algoritmos basados en ellos orientando su aplicación hacia la gestión de la calidad del agua. Los enfoques convencionales generalmente utilizan conceptos de conjunto clásicos, pero en el escenario de la vida real, en lugar de tener valores fijos, los parámetros tienen rango de valores y conceptos establecidos de manera que muchas veces conducen a resultados erróneos. Para superar este problema los investigadores en (Prato 2005) han demostrado la aplicabilidad de proposiciones difusas para los atributos. En (Lobbrecht, Dibike, and

Solomatine 2002) se han estudiado las aplicaciones de redes neuronales, lógica difusa y los enfoques híbridos hacia la gestión de la calidad del agua. Los autores han identificado la aplicabilidad de los diferentes modelos en diferentes cuerpos de agua. También se han sugerido diferentes procedimientos para hacer frente a ciertos problemas.

## Resumen y Recopilación de Trabajos Relacionados

A continuación, se hace un resumen de algunos trabajos desarrollados alrededor del análisis de datos de diferentes variables medidas sobre cuerpos de agua y sus posibles aplicaciones:

**Tabla 1.** Determinación de las cargas diarias Máxima Total TMDLs para cuerpos de agua.

Modelo/ Técnica	Resultado	Falencias
Modelado MLP ANN entrenado utilizando algoritmos de propagación y de gradiente conjugado. Un optimizador numérica no lineal, basado restricción a indicar en TMDLs , (Survey and Road 2003), (Conrads, Roehl, and Cook 2002), (Conrads et al. n.d.)	Efecto de la carga en la DO (Oxígeno disuelto) para regular TMDLs y tratamiento de aguas residuales.	Dependen solamente de la variable DO.
La aplicación de la minería de datos con el modelado MLP ANN entrenado retroalimentado con algoritmos de propagación y de gradiente conjugado. (Paul Conrads 2014)	Efecto de las precipitaciones, SC, WL y amplitud de la marea en la DO y por lo regular TMDLs para la gestión de las aguas costeras.	No se han utilizado parámetros distintos de la calidad del agua que demuestra DO.

**Tabla 2.** Gestión de la Calidad del Agua.

Modelo/ Técnica	Resultado	Falencias
Análisis vecindario utilizando dos redes KSOM dimensional. (Neighborhood Analysis in Water Pollution Estimation 2015)	Un mapa que proporciona un sistema que permite organizar la gestión de la calidad del agua.	Costo computacional con un enfoque heurístico.
Análisis de Dependencia y ajuste de curvas. (Karimipour 2005)	Dependencia e independencia de los parámetros de forma positiva o negativa.	Necesidades de datos precisos. La simulación de datos se ha realizado para los datos reales no disponibles.
Software de aplicación con lógica difusa utilizando funciones de pertenencia trapezoidales. (Edwin A. Roehl, Jr., John B. Cook, PE, Greenville 2013)	Índice difuso de calidad del agua (FWQ) para evaluar la calidad del agua y la comparación con WQL (índice de calidad del agua) real.	Los detalles relativos a la asignación de los factores de peso son desconocidos.

**Tabla 3.** Determinar el impacto del proceso de AMD (Drenaje Acido de Minas)

<b>Modelo/ Técnica</b>	<b>Resultado</b>	<b>Falencias</b>
Generador de reglas predictivas difusas y algoritmo de clustering (basado en lógica difusa y técnicas de minería de datos) (V, Bouwmeester, and Mohan 2009)	Gráficas que muestran el comportamiento de pH en comparación con el resto de las variables como un todo.	Sólo funciona con los datos actuales no usa de datos históricos ni predicción.
Min-max sistema de inferencia difuso con transformación logarítmica y procedimiento de normalización (Chau 2006)	Un modelo de clasificación y muestras difuso e índice numérico que muestra el grado de contaminación. Monitoreo futuro de impacto ambiental para las condiciones de AMD.	Principalmente orientado hacia vertederos de desechos ricos en sulfuros.

**Tabla 4.** Predicción de los parámetros de calidad del agua y funciones de pertenencia para diferentes objetivos.

<b>Modelo/ Técnica</b>	<b>Resultado</b>	<b>Falencias</b>
Modelo de predicción que usa MLP ANN para TDS (total de solidos disueltos) EC (conductividad eléctrica) y turbulencia (Aroba et al. 2007)	Modelo predictivo con un error en la media del 10%. Correlación significativa de TDS con CE y turbidez con SS.	Depende mucho de conjunto de datos precisos y no duplicados.
GFWLAM (Lackawanna, Bruns, and Sweet 2004) y de dos fases GFWLAM (Karmakar and Mujumdar 2006) con el concepto de la programación de intervalo, programación inexacta, programación lineal difusa, técnica difusa de optimización multi-objetivo y el operador max-min.	Las comparaciones entre el caso incierto determinista con números grises indican intervalos existencia de múltiples soluciones.	Dependen solamente de la variable DO.

**Tabla 5.** Preservación de los Ecosistemas de Cuerpos de Agua.

<b>Modelo/ Técnica</b>	<b>Resultado</b>	<b>Falencias</b>
Modelos 3DM, MSM y M2M aplicando modelado MLP ANN (Karmakar and Mujumdar 2007) y (Paul Conrads 2014)	3DM predice el cambio en el nivel del agua y la salinidad. Predicciones MSM generan un mapa de contorno 2D de conductancia especifica en el área de estudio	Gran carga computacional al implementar 3 modelos.
Dos etapas de modelado ANN con sub modelos diarios y horarios (Paul Conrads 2014)	Modelos diarios simulando picos SC (conductancia especifica) afilados y modelos hora simulando señales de frecuencia SC proporcionando la gama de coeficientes determinísticos.	Detalles del modelado ANN no disponibles.

Modelo ANN de "Ganancia" (Roehl et al. 2006)	Las relaciones entre las VM para la comprensión cualitativa de los estuarios complejos.	El uso de sustituto del nivel del agua en los frentes de sal en lugar de la medición real
Herramienta GIS con análisis estático. (Lackawanna, Bruns, and Sweet 2004)	Ranking y clasificación de cuencas en función de sus características para cubiertas de minería y terrenos forestales. Predicciones de SIG sobre los beneficios de la reforestación.	Enfoque estadístico exige conjunto de datos preciso para un resultado confiable.

## Características del Río

La cuenca del río Bogotá, tiene un área de drenaje de 5,907 Km<sup>2</sup>, pasa por el departamento de Cundinamarca en sentido noreste – sureste, desde su nacimiento en el municipio de Villapinzón a 3.300 m.s.n.m., hasta su desembocadura en el río Magdalena en el municipio de Girardot a una altura 600 m.s.n.m.(CAR 2014)

La cuenca del río Bogotá se encuentra dividida tres tramos:

- La cuenca Alta desde su nacimiento hasta el norte de la zona urbana del Distrito Capital, con una longitud de 165 km.
- La Cuenca Media, desde el inicio de la zona urbana de Bogotá hasta el Salto de Tequendama, con una longitud de 90 km.
- La Cuenca Baja, aguas abajo del salto de Tequendama hasta su desembocadura en el río Magdalena, con una longitud de 55 km. (Subdirecci, Ambiental, and Rep 2011)

En las siguientes imágenes se muestra la distribución geográfica del río Bogotá con el nombre de los 81 puntos de monitoreo utilizados por la CAR.

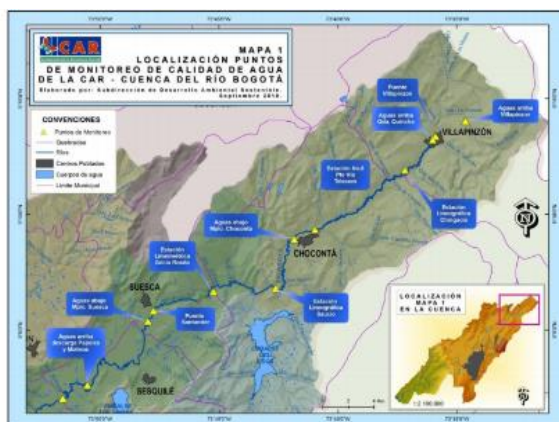


Figura 1. primeros 11 puntos de monitoreo de la cuenca alta.(Subdirecci, Ambiental, and Rep 2011)

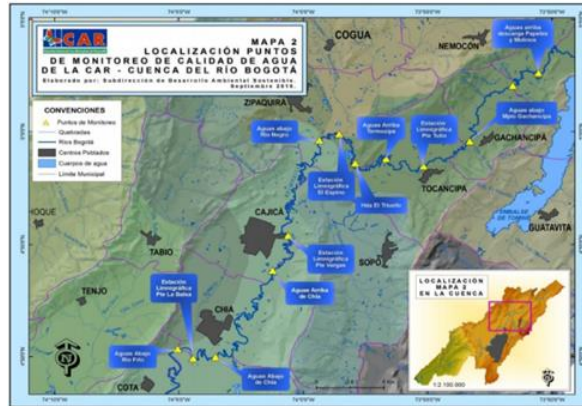


Figura 2. siguientes 12 puntos de monitoreo de la cuenca alta (Subdirecci, Ambiental, and Rep 2011).

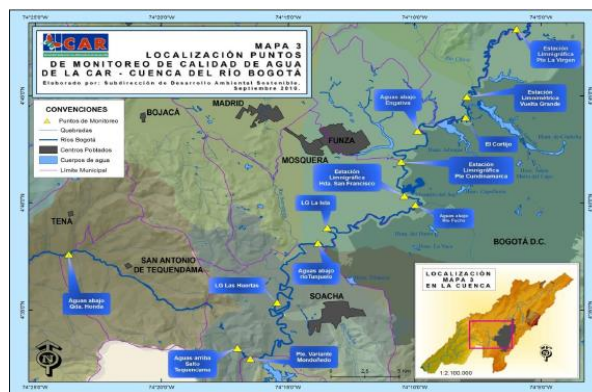


Figura 3. Localización Puntos de Monitoreo río Bogotá – Cuenca Media (Subdirecci, Ambiental, and Rep 2011)

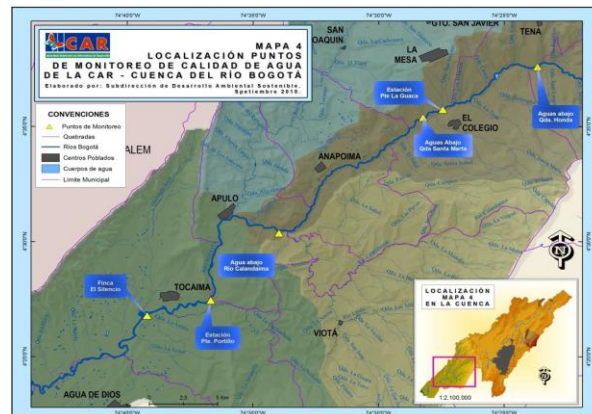


Figura 4. primeros 5 puntos de monitoreo de la cuenca baja. (Subdirecci, Ambiental, and Rep 2011)



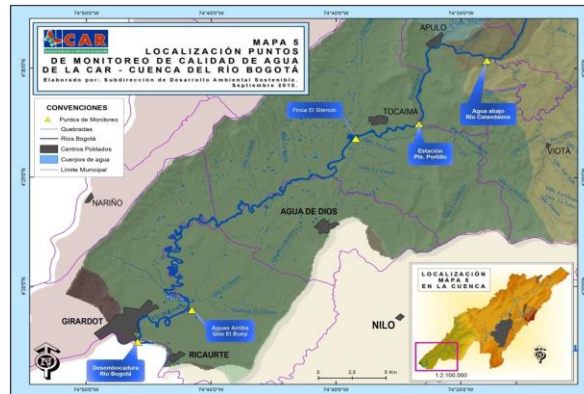


Figura 5. Siguiendo 5 puntos de monitoreo de la cuenca baja. (Subdirección, Ambiental, and Rep 2011)

### Decreto 1594 De 1984 Por El Cual Se Reglamentan Los Usos Del Agua Y Residuos Líquidos

Para el análisis de los datos se tomó como base la normativa planteada en este decreto por el cual se reglamentan los usos del agua y residuos líquidos en el territorio colombiano, en su capítulo IV. De Los Criterios De Calidad Para Destinación Del Recurso se relacionan los niveles máximos para el uso del agua en afluentes hídricos tales como consumo humano, uso pecuario, conservación de flora y fauna, uso agrícola y recreativo, a continuación, se relaciona la definición de cada uso y sus valores máximos.

Según el capítulo 3 del decreto 1594 se definen los siguientes artículos (Agricultura 1984):

ARTICULO 30. Se entiende por uso del agua para consumo humano y doméstico su empleo en actividades tales como:

- a. Fabricación o procesamiento de alimentos en general y en especial los destinados a su comercialización o distribución.
- b. Bebida directa y preparación de alimentos para consumo inmediato.
- c. Satisfacción de necesidades domésticas, individuales o colectivas, tales como higiene personal y limpieza de elementos, materiales o utensilios.
- d. Fabricación o procesamiento de drogas, medicamentos, cosméticos, aditivos y productos similares.

ARTICULO 31. Se entiende por uso del agua para preservación de flora y fauna,

su empleo en actividades destinadas a mantener la vida natural de los ecosistemas acuáticos y terrestres y de sus ecosistemas asociados, sin causar alteraciones sensibles en ellos, o para actividades que permitan la reproducción, supervivencia, crecimiento, extracción y aprovechamiento de especies hidrobiológicas en cualquiera de sus formas, tal como en los casos de pesca y acuicultura.

ARTICULO 32. Se entiende por uso agrícola del agua, su empleo para irrigación de cultivos y otras actividades conexas o complementarias, que el Ministerio de

Salud o la EMAR establezcan.

ARTICULO 33. Se entiende por uso pecuario del agua, su empleo para el consumo del ganado en sus diferentes especies y demás animales, así como para otras actividades conexas y complementarias que el Ministerio de Salud o la EMAR establezcan.

ARTICULO 34. Se entiende por uso del agua para fines recreativos, su utilización, cuando se produce:

- a. Contacto primario, como en la natación y el buceo.
- b. Contacto secundario, como en los deportes náuticos y la pesca.

Con estas definiciones se presentan los valores máximos permitidos para cada uso (ver Apéndice)

### Desarrollo del Modelo

Se desarrolló un modelo de clúster que permite reducir la cantidad de datos de manera que se analicen para un objetivo en particular. La idea es tratar los datos como una gran matriz de la cual se desglosan varias sub-matrices las cuales nos servirán para generar vectores que se analizarán con métodos numéricos y algoritmos para obtener una estimación del uso del agua según la cuenca y según el periodo analizado.

### Características del Dataset y Limpieza de Datos

Se parte del hecho de que los datos ya se convirtieron a un formato Excel (xlsx o xls) o CSV, en particular para el set de datos que se maneja los datos originales vienen en formato PDF y mediante herramientas de gestión de bases de datos se convierten a formato SQL y luego a CSV.

A continuación, se muestra es aspecto del archivo Excel que se manejara como formato de origen:

id	periodo	punto_monitoreo	variable	unidad	metodo_analisis	resultado_concentraci
14681	2009-01	AGUAS ARRIBA VILLAP CAUDAL		LPS		179
14682	2009-01	AGUAS ARRIBA VILLAP CLORUROS		MG CL- / L	ARGENTOMETRICO (4	1
14683	2009-01	AGUAS ARRIBA VILLAP CONDUCTIVIDAD	PS / CM		ELECTROMÉTRICO (2	101
14684	2009-01	AGUAS ARRIBA VILLAP DBO		MG O2 / L	PRUEBA DE 5 DÍAS (5;VMD	
14685	2009-01	AGUAS ARRIBA VILLAP DQO (REFLUJO A	MG O2 / L		REFLUJO ABIERTO (5;39.4	
14686	2009-01	AGUAS ARRIBA VILLAP DQO SOLUBLE		MG O2 / L	REFLUJO ABIERTO (5;2	12
14687	2009-01	AGUAS ARRIBA VILLAP DUREZA TOTAL		MG CACO3 / CÁLCULO (2340 B)		8.93

Figura 6. Formato original archivo .xlsx.

Es necesario para este caso eliminar algunas columnas que no nos aportan información para el caso de estudio particular que es determinar los usos potenciales del agua. Se eliminan entonces las columnas id, unidad y método de análisis y para poder procesar los datos es necesario cambiar el formato del periodo ya que el guion no es un carácter que nos permita realizar cálculos numéricos sobre este campo.

Finalmente, el archivo de entrada tendrá el siguiente aspecto:

periodo	punto_monitoreo	variable	resultado_concentracion
200801	AGUAS ARRIBA VILLAPINZÓN	SULFATOS*	2,8
200801	PUENTE VILLAPINZÓN	SULFATOS*	2,9
200801	AGUAS ARRIBA Q. QUINCHA	SULFATOS*	3,2
200801	Q. QUINCHA	SULFATOS*	0
200801	ESTACIÓN LM CHINGACIO	SULFATOS*	25,2
200801	AGREGADOS CHOCONTÁ	SULFATOS*	20,3
200801	RÍO TEJAR	SULFATOS*	3,3

Figura 7. Formato de entrada archivo xlsx.

Para este caso se encontraron problemas con los datos numéricos de la columna “Resultado\_Concentración” ya que muchos de los valores registrados no son numéricos debido a que corresponden a valores límite que dependen del método de análisis en la toma de los datos, valores como VMD (Valor Mínimo Detectable), LCM (Limite de Cuantificación), NR (No representativo), LCM (Limite de Cuantificación), ente otros. La forma de reemplazar estos datos fue basarse en los reportes de resultados de las variables presentes en los cuerpos de agua del Rio Bogotá publicado por la CAR en el 2013 y 2014 (Chorotega, Fax, and Brunca 2015).

De esta manera se obtienen todos números decimales que pueden ser fácilmente manipulados por una herramienta de cálculo, aunque para este caso se descartaron los datos del segundo periodo del año 2008 ya que presentaban demasiados datos similares que no tenían mucha información que aportar.

Finalmente, para facilitar los cálculos y optimizar el tiempo de procesamiento se eliminaron las variables que no estaban incluidas en el decreto 1594 de 1984 por el cual se regula los usos del agua en vertimientos superficiales. Este paso no es estrictamente necesario pues con el dataset completo funciona también el método.

### Definición del Modelo

Para la implementación del modelo se seleccionó el software Matlab ® por su versatilidad y capacidad de cálculo para grandes volúmenes de datos.

La idea es manejar las tablas del archivo de Excel como una gran matriz de 4 columnas por una gran cantidad de filas (para este caso 33881 columnas).

De esta matriz partiremos a clusterizar por el periodo a analizar según se seleccione con el siguiente código:

```
function [P_CHDes, P_CHConv, P_UAgric, P_Uff, P_Upec, P_Urec2, P_Urec1] =
usosdelagua(periodo,punto_monitoreo,datos)
```

%Periodo: periodo de tiempo en donde se va a realizar el análisis

%punto\_monitoreo: Cuenca del rio en donde se van a analizar los datos

%datos: nombre del archivo Excel en donde se encuentran los datos.

```
for j=1:size(numericData-1)
```

```
    if all(data{j,1})== periodo,
```

```
        base(j,1)=all(data{j,1});
```

```
        base(j,2)=all(data{j,2});
```

```

    base(j,3)=alldata(j,3);
    base(j,4)=alldata(j,4);
end
end

```

A partir de este análisis obtenemos una matriz más pequeña en donde se tienen los valores de todas las variables medidas en el periodo seleccionado. A continuación se genera un vector con el punto de monitoreo que relaciona los valores de las variables para este punto:

```

for i=1:size(numbase)-1
a(i)=strcmp(base(i,2),punto_monitoreo);
if a(i)==1,
ba2= numbase{i};
punto(i)= ba2;
else
punto(i==0) = [];
end
end

```

Finalmente, para cada variable analizada se crea una clase la cual acumula un vector que incluye el cruce del punto de monitoreo y el periodo del año, un ejemplo de la codificación de esta clase es:

```

coli=zeros(size(punto));
b(i)=strcmp(base(i,3),'COLIFORMES TOTALES');
if b(i)==1,
coli(i)=punto(i);
else
coli(i==0) = [];
end

```

Teniendo ya contruidos los vectores para analizar los datos, se toman los valores máximos de cada vector para cada variable y se comparan con los valores consignados en el decreto para determinar qué porcentaje de variables cumplen con la normativa y de esta manera reflejar una estimación de los usos del agua.

## RESULTADOS

Como resultado tenemos una función de Matlab que tiene como parámetros de entrada el periodo de tiempo a analizar, la cuenca del rio sobre la que se va a realizar el análisis y el dataset seleccionado. Y como salidas tendremos cada uno de los usos representado en

porcentaje dentro de una variable (P\_CHDes, P\_CHConv, P\_UAgric, P\_Uff, P\_Upec, P\_Urec2, P\_Urec1).

La función tiene el siguiente aspecto:

```
[P_CHDes, P_CHConv, P_UAgric, P_Uff, P_Upec, P_Urec2, P_Urec1] =
usosdelagua(201501,'DESEMBOCADURA RÍO
BOGOTÁ','VALORES_RIO_BOGOTA_MATLAB.xlsx').
```

Para este estudio se analizaron en el tiempo los puntos más bajos de cada cuenca en donde se supone que se encuentra la concentración más alta de elementos para el análisis de calidad y uso, para este caso se toman los datos de cada una de las cuencas en los siguientes puntos:

- a) Cuenca Alta: Estación LG – Puente La virgen
- b) Cuenca Media: Estación Puente La Guaca
- c) Cuenca Baja: Desembocadura del Rio Bogota

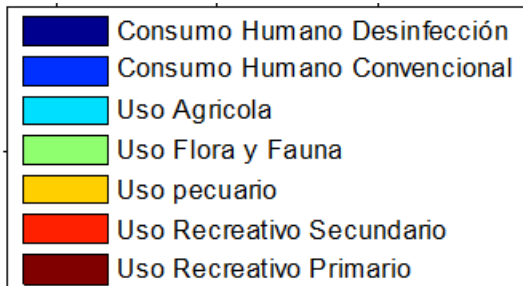


Figura 8. Convenciones de los diagramas

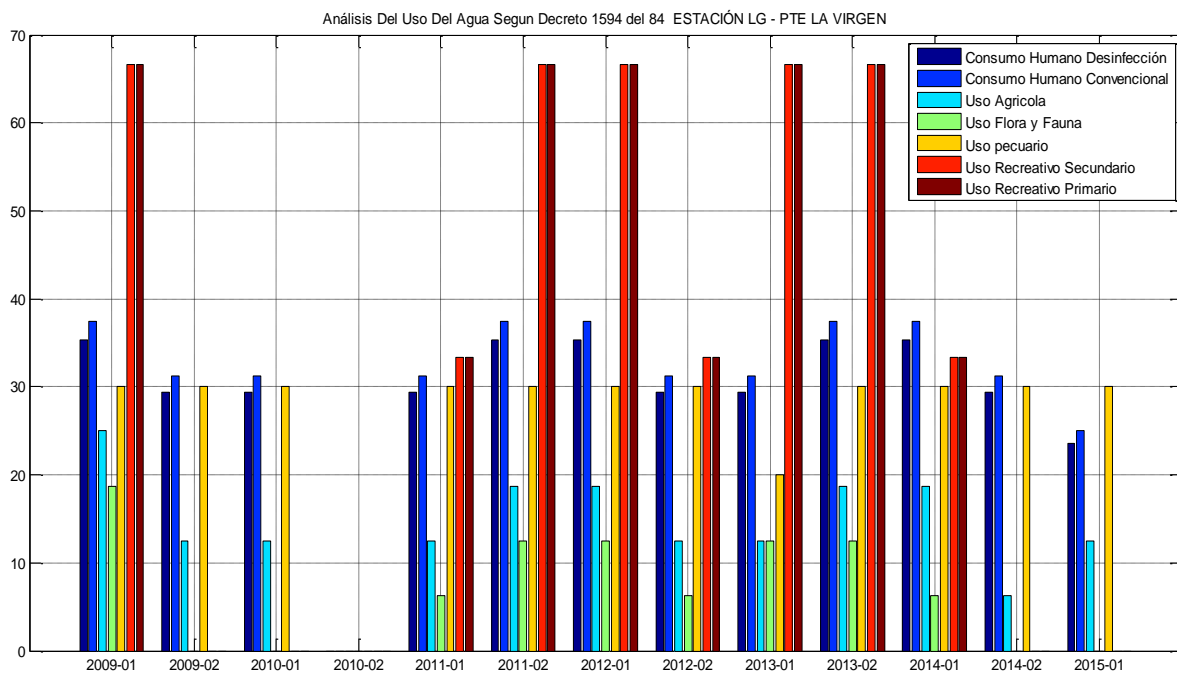


Figura 9. Diagrama de Barras análisis temporal Cuenca Alta.

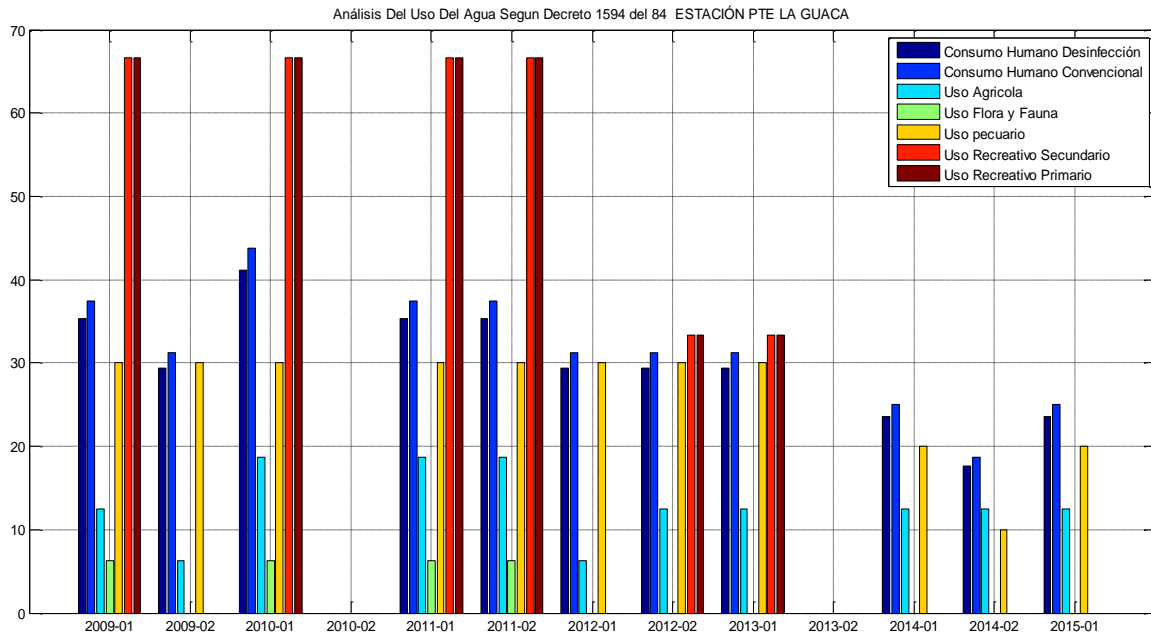


Figura 9. Diagrama de Barras análisis temporal Cuenca Media.

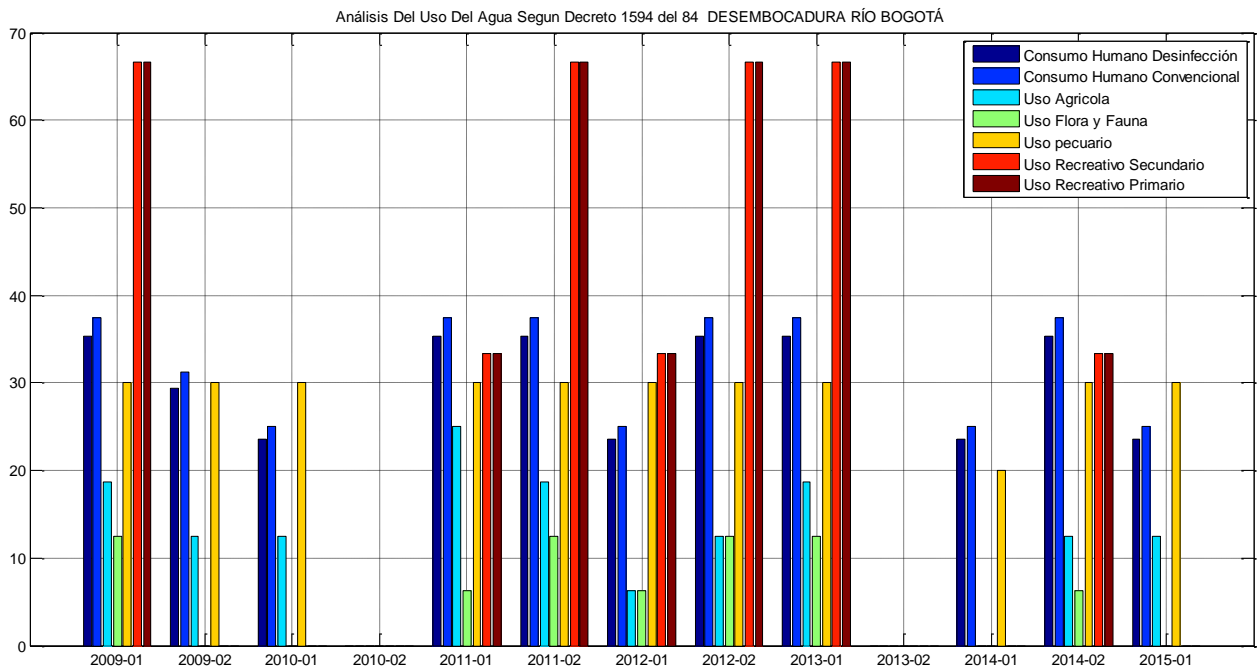


Figura 10. Diagrama de Barras análisis temporal Cuenca Baja.

## CONCLUSIONES

Como se puede ver en el análisis de las gráficas se puede afirmar que a partir del año 2014 según los tipos de usos de agua reflejados indican que la calidad del agua mejoró ya que se presenta menor porcentaje de usos recreativos y se incrementan los porcentajes de usos como el potable o el pecuario. Esto probablemente se deba a las actividades de recuperación del río que ha venido implementando la CAR en los últimos años.

Por la cantidad de datos y de cálculos que realiza la aplicación el análisis de los datos en cada punto de monitoreo tiene una duración de 3 minutos aproximadamente, esto quiere decir que se tiene un costo computacional alto que requiere de una máquina potente para poder analizar este volumen de datos, probablemente implementando este algoritmo en una herramienta de cómputo más ágil que Matlab pueda mejorarse el desempeño.

Realmente el decreto 1594 De 1984 Por El Cual Se Reglamentan Los Usos Del Agua Y Residuos Líquidos no da buenas pautas para poder predecir o analizar los usos del agua, ya que resultados como que se tengan porcentajes de usos recreativos similares desde la cuenca alta hasta la baja no resulta lógico y no brinda mucha información, esto llevaría a replantear la forma en que se regulan los usos del agua, bien sea considerando variables adicionales a las que se utilizaron o complementando la normativa con estudios adicionales a los que se han desarrollado.

Tampoco es lógico que en los últimos años en donde se ve una mejora en los indicadores se tenga un porcentaje más alto de uso recreativo en la cuenca alta que en la baja ya que generalmente en donde más se contamina el agua es en su paso por Bogotá por ser ciudad capital con mayor desarrollo industrial. Sin embargo es una buena aproximación a un modelo de estimación para proponer mejoras sobre la normatividad y las mediciones en campo que se hacen actualmente.

## REFERENCIAS BIBLIOGRÁFICAS

Agricultura, ministerio de. 1984. "Decreto 1594." Ley 9 de 1979 - Ley 2811 1974 1984(Junio 26): 55.

Aroba, J et al. 2007. "Application of Fuzzy Logic and Data Mining Techniques as Tools for Qualitative Interpretation of Acid Mine Drainage Processes." *Environmental Geology* 53(1): 135–45. <http://dx.doi.org/10.1007/s00254-006-0627-0>.

CAR. 2014. BOLETIN DE CALIDAD HÍDRICA 2014.

Castro, Mario, Juniel Almeida, Julio Ferrer, and Daissy Díaz. 2014. "Indicadores de La Calidad Del Agua : Evolución Y Tendencias a Nivel Global."

Chau, Kwok Wing. 2006. "A Review on Integration of Artificial Intelligence into Water Quality Modelling." *Marine Pollution Bulletin* 52(7): 726–33.

Chorotega, Región, T E L Fax, and Región Brunca. 2015. "Reporte de Resultados." (37): 25871760.

Conrads, Paul A, Edwin A Roehl, and John B Cook. 2002. "ESTIMATING POINT-SOURCE IMPACTS ON THE BEAUFORT RIVER USING ARTIFICIAL NEURAL NETWORK MODELS." Water Resources (803): 1–6.

Conrads, Paul A, Edwin A Roehl, William Martello, and Edward Saxon. "The Role of Artificial Neural Network Models in Developing a Regional Wastewater Reclamation Facility for." 303(d).

Dutta, Prasun, and Rituparna Chaki. 2012. "A Survey of Data Mining Applications in Water Quality Management." Proceedings of the CUBE International Information Technology Conference on - CUBE '12: 470. <http://dl.acm.org/citation.cfm?id=2381716.2381805>.

Edwin A. Roehl, Jr., John B. Cook, PE, Greenville, SC. 2013. "EFFECTS OF NONPOINT SOURCE MARSH LOADING ON COMPLEX ESTUARIES." The effects of brief mindfulness intervention on acute pain experience: An examination of individual difference 1(September 2013): 1–36.

Hammond, Allen et al. 1995. 36 Environmental Indicators: A Systematic Approach to Measuring and Reporting on Environmental Policy Performance in the Context of Sustainable Development. [http://pdf.wri.org/environmentalindicators\\_bw.pdf](http://pdf.wri.org/environmentalindicators_bw.pdf).

Karimipour, F. 2005. "Water Quality Management Using GIS Data Mining." Journal of Environmental Informatics 5(2): 61–71.

Karmakar, Subhankar, and P. P. Mujumdar. 2006. "Grey Fuzzy Optimization Model for Water Quality Management of a River System." Advances in Water Resources 29(7): 1088–1105.

Karmakar, Subhankar, and P. P. Mujumdar. 2007. "A Two-Phase Grey Fuzzy Optimization Approach for Water Quality Management of a River System." Advances in Water Resources 30(5): 1218–35.

Lackawanna, Upper Susquehanna, Dale Bruns, and Tom Sweet. 2004. "Geospatial Tools to Support Watershed Environmental Monitoring and Reclamation: Overview of Presentation." Atlantic.



Lobbrecht, a H, Y B Dibike, and D P Solomatine. 2002. "Applications of Neural Networks and Fuzzy Logic to Integrated Water Management Project Report."

"Neighborhood Analysis in Water Pollution Estimation." 2015. (October).

Oecd. 2004. "OECD Key Environmental Indicators." : 38.

Orjuela, Luz, Javier Saldarriaga, Martha Garcia, and Hernando Wilches. 2010. "Informe de Calidad Superficial Del Agua En Colombia." : 1–52. [https://www.siac.gov.co/documentos/DOC\\_Portal/DOC\\_Agua/20120928\\_Cali\\_agua\\_sup\\_ENA\\_2010.pdf](https://www.siac.gov.co/documentos/DOC_Portal/DOC_Agua/20120928_Cali_agua_sup_ENA_2010.pdf).

Paul Conrads, Edwin A. Roehl Jr. 2014. "Transforming Large Databases Into Critical Knowledge Using Data Mining– Three Case Studies in South Carolina and Georgia." *Psychological Science* 25(9): 1682–90. <http://www.ncbi.nlm.nih.gov/pubmed/15003161>.

Prato, Tony. 2005. "A Fuzzy Logic Approach for Evaluating Ecosystem Sustainability." *Ecological Modelling* 187(2–3): 361–68.

Roehl, Edwin A et al. 2006. "Integrating 3D Hydrodynamic Transport and Ecological Plant Models of the Savannah River Estuary Using Artificial Neural Network Models." *Transport*: 1–8.

Subdirecci, C A R, Desarrollo Ambiental, and Sostenible Rep. 2011. "Boletín de Calidad de Las Cuencas de La Jurisdicción Car 2011 Bogotá, D.c." (36).

Survey, U S Geological, and Gracern Road. 2003. "DEVELOPMENT OF AN EMPIRICAL MODEL OF A COMPLEX, TIDALLY AFFECTED RIVER USING ARTIFICIAL NEURAL NETWORKS."

V, Raman Bai, Reinier Bouwmeester, and S Mohan. 2009. "Fuzzy Logic Water Quality Index and Importance of Water Quality Parameters." : 51–59.