

ANÁLISIS GEOESTADÍSTICO Y PREDICCIÓN ESPACIAL PARA EL ÍNDICE DE CALIDAD DEL AGUA EN EL RÍO BOGOTÁ

Cindy Daniela Mesa Wilchez, Estudiante ICG, Andrés David Mesa Ospina, Estudiante ICG, Carlos Enrique Montenegro Marín, Paulo Alonso Gaona Garcia.

Universidad Distrital Francisco José de Caldas, Bogotá D.C, Colombia, cdmesaw@correo.udistrital.edu.co

Universidad Distrital Francisco José de Caldas, Bogotá D.C, Colombia, admesao@correo.udistrital.edu.co

Universidad Distrital Francisco José de Caldas, Bogotá D.C, Colombia, cemontenegrom@udistrital.edu.co

Universidad Distrital Francisco José de Caldas, Bogotá D.C, Colombia, pagaonag@udistrital.edu.co

RESUMEN– la finalidad de este artículo es dar a conocer la implementación de conceptos geoestadísticos en la cuenca del río Bogotá, mediante el conocimiento de un conjunto de datos capturados en 59 estaciones localizadas a lo largo de todo el afluente del río Bogotá. Cada una de las estaciones de monitoreo mide la concentración de 5 o 6 variables ambientales con las cuales posteriormente se realiza el cálculo del ICA (Índice de calidad del agua) para cada una de las estaciones. Finalmente mediante estos valores se realiza el análisis geoestadístico para generar la predicción del ICA de manera continua sobre todo el río. Para lo anterior se emplean diferentes métodos de interpolación y se generan los mapas correspondientes para la visualización de la variabilidad de la calidad del agua sobre todo el río Bogotá determinando posibles áreas con bajos, medios y altos índices de calidad de agua.

Palabras Claves: ICA, Geo-estadística, Interpolación, predicción.

I. INTRODUCCIÓN

Es evidente para todos los habitantes de la ciudad de Bogotá y sus alrededores que el Río Bogotá, presenta serios problemas de contaminación. Esta problemática no se puede dejar a un lado, dado que el río Bogotá es la cuenca principal del sistema hídrico de la ciudad y que por políticas externas y sobre todo por una pésima conciencia social, en el año 2017 la capital de Colombia cuenta con el río más contaminado de todo el país, siendo el botadero de industrias, empresas y ciudadanos del común. Por lo anterior desde hace varias décadas, durante las administraciones públicas de la ciudad, se han planteado un sinnúmero de proyectos dirigidos a la intervención del río, con el fin de frenar y mitigar la contaminación presente en este.

Actualmente se cuenta con diferentes estudios que determinan cual es el nivel de contaminación que se presenta, tanto en el agua como en el aire de la ronda del Río Bogotá. El IDEAM (Instituto de hidrología, Meteorología y estudios ambientales) reglamenta el índice ICA (Índice de Calidad del Agua) el cual proporciona valores para determinar cuál es el nivel de contaminación del agua en un punto específico. Este Índice es el cálculo resultante de la toma de variables como Oxígeno disuelto, Sólidos suspendidos, Demanda química del oxígeno, nitrógeno, conductividad eléctrica y PH. Este conjunto de variables nos permite obtener un valor (ICA) a través de métodos de minería de datos y modelos de regresión lineal, permitiendo clasificar la calidad del agua en cada punto de

investigación como Muy mala, mala, regular, aceptable o buena.

El cálculo ICA para este caso se desarrolló para un total de 56 estaciones a lo largo del río, y es aquí donde la geoestadística entra a jugar un valor importante para un análisis más detallado de su comportamiento, ya que esta nos permite describir la continuidad espacial de cualquier fenómeno, permitiendo de esta manera la observación por medio de la variabilidad de la calidad del agua a nivel de contaminación de manera continua sobre todo el río y no solamente en las 59 estaciones de monitorio .

II. MARCO TEÓRICO

A. Índice de Calidad del Agua-ICA

El ICA es un valor numérico determinado entre 0-1 el cual califica en cinco categorías la calidad del agua de una corriente superficial. Las categorías son: Muy mala, mala, regular, aceptable o buena. Las anteriores categorías dependen del valor calculado a partir de cinco o seis variables capturadas en una estación de monitoreo durante un tiempo específico.

B. Análisis Geoestadístico

La geoestadística estudia aquellos fenómenos que varían de forma continua sobre una región delimitada de estudio. Los Análisis geoestadísticos comienzan con un análisis estructural que corresponde al análisis del variograma, que permite obtener un variograma teórico (ajuste de los valores de varianza o covarianza en una función específica) el cual es utilizado para la interpolación de la variable a analizar, generando así finalmente un mapa de predicción utilizado para el análisis y la toma de decisiones.

C. Interpolación espacial

La interpolación es el proceso por el cual se generan predicciones en puntos donde originalmente no se fueron capturados ningún tipo de información, Como resultado de esto se obtienen mapas de pronóstico que permiten visualizar el comportamiento de una variable espacial en una zona de estudio [1]

III. METODOLOGÍA

Para este caso en particular se analizó la información referente a la determinación del índice de ICA brindada por el

Análisis geoestadístico y predicción espacial para el índice de calidad del agua en el río Bogotá

grupo de investigación GIIRA de la Universidad Distrital Francisco José de Caldas. La base de datos comprende valores determinados por el grupo de investigación QUINOA como apoyo para el desarrollo de minería de datos. Para realizar este estudio se filtró la información, por lo tanto en este caso la base de datos seleccionada consta de 59 registros promediados, para los cuales se presenta la siguiente información: ID, Período, Nro. Estación, Punto monitoreo, Cuenca, variable, unidad, Método analizado, Conductividad, DQO, Oxígeno disuelto, PH, Sólidos suspendidos, Temperatura agua, ICA, Nortes y Estes respectivamente [2]. En donde:

- pH: Hace referencia al nivel de acidez o alcalinidad presentada por el agua, en las diferentes estaciones analizadas, cabe destacar que esta variable se mide a través de procedimientos electrométricos.
- Sólidos suspendidos: Hace referencia a la cantidad de sólidos encontrados en la superficie del agua. Este valor se calcula mediante el método de secado a 180° (CAR) y presenta unidades de mg/l. Se entiende principalmente como una variable que indica contaminación.
- Conductividad: Esta variable se caracteriza por reflejar las propiedades presentes del agua para conducir la corriente eléctrica, esto depende no solamente del tipo de materiales disueltos en la zona, sino también de los niveles de salinidad del agua y cantidad de materia vegetal. Esta se encuentra en unidades de $\mu\text{S}/\text{cm}$. Es posible calcular este valor por el método electrométrico (2510 B).
- DQO o demanda química de oxígeno: Es posible definirla como la cantidad de oxígeno requerido para oxidar la materia orgánica de la muestra de agua residual, esta se determinó mediante la metodología de reflujo abierto y se presenta en (Mg O₂/l).
- Oxígeno disuelto: Hace referencia a la cantidad de oxígeno disuelto en el agua, este también se presenta como un indicador de contaminación y soporta la cantidad de vida tanto animal como vegetal presente en la zona. Se determina mediante el método de Modificación ácida y presenta unidades de (mg O₂/l).
- Temperatura del agua: Esta se presenta en unidades de °C en la cual se determinan características intrínsecas del ecosistema en el que se encuentra, ya que la temperatura es la que define el equilibrio térmico del medio.

Por otro lado referente a información espacial necesaria para la espacialización e interacción con la base de datos, fue necesario la descargar de un conjunto de shapefile del IDECA (Infraestructura de datos espaciales de la ciudad de Bogotá), en el cual se encuentra catalogada toda la información espacial referente a la ciudad de Bogotá.

Aparte de esta información espacial fue necesario georreferenciar los puntos de investigación es decir, las estaciones donde fueron capturados los datos, con el fin de poder generar una salida gráfica en donde se visualizaran estas estaciones. Por último se generó un buffer o área de influencia

de aplicación en donde se pudiera visualizar el desarrollo de los métodos de interpolación (Kriging, IDW, RBF).

A. Análisis Geo-estadístico

Primero que todo, para el desarrollo de este análisis se implementó el software RStudio, en el cual es posible desarrollar cada una de las verificaciones de supuestos y espacialización de los datos, aunque cabe resaltar que alguna de la información base como los shapefile, fue manejada en ArcGis 10.3 en el cual se expresa la mayoría de información cartográfica para este análisis.

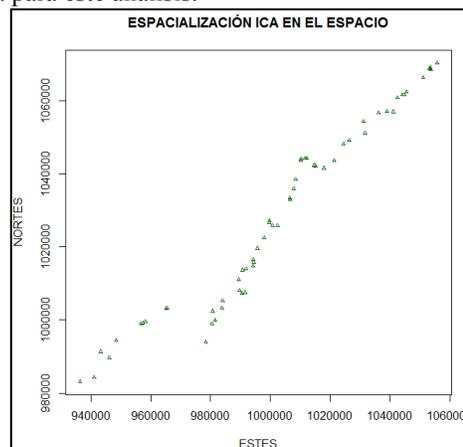


Fig. 1 Espacialización de los datos

En la Fig. 1 es posible visualizar la espacialización de los datos, en el cual se observa la distribución de las 59 estaciones a lo largo del río Bogotá, las cuales casan perfectamente con respecto al shapefile del mismo, dando a entender que la captura de los datos se hizo de manera correcta además de la debida proyección respecto al sistema de referencia a manejar, es importante tener en cuenta la orientación de la misma (Norte, Estes).

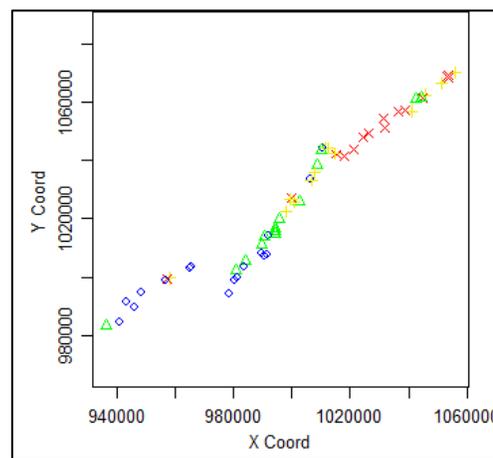


Fig. 2 Espacialización de los datos-GEODATO

En la Fig. 2 se encuentra representado el valor del ICA en cada una de las estaciones. Los colores representan la distribución de los datos, es decir en que categoría se encuentra el valor calculado del ICA. Se identifican presencia de datos medios localizados principalmente en la zona central del afluente del río en tonalidades de verdes, mientras que en los

valores de los extremos el caso es inversamente proporcional, es decir en la zona sur-oeste se acumulan datos bajos aunque hay algunos datos atípicos respecto a la localización, mientras que en la zona nor-este es posible notar predominancia de valores altos del ICA aunque hay presencia de uno pocos valores neutrales (variación degradada).

Por otro lado, también permite analizar la distribución espacial de los datos del ICA con respecto a su espacialización, dando la posibilidad de analizar la presencia de tendencia de la muestra respecto a la zona, con respecto a variables netamente espaciales. En este caso se identifica en los gráficos de la fig. 3 que existe tendencia espacial en los datos, en los cuales prima una tendencia generalizada respecto a las coordenadas x , mientras que en las coordenadas y se podría decir que esta tendencia no existe, por lo cual a continuación se procede a realizar la respectiva corrección de la tendencia y a generar análisis de anisotropía en los datos. Se entiende como tendencia cuando los datos de interés tienden a comportarse de cierta forma. Por ejemplo en la Fig. 3 se observa que a medida que la distancia en la coordenada x aumenta el valor de los datos tiende a aumentar.

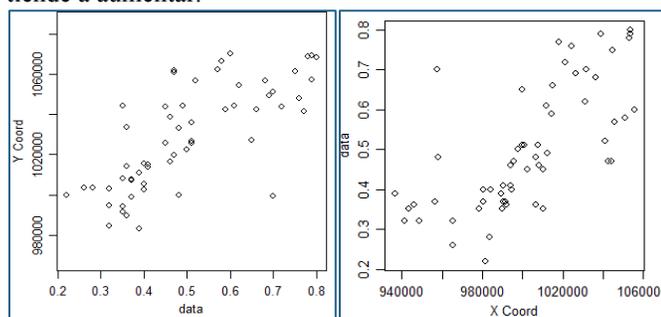


Fig. 3 Distribución espacial de los datos

La existencia de tendencia presenta problemas en la estimación de las predicciones, debido a la existencia de valores dependientes, por lo que es importante previamente generar un análisis para determinar si es necesario remover tendencia e inclusive analizar si se presenta anisotropía en la información. Se entiende por anisotropía como la invarianza que se presenta en los datos al aplicar rotaciones, se pueden presentar dos tipos de anisotropía: geométrica y zonal. Primero que todo se procede a evaluar la anisotropía y determinar el tipo de anisotropía que se presenta y así identificar el tipo de corrección que se debe aplicar al conjunto de datos.

Primero se realiza el análisis de la anisotropía zonal, la cual expresa que el comportamiento del modelo de semivariograma para los datos presenta un valor constante en su rango (a) y una variación en la meseta (c_0, c_1, \dots, c_n), expresando de esta manera que el ajuste va a generar algunos cambios significativos en esta zona, al presentar estos problemas es posible generar su corrección a partir de un proceso de rotación en donde se genera un desplazamiento angular [1]. Para este caso el modelo indica que no hay existencia de anisotropía zonal en los datos, como se puede observar en la tabla 1, por lo que no será necesario aplicar una rotación de la información.

\$R	1
\$Grado del ángulo theta	0

\$Q	0 0 0
\$Dirección de rotación	FALSE

Tabla 1. Anisotropía Local

Por otro lado también es de vital importancia verificar que no existan problemas de anisotropía geométrica, la cual se comporta de manera inversa con respecto a la anisotropía zonal, es decir que presenta un valor constante en la meseta (c), mientras que el rango presenta valores variables (a_1, a_2, \dots, a_n) [1]. Para este caso podemos visualizar a continuación el comportamiento del variograma en diferentes direcciones del río con el ICA transformado ($0^\circ, 20^\circ, 160^\circ, 180^\circ, 200^\circ, 340^\circ$).

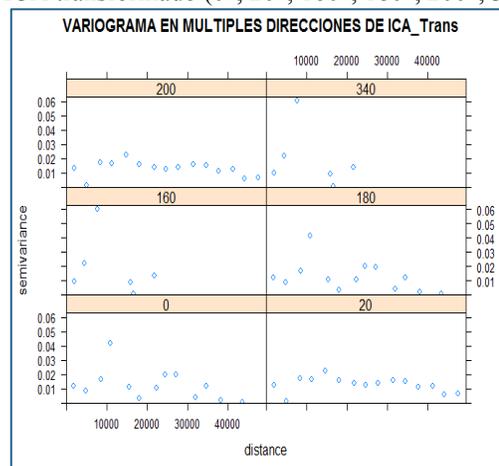


Fig 4. Variograma en múltiples dimensiones- ICA Transformado

Finalmente al aplicar un modelo de regresión lineal para el ICA con respecto a las coordenadas del río (NORTES y ESTES) de la base de datos seleccionada, es posible identificar una mejora entre cada una de las diferenciaciones mencionadas anteriormente, con lo cual se concluye que existe anisotropía geométrica, además que visualmente podemos observar que el comportamiento del variograma no es contante en todas las direcciones. La anisotropía geométrica se puede corregir a través de la aplicación de un modelo de regresión lineal, generando así las variaciones presentadas anteriormente.

Con la finalidad de eliminar la tendencia es necesario analizar la existencia de normalidad en los datos, para ello se aplica el test shapiro-wilk, dando como resultado un p-value de 0.00439 lo cual indica que no existe evidencia suficiente para rechazar la hipótesis nula, es decir que la variable ICA no presenta normalidad en los datos, esto contribuye de la misma manera a problemas en la distribución de los datos.

Por lo que se debe realizar el modelo de regresión en el cual se anexan las variables de Estes y Nortes respectivamente, obteniendo como resultado un modelo donde la variable Este no es significativa estadísticamente, por lo tanto se debe remover esta variable del modelo y finalmente trabajar con un modelo de la siguiente forma funcional:

```

Call:
lm(formula = Proyecto_ICA$ICA ~ Proyecto_ICA$NORTES)

Residuals:
    Min       1Q   Median       3Q      Max
-0.22280 -0.07150  0.00220  0.07179  0.33089

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -4.172e+00  5.576e-01  -7.483 4.98e-10 ***
Proyecto_ICA$NORTES  4.545e-06  5.422e-07  8.383 1.59e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1057 on 57 degrees of freedom
Multiple R-squared:  0.5522,    Adjusted R-squared:  0.5443
F-statistic: 70.28 on 1 and 57 DF,  p-value: 1.593e-11

```

Fig. 5 Modelo de regresión lineal inicial

Se debe analizar la normalidad en la regresión planteada y en los residuos. Para la regresión se obtiene que la variable ICA presenta un p-value de 0.00439, lo anterior indica la no normalidad en los datos, pero al aplicar el test a los residuos salvados, se obtiene un p-value de 0.4497 concluyendo que no existe evidencia suficiente para aceptar hipótesis nula, por lo cual los residuos presentan normalidad. Por esta razón se continua a desarrollar la transformación BOX-COX la cual arroja un $\lambda = 0.1$ por tal razón se le aplica logaritmo a la variable respuesta para transformarla. Obteniendo como resultado el siguiente modelo:

```

Call:
lm(formula = Proyecto_ICA$ICA_trans ~ Proyecto_ICA$NORTES)

Residuals:
    Min       1Q   Median       3Q      Max
-0.51337 -0.12825  0.01529  0.10788  0.65111

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.025e+01  1.106e+00  -9.269 5.63e-13 ***
Proyecto_ICA$NORTES  9.251e-06  1.075e-06  8.602 6.94e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2096 on 57 degrees of freedom
Multiple R-squared:  0.5649,    Adjusted R-squared:  0.5572
F-statistic: 74 on 1 and 57 DF,  p-value: 6.937e-12

```

Fig. 6 Modelo de regresión lineal transformado

A este resultado se debe nuevamente aplicar el test Shapiro-wilk para evidenciar la existencia de normalidad tanto en los residuos como en los datos, en donde arroja un p-value de 0.4212 para el caso de los residuos y un p-value de 0.0841 para los datos. Para los dos casos no existe evidencia suficiente para aceptar hipótesis nula por lo cual existe normalidad en los datos, obteniendo como resultado el siguiente geodato:

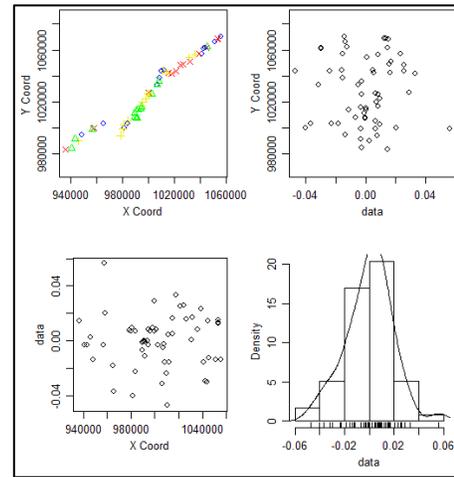


Fig. 7 Geodato Final

En la Fig. 7 mostrada anteriormente es posible identificar que se ha eliminado la tendencia en los datos (distribución aleatoria de los datos), esto gracias a que se normalizaron los errores. Al tener los datos sin ningún rastro de tendencia y estar seguros que no existe anisotropía se procede a desarrollar la estimación del modelo de semivariograma.

B. Estimador Clásico

Para comenzar con el desarrollo de este procedimiento, se debe generar un variograma, en el cual se especifican valores en todas las direcciones con el método de puntos. Este variograma inicial no permite analizar fácilmente el comportamiento de la información, por lo cual es necesario desarrollar el semivariograma en una dirección específica, cabe resaltar que la distancia recomendada para el análisis de los semivariogramas equivale a un medio de la distancia máxima identificada [1], equivaliendo está a 75000, como es posible identificar en la Fig. 8 mostrada a continuación.

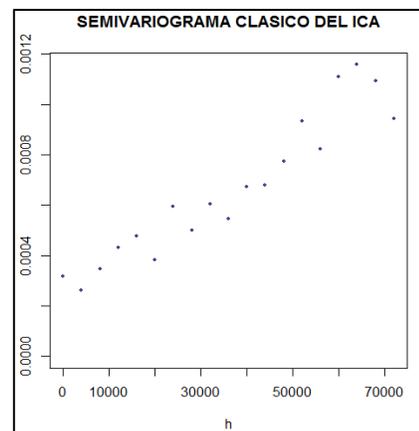


Fig. 8 Semivariograma clasico del ICA

Finalmente para revisión del semivariograma clásico generamos un gráfico Box-plot en el cual podemos observar la existencia de datos atípicos (aquellos que salen de la probabilidad esperada en los percentiles), los cuales pueden llegar a atrofiar la medición.

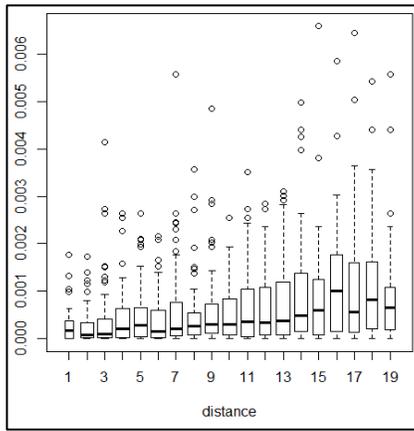


Fig. 9 Semivariograma clásico ICA- Grafico Box-plot

En la Fig. 10 es posible identificar el desarrollo de siete modelos de semivariograma experimental, en los que cada uno de ellos presenta una estructura determinada previamente, además que es posible llegar a modificar los parámetros para lograr un ajuste óptimo de cada modelo. Visualmente se observa cuáles de los modelos se ajusta mejor a los datos. Pero como complemento se observan los valores de AIC y BIC para cada uno de los modelos.

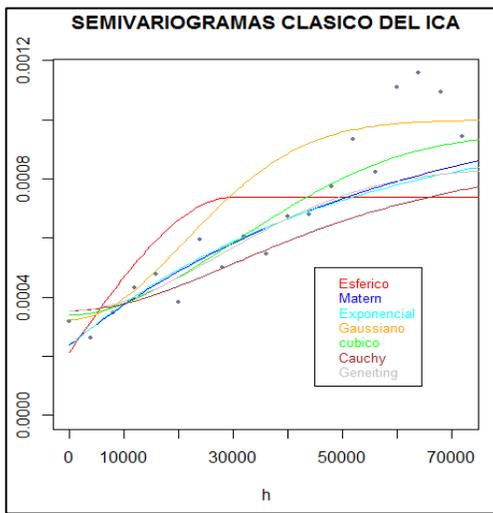


Fig. 10 Modelos de semivariogramas ajustados

Los modelos que mejor se ajustan son aquellos que tienen los valores de AIC Y BIC más pequeños por lo tanto se determinan que estos modelos son: Cauchy, Gneiting y Cubico respectivamente. Lo anterior puede ser corroborado en la tabla A estos modelos se les realiza el ajuste por máxima verosimilitud y máxima verosimilitud restringida.

MODELO	AIC	BIC
Esférico	-285.1771	-276.8669
Matern	-283.2431	-272.8554
Exponencial	-285.2431	-276.9329
Gaussiano	-279.4238	-271.1137

Tabla 2 Valores de AIC y BIC de los modelos ajustados

C. Modelo Seleccionado

El modelo que mejor se acopla posteriormente al ajuste de máxima verosimilitud y máxima verosimilitud restringida es el modelo GNEITING. Esto se determinó observando nuevamente los valores de AIC y BIC más pequeños de los tres modelos, obteniendo un mejor ajuste por máxima verosimilitud restringida teniendo un valor de AIC -276.6746 y un BIC 268.3644.

Ahora bien, es importante tener en cuenta que aunque el modelo Gneiting ajustado por máxima verosimilitud restringida es el que mejor ajusta, este modelo no puede ser utilizado en la función vgm de la librería gstat del software R para generar las diferentes predicciones. Por lo anterior es necesario evaluar cuál de los modelos con los que trabaja esta función mejor ajusta. De tal manera tenemos cuatro modelos (Esférico, Matern, exponencial y gaussiano). El ajuste de estos modelos se puede visualizar en la Fig. 11.

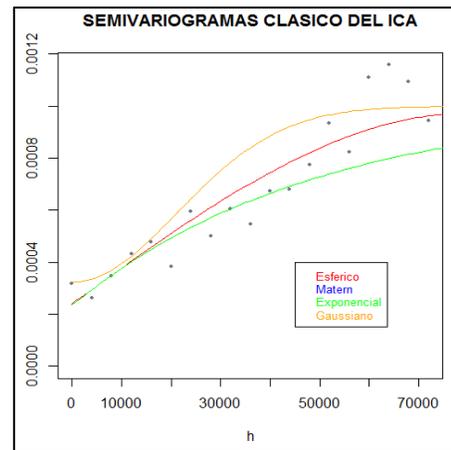


Fig. 11 Modelos ajustados

Se debe realizar de nuevo el análisis, pero solo teniendo en cuenta los modelos que utiliza la función vgm. Los modelos con menor valor en AIC Y BIC son Matern y Gaussiano a los cuales se les realizara el ajuste respectivo

MODELO	AIC	BIC
Matern	-283.142	-272.7543
Exponencial	-285.2431	-276.9329
Cúbico	-280.3767	-272.0665
Cauchy	-277.886	-267.4984
Gneiting	-279.2838	-270.9737

Tabla 3. Valores AIC y BIC-Modelos de interpolación

Finalmente se concluye que el mejor modelo es el GAUSSIANO, para poder utilizar este modelo para las interpolaciones, es necesario realizar un ajuste por máxima verosimilitud, máxima verosimilitud restringida, mínimos cuadrados ordinarios y mínimos cuadrados ponderados. Al realizar este ajuste primero se observa visualmente cual es el

modelo ajustado que mejor se asemeja al semivariograma. Lo anterior puede ser visualizado en la Fig. 12

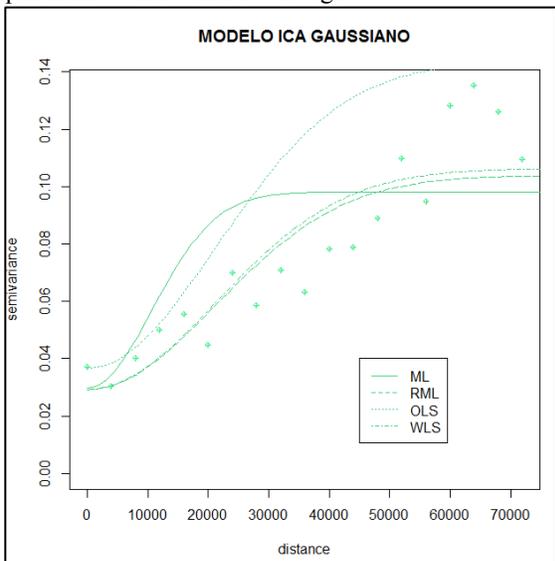


Fig. 12 Ajustes modelo Gaussiano

Al observar la fig. 12 se determina que para el modelo gaussiano los mejores ajustes son el de máxima verosimilitud y máxima verosimilitud restringida. Determinados los anteriores modelos finalmente se revisan los valores de AIC y BIC que se encuentran en la tabla 4

METODO DE AJUSTE	AIC	BIC
Máxima verosimilitud	-0.09838819	10.2893
Máxima verosimilitud restringida	0.8550298	11.24272

Tabla 4. Valores de AIC Y BIC para ML y RML

Como se puede observar en la tabla anterior el modelo que mejor ajusta es el gaussiano, por el ajuste de máxima verosimilitud ya que tiene los valores de AIC y BIC más pequeños. Este modelo será el empleado para realizar los métodos de interpolación.

IV. MÉTODOS DE INTERPOLACIÓN

A. KRIGING

Al observar todo el análisis realizado a los datos, se concluye que es necesario emplear el método de Kriging Universal cuando hay presencia de tendencia [1]. Por esta razón se emplea este método aplicado en el Software ArcGis y se obtiene como resultado el siguiente mapa:

http://giira.udistrital.edu.co/mapas/KRIGING/Kriging_universal.jpg

Como se puede observar en el mapa anterior los valores más altos del ICA se encuentra en la cuenca alta donde el río

nace. Esto es totalmente cierto, ya que teniendo en cuenta la clasificación para los valores del ICA se dice que un ICA de 0,71 a 1 se puede catalogar la calidad del agua como Buena o aceptable. Si observamos las convenciones del mapa vemos que el color rojo refleja valores que están entre 0,69 a 0,8. En la cuenca media donde debemos tener en cuenta que el río pasa por la ciudad de Bogotá observamos colores amarillos y azules, que comprenden valores entre 0.27 hasta 0.43, según la clasificación del ICA estas zonas presentarían una calidad del agua mala y regular. Es bueno tener en cuenta que la clasificación del agua como Muy mala se encuentra entre 0 hasta 0,25.

B. IDW

Este método de interpolación se obtiene por medio de la suma ponderada de los valores de la variable ICA. Es importante tener en cuenta en que si la distancia entre un punto y el punto a predecir es muy grande el peso disminuirá. A continuación se presentan los mapas de predicción de la interpolación de IDW para un $p=3$ y un $p=5$ (parámetros de definición del método). En estos mapas se puede observar un comportamiento muy similar al de la interpolación por Kriging siendo la cuenca media la que cuenta con los valores más bajos del ICA entre colores morado a azules. Es importante resaltar los puntos de color negro donde se presenta los valores más bajos del ICA son zonas aledañas a la ciudad de Bogotá expresando así valores más elevados de contaminación.

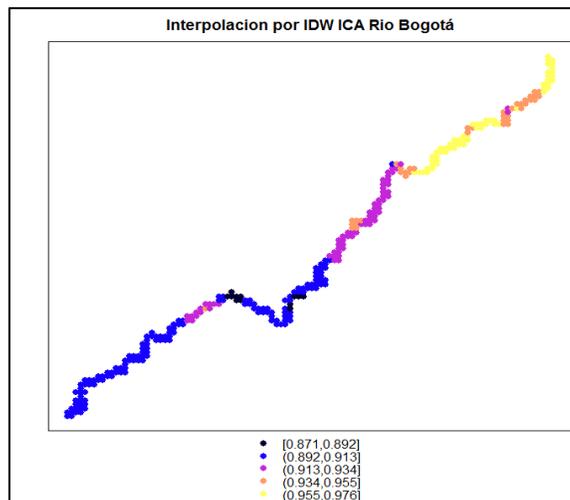


Fig. 13 Interpolación IDW

De igual forma también se genera la salida gráfica del método en el software ArcGis que se encuentra en la siguiente URL:

http://giira.udistrital.edu.co/mapas/IDW/IDW_ICA.jpg

C. Funciones de Base Radial

Las funciones de base radial requieren de los parámetros “eta” y “rho” que se deben optimizar utilizando (eta.opt= TRUE, rho.opt= TRUE) dentro de la función graph.rbf en la programación del código en el software RStudio. Lo anterior

junto con el procedimiento desarrollado en el código permite obtener el mapa de predictores para diferentes funciones como TPS, ST, Exponencial, Multicuadratica, CRS Y Gaussiana.

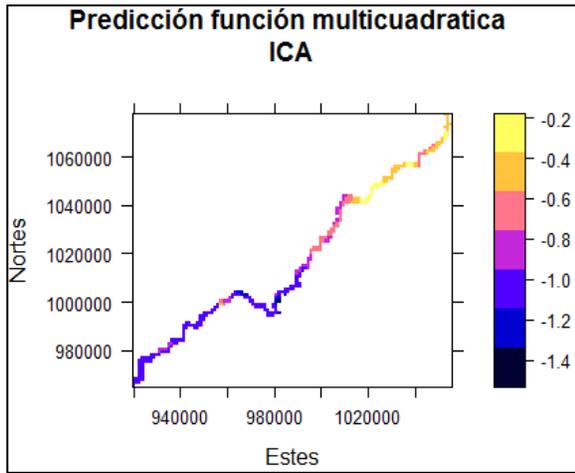


Fig. 14 Predicción Función multicuadrática-ICA-

Como se puede observar en cada uno de los mapas de predicciones se observan varias diferencias en la predicción del ICA sobre todo hacia la cuenca media y baja del río. Como por ejemplo TPS (Fig. 15) y Multicuadratica (Fig. 14) para la parte de la cuenta baja que comienza saliendo de la sabana de Bogotá predicen valores donde la calidad del agua se clasificaría entre mala y regular. Ahora analicemos la cuenca alta del río donde en general todas las funciones predicen un comportamiento muy similar, el cambio se encuentra hacia la cuenta media y baja.

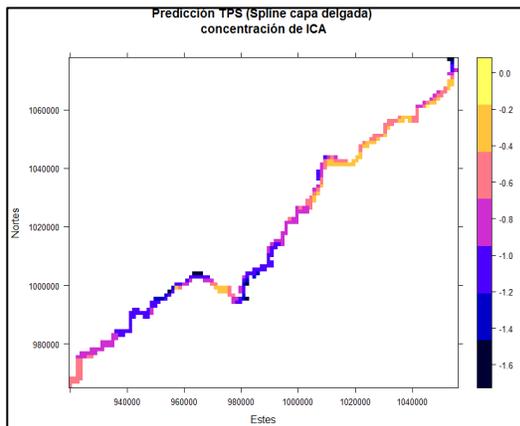


Fig. 15 Función TPS

En el software ArcGis se generaron de igual manera las salidas gráficas para las funciones CRS, Multicuadratica, Multicuadratica Inversa y ST. Las cuales pueden ser visualizadas en las siguientes URL:

- <http://giira.udistrital.edu.co/mapas/RBF/CRS.jpg>
- <http://giira.udistrital.edu.co/mapas/RBF/Multicuadratica.jpg>
- http://giira.udistrital.edu.co/mapas/RBF/Multicuadratica_inv.jpg

http://giira.udistrital.edu.co/mapas/RBF/Spline_tensor.jpg

D. DISEÑO DE RED

El diseño de red se realiza para determinar si se pudo haber obtenido una predicción mucho mejor. Por lo tanto se desarrollan mayas de puntos como una cantidad de puntos en el eje x y en el eje y.

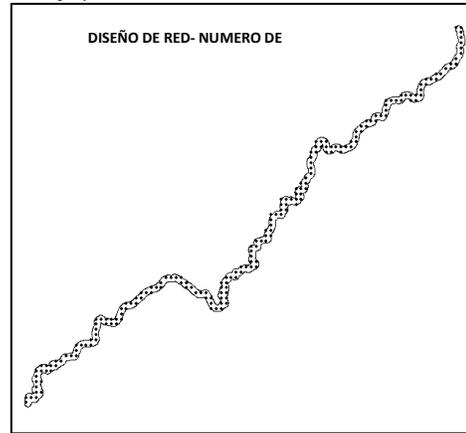


Fig. 16 Diseño de Redes

Es necesario determinar el ASEPE que corresponde al Error estándar promedio de los errores de predicción. Por lo tanto se puede concluir que entre mayor sea la malla de puntos menor error se presentará, esto se sustenta con la tabla respectiva donde las mallas de 800x8000 y 500x500 son las que tienen menor error. Lo anterior se visualiza en la tabla 5.

Tamaño de Red	ASEPE
50 X 50	0.0003836804
200 X 200	4.246205e-05
500 X 500	1.070268e-05
800 X 800	4.482714e-06

Tabla 5. ASEPE Diseño de redes

El diseño de red se aplica ya que los métodos de interpolación Kriging son exactos y los valores pronosticados son iguales a los valores tomados en los puntos, lo cual nos permite determinar que tan buenos son estos pronósticos. Por lo tanto entre mayor número de puntos menor será el error.

V. CONCLUSIONES

Como se pudo observar en todos los mapas de predicción, se confirma que las zonas donde se debería producir un índice de contaminación más alto por la intervención de industrias y en general por cercanía a las ciudades predicen valores del ICA muy bajos, que indicarían que el agua se encuentra en pésimas condiciones. Mientras que se verifica que en la zona donde nace el río es lógico que se tengan valores del ICA altos ya que el agua de esta zona no tiene la intervención de industrias. La zona de la cuenca baja donde se llega al río Magdalena tiende a tomar valores medios a altos del ICA, esto da a entender que a medida que el agua transcurre realiza un proceso de limpieza y que además estas zonas no le generan al río vertimiento de desechos ni malos hábitos de la sociedad que puedan generar

contaminación al río. Los métodos de interpolan dan a conocer con el agua se comporta en todo el río solamente con los datos tomados en las 59 estaciones. Con estos resultados se podrían determinar las zonas exactas del río donde planes de limpieza serian efectivos y las zonas en donde la vigilancia se debería incrementar para evitar las acciones que sigan contaminando las aguas.

VI. REFERENCIAS

- [1] Melo, C. (2012). *Análisis Geoestadístico Espacio Tiempo Basado en Distancias y Splines con Aplicaciones*. Barcelona: Universidad de Barcelona.
- [2] Instituto de Hidrología, M. y.-I. (2011). *Hoja metodológica del indicador Índice de calidad del agua (Version 1,00)*. . Sistema de Indicadores Ambientales de Colombia- Indicadores de Calidad del agua Superficial.
- [3] Castro, M., Almeida, J., Ferrer, J., & Díaz , D. (2014). Indicadores de la calidad del agua: Evolucion y tendencias a nivel global. *Ingenieria Solidaria*, 111-124.
- [4] Instituto de hidrologia, m. y. (2012). Índice de alteración potencial de la calidad del agua (IACAL). (*Hoja metodológica versión 1,00*).
- [5] Ministerio de Ambiente , y. (2012). *Decreto 1640 del 2012*. Republica de Colombia.
- [6] Regional, C. A. (2006). PLAN DE ORDENACIÓN Y MANEJO DE LA CUENCA HIDROGRÁFICA DEL RÍO BOGOTÁ. *Cuenca hidrografica Rio Bogota*.