

**METODOLOGÍA PARA LA ESTIMACION DE ZONAS CON NECESIDAD DE  
REHABILITACIÓN EN REDES DE DISTRIBUCIÓN DE AGUA POTABLE  
USANDO GEOSTADÍSTICA**

**ANDRÉS FERNANDO ARIZA BELTRÁN  
ALVARO SEBASTIÁN BELTRÁN MANTILLA**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS  
ESCUELA DE INGENIERÍA CIVIL  
BUCARAMANGA**

**2015**

**METODOLOGÍA PARA LA ESTIMACION DE ZONAS CON NECESIDAD DE  
REHABILITACIÓN EN REDES DE DISTRIBUCIÓN DE AGUA POTABLE  
USANDO GEOSTADÍSTICA**

**ANDRÉS FERNANDO ARIZA BELTRÁN  
ALVARO SEBASTIÁN BELTRÁN MANTILLA**

**Trabajo de Grado para optar al título de  
Ingeniero Civil**

**Director  
LUIS FERNANDO CASTAÑEDA GALVIS  
Ingeniero Civil, M.SC**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS  
ESCUELA DE INGENIERÍA CIVIL  
BUCARAMANGA  
2015**

Dedicamos este trabajo de grado a primeramente a Dios y a nuestros padres, por brindarnos apoyo incondicional a lo largo de nuestra formación como profesionales y por darnos la motivación suficiente para salir adelante ante las adversidades.

Agradecemos a nuestro director de trabajo de grado, el ingeniero Luis Fernando Castañeda Galvis por haber puesto a nuestra disposición su tiempo y conocimientos para la realización de este trabajo de grado.

Además agradecemos a todos nuestros compañeros y docentes de pregrado que contribuyeron de una u otra manera en nuestra formación como profesionales.

# CONTENIDO

	Pág.
INTRODUCCIÓN .....	15
1. DAÑOS REGISTRADOS POR EL AMB S.A. E.S.P. EN LA RED DE DISTRIBUCIÓN DE AGUA POTABLE DE BUCARAMANGA.....	17
1.1 UNIFICACIÓN DE LOS ATRIBUTOS DE LOS DAÑOS .....	22
1.2 DATOS INCIDENTES .....	24
2. ESTIMACIÓN DE ZONAS CON MAYOR NECESIDAD DE REHABILITACIÓN A PARTIR DE LA GEOSTADISTICA .....	26
2.1 ANÁLISIS DE PUNTOS CALIENTES .....	26
2.1.1 Agrupamiento de los daños. ....	28
2.1.2 Incremental Spatial Autocorrelation. ....	30
2.1.3 Resultados del análisis de puntos calientes .....	32
3. ESTIMACIÓN DE LOS FACTORES QUE INFLUYEN EN LA PRESENCIA DE DAÑOS CON BASE EN ANÁLISIS DE REGRESIÓN .....	38
3.1 ANÁLISIS DE REGRESIÓN .....	38

3.1.1 Factores influyentes considerados para el análisis de regresión .....	40
3.1.2 Zonas censales .....	43
3.2 VALIDACIÓN DE LOS RESULTADOS .....	45
3.3 ANÁLISIS DE RESULTADOS DE REGRESIÓN OLS .....	45
3.3.1 Factor material .....	45
3.3.2 Factor edad .....	46
3.3.3 Factor diámetro .....	47
3.3.4 Factor presión .....	48
3.4 IDENTIFICACIÓN DE TUBOS CON FACTORES IDENTIFICADOS MEDIANTE EL ANÁLISIS DE REGRESIÓN .....	49
4. METODOLOGÍA PARA LA ESTIMACIÓN DE ZONAS CON NECESIDAD REHABILITACIÓN DE LAS REDES DE DISTRIBUCIÓN DE AGUA POTABLE ...	50
5. CONCLUSIONES .....	52
BIBLIOGRAFIA .....	54
ANEXOS .....	56

## LISTA DE FIGURAS

Pág.

Figura 1. Representación del conjunto de daños registrados por el amb S.A. E.S.P. desde el año 2006 hasta el 2013 mediante puntos en un Shapefile.....	18
Figura 2. Distribución de los daños presentados entre los años 2006 y 2013. ....	25
Figura 3. Representación gráfica datos de entrada y salida herramientas Integrate - Collect Events.....	29
Figura 4. Datos incidentes agrupados desde 2006 al año 2013. ....	30
Figura 5. Reporte generado por Incremental Spatial Autocorrelation para los datos agrupados desde el año 2006 hasta el año 2013 .....	32
Figura 6. Análisis de puntos calientes de los daños agrupados desde el año 2006 hasta el año 2013. ....	33
Figura 7. Visualización del análisis de puntos calientes de los todos daños presentados desde el año 2006 hasta el año 2013. ....	34
Figura 8. Visualización prioridad de rehabilitación de los distritos. ....	37
Figura 9. Ecuación análisis de regresión. ....	38
Figura 10. Tipos de relación .....	40
Figura 11. Distribución del factor material.....	42
Figura 12. Distribución del factor edad. ....	42
Figura 13. Distribución del factor diámetros.....	43
Figura 14. Distribución del factor presiones.....	43
Figura 15. Visualización zonas censales. ....	44
Figura 16. Tubos con factores identificados con el análisis de regresión. ....	49

## LISTA DE TABLAS

	Pág.
Tabla 1. Fields año 2008. ....	19
Tabla 2. Fields año 2007. ....	20
Tabla 3. Fields año 2008. ....	20
Tabla 4. Fields año 2009. ....	21
Tabla 5. Fields año 2010. ....	21
Tabla 6. Fields año 2011, 2012 y 2013. ....	22
Tabla 7. Diccionario con unificación de fields. ....	23
Tabla 8. Tipos de daños usados para el desarrollo de la metodología. ....	24
Tabla 9. Prioridad de rehabilitación por distrito. ....	36
Tabla 10. Factores considerados para el análisis de regresión. ....	41

## LISTA DE ANEXOS

	<b>Pág.</b>
Anexo A. Resultados numéricos análisis de regresión factor material.....	56
Anexo B. Resultados numéricos análisis de regresión factor edad. ....	56
Anexo C. Resultados numéricos análisis de regresión factor diámetro. ....	57
Anexo D. Resultados numéricos análisis de regresión factor presión. ....	57
Anexo E. Distribución residuos análisis de regresión factor material.....	58
Anexo F. Distribución residuos análisis de regresión factor edad.....	59
Anexo G. Distribución residuos análisis de regresión factor diámetro. ....	60
Anexo H. Distribución residuos análisis de regresión factor presión.....	61

## RESUMEN

**TÍTULO:** METODOLOGÍA PARA LA ESTIMACION DE ZONA CON NECESIDAD DE REHABILITACIÓN EN REDES DE DISTRIBUCIÓN DE AGUA POTABLE USANDO GEOSTADÍSTICA \*

**AUTORES:** ANDRÉS FERNANDO ARIZA BELTRÁN  
ÁLVARO SEBASTIÁN BELTRÁN MANTILLA \*\*

**PALABRAS CLAVE:** GEOSTADÍSTICA, ANÁLISIS DE PUNTOS CALIENTES, ANÁLISIS DE REGRESIÓN, DAÑOS EN REDES DE DISTRIBUCIÓN, REDES DE DISTRIBUCIÓN, DISTRITOS HIDRÁULICOS.

### DESCRIPCIÓN:

Este trabajo de investigación presenta una metodología para la estimación de zonas con necesidad de rehabilitación a partir de un análisis geostadístico y estadístico espacial de la información histórica de daños presentados en la red de distribución de agua potable del área metropolitana de Bucaramanga. La metodología se crea a partir de un análisis de puntos calientes aplicado a la distribución geográfica de los daños presentados en los distritos hidráulicos.

La rotura de los elementos de conducción en las redes distribución obedecen a múltiples factores, por tal motivo y en segunda instancia se realiza un análisis de regresión que permite la creación de un modelo que se ajusta a la relación entre el número de daños y cada factor influyente. En cada modelo se presenta su respectiva validación de los resultados con el fin de poder confiar y utilizar el mismo. Además, se realiza una representación gráfica de la relación entre los elementos de conducción existentes y los factores con mayor influencia estimados mediante el modelo originado por el análisis de regresión.

Finalmente, se expone una clasificación de los distritos hidráulicos de acuerdo a la prioridad de rehabilitación y los factores con mayor influencia en la presencia de daños en las redes de distribución de agua potable. Los análisis y modelos presentados en el presente trabajo se realizaron mediante el uso del software ArcGIS.

---

\* Trabajo de Grado.

\*\* Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingeniería Civil. Director: Luis Fernando Castañeda Galvis, Ingeniero Civil, M.SC.

## ABSTRACT

**TITLE:** METODOLOGY TO ESTIMATE THE ZONE WHICH NEEDS TO BE REHABILITATED IN THE DRINKING WATER DISTRIBUTION NETWORKS USING GEOSTATISTICS \*

**AUTHORS:** ANDRÉS FERNANDO ARIZA BELTRÁN  
ÁLVARO SEBASTIÁN BELTRÁN MANTILLA\*\*

**KEYWORDS:** GEOSTATISTICS, HOT SPOTS ANALYSIS, REGRESSION ANALYSIS, DAMAGES IN THE DISTRIBUTION NETWORKS, DISTRIBUTION NETWORKS, HYDRAULIC DISTRICTS.

### DESCRIPTION:

This research work presents a methodology to estimate the zones which need to be rehabilitated based on a geostatistical and spatial- statistical analysis of the historical information of the damages presented in the drinking water distribution network of Bucaramanga's metropolitan area. This methodology was created from a hot spots analysis applied to the geographical distribution of the damages presented in the hydraulic districts.

The breaking of the elements of conduction in the distribution networks are due to multiple factors; for this reason and in second place, a regression analysis was done. This allows the creation of a model that is adjusted to the relationship between the number of damages and each influencing factor. In each model, it is presented its corresponding validity of the results to make reliable and use it. Besides, it is done a graphic representation of the relationship between the conducting existing elements and the factors with the greatest influence which were estimated through a model originated by the regression analysis.

Finally. It is exposed a classification of the hydraulic districts according to the priority of rehabilitation and the factors with the greatest influence in the presence of damages in the drinking water distribution networks. The analysis and models presented in this work were done by using software ArcGIS.

---

\* Bachelor Thesis.

\*\* Faculty of Physical-Mechanical Engineerings. Civil Engineering School. Director: Eng. M. Sc. Luis Fernando Castañeda Galvis.

## INTRODUCCIÓN

Una red de distribución de agua potable permite el transporte del agua de los tanques de almacenamiento del acueducto hasta cada una de las viviendas de la comunidad. Esto es posible gracias a un conjunto de tuberías que junto a sus respectivos accesorios (válvulas, medidores y empalmes) conforman una sola red con el fin de distribuir el agua hasta distintas zonas de la localidad. En las redes de distribución de agua potable ocurren daños frecuentemente en los elementos de conducción y accesorios, los cuales generan pérdidas económicas e inconvenientes en el servicio. Los daños se pueden presentar por características propias de los elementos hidráulicos, la falta de mantenimiento mediante labores preventivas o debido a factores externos al sistema de conducción.

El constante crecimiento de la población y la escasez del recurso hídrico hacen que los sistemas de distribución de agua potable requieran de una mayor complejidad para suplir las necesidades de la comunidad. Es por ello que las empresas prestadoras del servicio se ven en la tarea de buscar nuevas estrategias que permitan optimizar procesos y reducir costos, como lo es la creación de metodologías para la implementación de actividades de mantenimiento, el cual consiste típicamente en la rehabilitación, reparación y renovación de las tuberías que no permiten la eficaz entrega de los niveles de servicio. La mayoría de las políticas de mantenimiento combinan el desarrollo de una técnica de soporte de decisión a partir del reemplazo de líneas de tubería basada en el juicio técnico bajo el buen conocimiento hidráulico del sistema<sup>1</sup>. Las políticas enfocadas para la rehabilitación de redes de distribución presentan dificultades para determinar qué zonas requieren de estas labores y cuáles deben realizarse con mayor prioridad.

---

<sup>1</sup> ARAQUE, Diego Alejandro. Optimización de Redes de Acueducto con el fin de Uniformizar el Estado de presiones. Trabajo de grado Maestría. Bogotá D.C: Universidad de los Andes. Facultad de Ingeniería. Departamento de Ingeniería Civil y Ambiental, 2004. 3 p.

En la actualidad, el registro de los daños de la redes de distribución se realiza en bases datos digitales por parte de la mayoría de las empresas prestadoras del servicio. El uso de herramientas digitales y software de última generación pueden ser aplicados para realizar modelaciones y análisis de la información de los daños registrados que permitan conocer y/o entender patrones y causas de los mismos. Esta información puede ser utilizada y analizada para determinar y generar nuevas pautas en los tipos de mantenimiento.

El presente trabajo desarrolla una metodología que permite estimar cuáles zonas presentan mayor necesidad de rehabilitación a partir de análisis geoestadísticos. De manera conjunta se estiman los factores con mayor influencia en la presencia de daños en la redes de distribución mediante análisis de regresión. Se hace uso del registro de daños de las redes de distribución del área metropolitana de Bucaramanga y se emplea el software ArcGIS, el cual permite organizar y analizar información geográfica como herramienta para el desarrollo de los análisis especificados.

## 1. DAÑOS REGISTRADOS POR EL AMB S.A. E.S.P. EN LA RED DE DISTRIBUCIÓN DE AGUA POTABLE DE BUCARAMANGA

El amb S.A. E.S.P. es una empresa estatal cuyo propósito es el prestar servicios de alcantarillado y saneamiento básico además de la producción y distribución de agua potable a las ciudades colombianas de Bucaramanga, Girón y Floridablanca.

Desde el año 2006 el amb S.A. E.S.P. se ha encargado de la tarea de registrar los daños que se ha presentado en las redes de distribución de agua potable de la ciudad de Bucaramanga.

Hasta el año 2013 el amb S.A. E.S.P. registró los daños en una base de datos compatible con el software ArcGIS el cual permite recopilar organizar, administrar, analizar y distribuir información geográfica<sup>2</sup>. Los daños se agruparon en un archivo de formato vectorial conocido como *shapefile*, el cual permite guardar la localización de elementos geográficos y los atributos asociados a ellos. Estos elementos pueden ser representados gráficamente mediante polígonos, líneas o puntos<sup>3</sup>. Los atributos hacen referencia a cualidades o características que pueden ser asignadas a cada uno de los daños mediante *fields* en una tabla de atributos.

Los *fields* son campos en una tabla de atributos los cuales pueden ser de diferente tipo y almacenar distintas propiedades o características de un datos almacenado.

---

<sup>2</sup> ESRI. ¿Qué es ArcGIS? [en línea]. ESRI, 1995-2013. [New York, E.U.] ESRI, 2013. [Citado 19 de Octubre de 2015]. Disponible en Internet: <http://resources.arcgis.com/es/help/getting-started/articles/026n00000014000000.htm>

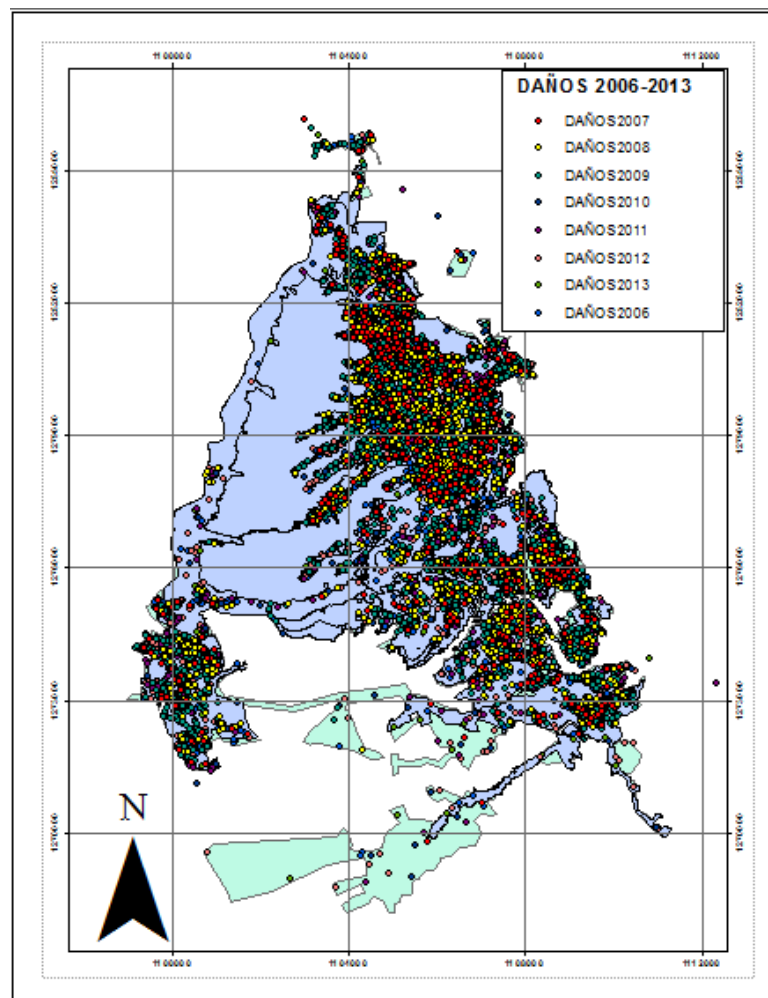
<sup>3</sup> ESRI. Why Shapefiles?. En: ESRI Shapefile Technical Description. New York, E.U, July 1998, p. 1.

Los tipos de *fields* empleados por el amb S.A. E.S.P. para el almacenamiento de la información se describen a continuación:

- **String:** permite representar una serie de símbolos alfanuméricos.
- **Double:** permite almacenar números.
- **Date:** permite almacenar fechas, horas o fechas y horas.

En la Figura 1 se muestra la representación para los daños registrados desde el año 2006 hasta el año 2013.

Figura 1. Representación del conjunto de daños registrados por el amb S.A. E.S.P. desde el año 2006 hasta el 2013 mediante puntos en un Shapefile.



La información suministrada por el amb S.A. E.S.P. contaba con *fields* de diferentes nombres para una misma característica y atributos sin asignar. Esto imposibilitó realizar los análisis de los datos empleando el software ArcGIS.

Debido a que el registro de los daños fue realizado por diferente personal del Acueducto Metropolitano de Bucaramanga y en distintas épocas, se presentaron diferencias entre los atributos asignados a la información recopilada para cada año en cuanto al tipo de field en el cual se registraron y la descripción del mismo.

En las Tablas 1, 2, 3, 4, 5, 6 se observa la diferencia presente entre los nombres de los *fields* asignados a los daños en diferentes años.

Tabla 1. Fields año 2008.

<b>AÑO 2006</b>			
<b>No.</b>	<b>Field</b>	<b>No.</b>	<b>Field</b>
1	OBJECTID	15	HORA_SOLIC
2	X	16	FECHA__IMP
3	Y	17	HORA__IMPR
4	CODIGO	18	FECHA__DES
5	CODIGO_MUN	19	HORA__DESC
6	CODIGO_BAR	20	FINIC_EJEC
7	DIRECCION	21	HORA_INIC
8	NOMBRE_MUN	22	FFINAL_EJE
9	ORDEN	23	Campo17
10	NOMBRE_BAR	24	IMP__SOLI
11	MEDIO	25	INIC__IMP
12	R_E_S_U_L	26	SOL_INIC
13	DIAMETRO	27	FINAL__IN
14	FECHASOLIC	28	DESC__FIN

Tabla 2. Fields año 2007.

<b>AÑO 2007</b>					
<b>No.</b>	<b>Field</b>	<b>No.</b>	<b>Field</b>	<b>No.</b>	<b>Field</b>
1	OBJECTID	19	HORAINIC	37	PORC_UTIL
2	A_CODIGO	20	HORAFINAL	38	VALOR_IVA
3	X	21	LECTURA	39	LIQ_FRAUDE
4	Y	22	CUADRILLA	40	CUOTAS
5	DAÑOS_ABR	23	RESULTADO	41	ASIGNADO
6	Bucaramang	24	OBSERVAC	42	FECHASYS
7	Bucarama_1	25	VALVISITA	43	HORASYS
8	NUMERO	26	VALSERVICI	44	FECHAIMPRE
9	USUARIO	27	VALMATERIA	45	HORAIMPRE
10	FECHASOLIC	28	VALMEDIDOR	46	CONSECUTIV
11	HORASOLIC	29	TIPOSERSOL	47	DIAMTUBE
12	SOLICITA	30	LECTUSOLIC	48	FECHAEJECF
13	MEDIOSOL	31	USERSOLIC	49	DAÑOSNUEV
14	NUMESOLIC	32	USERREPOR	50	ORDEN
15	OBSERV118	33	TIPOMEDI	51	BARRIO
16	IMPRESION	34	NUMEROMEDI	52	MEDIO
17	TIPOSERV	35	SERIEMEDI	53	R_E_S_U_L
18	FECHAEJEC	36	PORC_IVA	54	DIAMETRO

Tabla 3. Fields año 2008.

<b>AÑO 2008</b>			
<b>No.</b>	<b>Field</b>	<b>No.</b>	<b>Field</b>
1	OBJECTID	10	TIPOSERV
2	CODIGO	11	ID_TUBERIA
3	X	12	TIPO_TUBO
4	Y	13	DISTRITO
5	COD_MUNICI	14	BARREAL
6	DIRECCION	15	DIAMETRO_1
7	ORDEN	16	UNIDAD
8	NOMBRE_MUN	17	Expr1015
9	FECHASOLIC	18	

Tabla 4. Fields año 2009.

<b>AÑO 2009</b>			
<b>No.</b>	<b>Field</b>	<b>No.</b>	<b>Field</b>
1	OBJECTID	12	TipoDaño
2	FID_	13	Diametro
3	Codigo_usu	14	FechaSolic
4	X	15	HoraSolici
5	Y	16	ID_Tuberia
6	Cod_Munici	17	Tipo_Tubo
7	Direccion	18	Diametro_T
8	Municipio	19	Unidad_dia
9	Orden	20	Sector
10	Barrio	21	Distancia_
11	Medio	22	Mes_Daño

Tabla 5. Fields año 2010.

<b>AÑO 2010</b>			
<b>No.</b>	<b>Field</b>	<b>No.</b>	<b>Field</b>
1	OBJECTID	10	tipodaño
2	x	11	diametro
3	y	12	fecha_soli
4	Expr1002	13	id_tuberia
5	direccion	14	tipo_tubo
6	orden	15	diametro_1
7	barrio	16	sector
8	medio	17	distance_1
9	Expr1015	18	

Tabla 6. Fields año 2011, 2012 y 2013.

AÑO 2011			
No.	Field	No.	Field
1	OBJECTID	13	HORSOL_OT
2	NUM_OT	14	FCHINIE_OT
3	DIREC_OT	15	HORINE_OT
4	BARRIO_OT	16	FCHFINE_OT
5	MUNIC_OT	17	HORFINE_OT
6	RESULT_OT	18	COORDX_GEO
7	UBICAC_OT	19	COORDY_GEO
8	DIAM_OT	20	ID_TUB_GEO
9	MATER_OT	21	TIPTUB_GEO
10	MATREMP_OT	22	SECTOR_GEO
11	PROFUND_OT	23	FCHTUB_GEO
12	FCHSOL_OT	24	

### 1.1 UNIFICACIÓN DE LOS ATRIBUTOS DE LOS DAÑOS

Con el objetivo de poder llevar a cabo los análisis de los datos se hizo necesario una unificación o estandarización de los *fields* de todos los años, determinando el mismo nombre y tipo para cada uno de ellos. Se utilizó la extensión de ArcGIS “ETGeoWizards” la cual permite editar la tabla de atributos de los *shapefiles*, y la herramienta *Spatial Join*, la cual permite unir los atributos de un *shapefile* a otro basándose en una relación espacial<sup>4</sup>.

En la Tabla 7 se presenta un diccionario el cual contiene la unificación de los fields junto con el tipo y la descripción de su contenido.

<sup>4</sup> ESRI. Spatial Join (Analysis) [en línea]. ESRI, 1995-2013. [New York, E.U.] ESRI, 2013. [Citado: 22 de junio 2015] Disponible en Internet: <http://resources.arcgis.com/en/help/main/10.1/index.html#//00080000000q000000>

Tabla 7. Diccionario con unificación de fields.

DICCIONARIO		
NOMBRE	TIPO	CONTIENE
<b>COORD_X</b>	DOUBLE	Coordenada X en el sistema coordenado MAGNA San Bucaramanga.
<b>COORD_Y</b>	DOUBLE	Coordenada Y en el sistema coordenado MAGNA San Bucaramanga.
<b>MUNICIPIO</b>	STRING	Municipio en el cual se encuentra ubicado el daño.
<b>DISTRITO</b>	STRING	Distrito en el cual se encuentra ubicado el daño.
<b>JERARQUIA</b>	STRING	Jerarquía del tubo en el cual se presentó el daño.
<b>EDAD</b>	STRING	Edad del tubo en el cual se presentó el daño.
<b>BARRIO</b>	STRING	Barrio en el cual se encuentra ubicado el daño.
<b>DIRECCION</b>	STRING	Ubicación del daño mediante nomenclatura de calles y carreras.
<b>RESULTADO</b>	STRING	Resultado del diagnóstico. Tipo de daño.
<b>TIPTUB_GEO</b>	STRING	Tipo de tubo en el cual se presentó el daño asignado el spatial join.
<b>TIPO_TUBO</b>	STRING	Tipo de tubo en el cual se presentó el daño.
<b>DIAMETRO</b>	STRING	Diámetro del tubo en el cual se presentó el daño.
<b>DIAM_GEO</b>	STRING	Diámetro del tubo en el cual se presentó el daño asignado el spatial join.
<b>FECHA_SOL</b>	STRING	Fecha en la que se solicitó la reparación del daño.
<b>HORA_SOL</b>	STRING	Hora en la cual se solicitó la reparación del daño.
<b>FECHA_EJEC</b>	STRING	Fecha en la que se inició la reparación del daño.
<b>HORA_EJEC</b>	STRING	Hora en la que se inició la reparación del daño.
<b>FECHA_INST</b>	STRING	Fecha en la que se instaló el tubo en el que se presentó el daño.
<b>FINST_GEO</b>	STRING	Fecha en la que se instaló el tubo en el que se presentó el daño asignado por el join.

## 1.2 DATOS INCIDENTES

El amb S.A. E.S.P. ha registrado un total de 17232 daños presentados en las redes de distribución de agua potable en el periodo comprendido desde el año 2006 hasta el año 2013, los cuales han obedecido a múltiples causas. En la Figura 2 se presenta un histograma de frecuencia que muestra la distribución de los tipos de daños presentados.

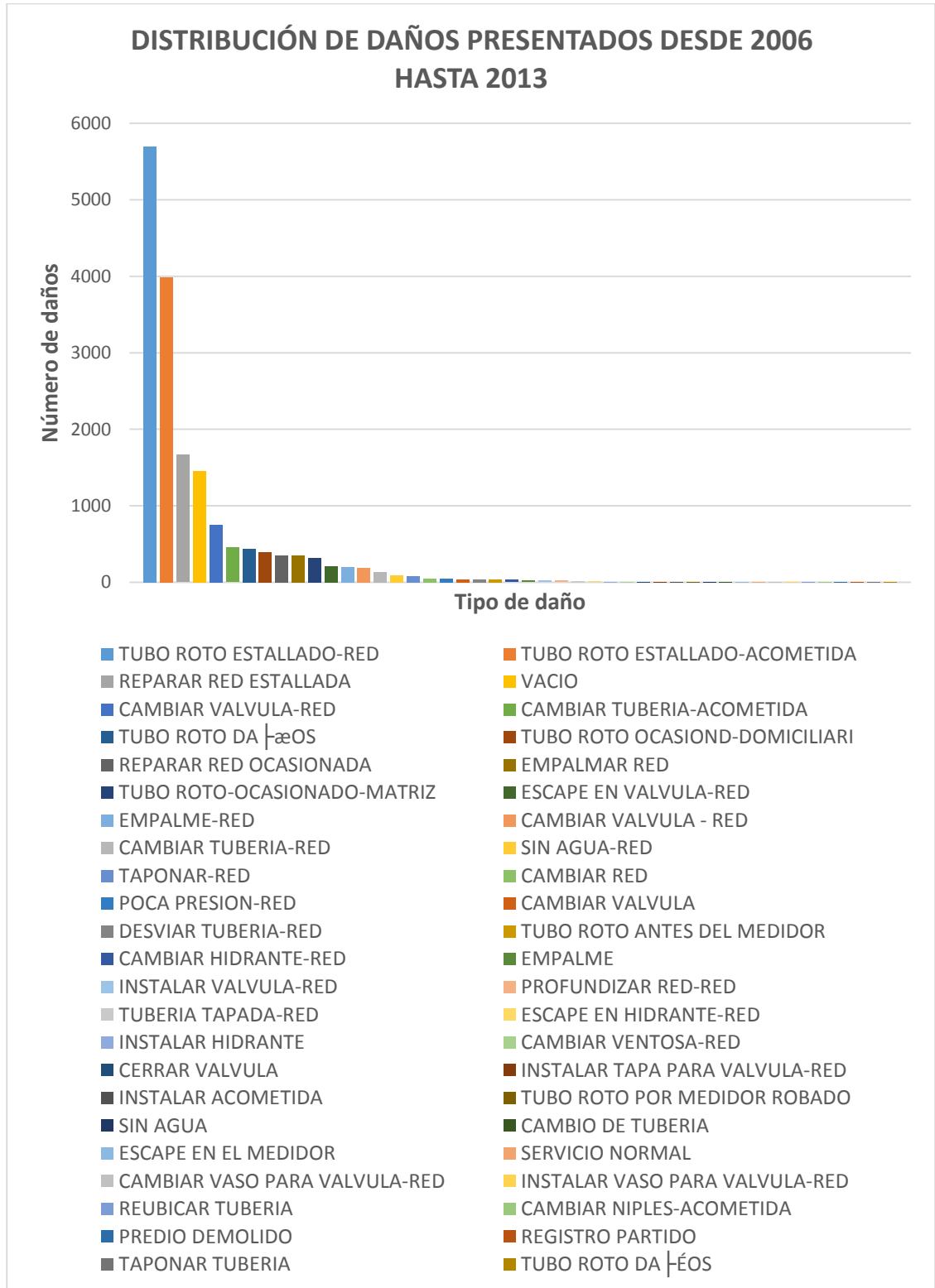
En base a la distribución de los daños y a que la metodología se centra en los daños en los elementos de conducción, se desarrolló la metodología con los tipos de daños mostrados en la Tabla 8 debido a que estos representan más del 70% de la totalidad de los daños, siendo estos tipos de daños los más representativos.

Tabla 8. Tipos de daños usados para el desarrollo de la metodología.

<b>TIPO DE DAÑO</b>	<b>%</b>
REPARAR RED ESTALLADA	9,69
TUBO ROTO ANTES DEL MEDIDOR	0,22
TUBO ROTO DAÑOS	2,56
TUBO ROTO ESTALLADO-ACOMETIDA	23,17
TUBO ROTO ESTALLADO-RED	33,09
TUBO ROTO OCASIOND-DOMICILIARI	2,32
TUBO ROTO POR MEDIDOR ROBADO	0,03
TUBO ROTO-OCASIONADO-MATRIZ	1,84
TUBO ROTO DAÑOS	0,01
<b>TOTAL</b>	<b>72,93</b>

A partir de estos tipos de daños se procedió a realizar el análisis con un total de 12550 daños, dando prioridad a los daños presentados debido a la rotura de la tubería de las redes de distribución.

Figura 2. Distribución de los daños presentados entre los años 2006 y 2013.



## **2. ESTIMACIÓN DE ZONAS CON MAYOR NECESIDAD DE REHABILITACIÓN A PARTIR DE LA GEOSTADÍSTICA**

La geostatística es una rama de la estadística aplicada cuyo objetivo es el análisis y predicción de valores asociados a fenómenos espaciales o espacio-temporales. La geostatística incorpora las coordenadas espaciales de los datos dentro del análisis. Numerosas herramientas geostatísticas fueron originalmente desarrolladas como un medio práctico para describir patrones espaciales e interpolar valores en lugares donde no se tomaron muestras<sup>5</sup>.

Dichas herramientas y métodos han evolucionado para proporcionar no solo valores interpolados sino también medidas de incertidumbre para estos valores. La medición de la incertidumbre es fundamental para la toma de decisiones puesto que proporciona información sobre los posibles valores para cada ubicación en lugar de solo un valor interpolado. El análisis geostatístico ha evolucionado de univariable a multivariable, ofreciendo mecanismos para incorporar conjuntos de datos secundarios los cuales complementen la variable primaria de interés, permitiendo la construcción de modelos de interpolación y de incertidumbre más precisos.

Esta metodología se desarrolló en base a un análisis geostatístico con el fin de estimar las zonas con necesidad de rehabilitación en las redes de distribución.

### **2.1 ANÁLISIS DE PUNTOS CALIENTES**

El análisis de puntos calientes también conocido como estadístico  $G_i^*$  es un análisis el cual permite la identificación de patrones en la distribución de una serie de datos, o la existencia de una autocorrelación espacial.

---

<sup>5</sup> DÍAZ VIERA, Martín Alberto. INTRODUCCIÓN. En: Geostatística aplicada. Cuba, 2002. p. 1.

Los análisis de puntos calientes son usados frecuentemente para análisis espacial de la distribución de crímenes, concentración de lluvias, propagación de enfermedades, incendios, accidentes y llamadas de emergencia, permitiendo encontrar patrones que sirva de fundamento para la toma de decisiones.

El análisis de puntos calientes permite a partir de un conjunto de datos de características ponderadas identificar los puntos calientes (*Hot spots*) estadísticamente significativos y los puntos fríos mediante la estadística Getis-Ord  $G_i^*$ <sup>6</sup>.

El análisis de puntos calientes funciona examinando cada punto en contexto con las características de los puntos cercanos, generando un punto caliente estadísticamente significativo cuando éste tenga un valor alto y esté rodeado de otros puntos con valores altos, lo mismo ocurre para los puntos fríos los cuales se generan a partir de valores bajos rodeados de otros valores bajos.

La herramienta *Hot Spot analysis* utiliza la estadística Getis-Ord  $G_i^*$  para cada una de las características del conjunto de datos incidentes. De la estadística Getis-Ord  $G_i^*$  se obtienen puntuaciones Z resultantes y valores p que muestran donde se agrupan espacialmente los valores altos y bajos. Cuando las puntuaciones Z son altas y positivas es más intensa la agrupación de los valores altos (puntos calientes), por el contrario cuando las puntuaciones Z sean bajas y negativas será más intensa la agrupación de valores bajos (puntos fríos)<sup>7</sup>.

---

<sup>6</sup> ANSELIN, Luc. y REY, Sergio J. Perspectives on Spatial Data Analysis. En: Perspectives on Spatial Data Analysis. Tempe AZ USA, Springer, 2010. p. 20.

<sup>7</sup> GETIS, Arthur y ORD J. Keith, The analysis of spatial association by use of distance statistics. En: Geographical analysis, 1992, vol. 24, no 3, pp. 189-206.

La estadística de Getis-Ord  $G_i^*$  está dada por las Ecuaciones 1,2 y 3:

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\left[ \frac{n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2}{n-1} \right]}} \quad (1)$$

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (2)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2} \quad (3)$$

Donde  $x_j$  es el valor del atributo para la característica  $j$ ,  $w_{ij}$  es el peso espacial entre la característica  $i$  y  $j$ , y  $n$  es el número total de características.

Con el fin de estimar las zonas con necesidad de rehabilitación se realizó un análisis de puntos calientes para el conjunto de datos registrados en la zona metropolitana de Bucaramanga desde el año 2006 hasta el año 2013.

**2.1.1 Agrupamiento de los daños:** Los datos se encuentran proyectados en el sistema coordenado de MAGNA San Bucaramanga.

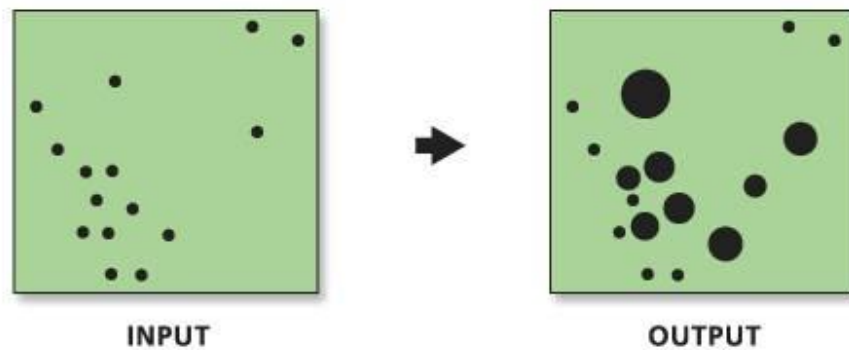
Las herramientas *Integrate* y *Collect Events* permiten agrupar los datos incidentes para ser utilizados en el análisis de puntos calientes con la herramienta *Hot Spot analysis* del software ArcGIS.

La herramienta *Integrate* permite integrar daños que tengan características compartidas, haciendo que los datos sean coincidentes si se encuentran dentro de

una tolerancia X y Y especificada<sup>8</sup>, esta distancia debe ser de un valor tal que permita la conformación de mínimo 30 grupos de daños. Para los datos objeto de estudio se asignó una tolerancia X y Y de 200 metros, distancia con la cual se conformaron un total de 69 grupos de daños.

La herramienta *Collect Events* permite crear una nueva clase de entidad la cual contiene un punto en cada lugar único con un atributo recuento asociado para indicar el número de eventos<sup>9</sup>.

Figura 3. Representación gráfica datos de entrada y salida herramientas Integrate - Collect Events.



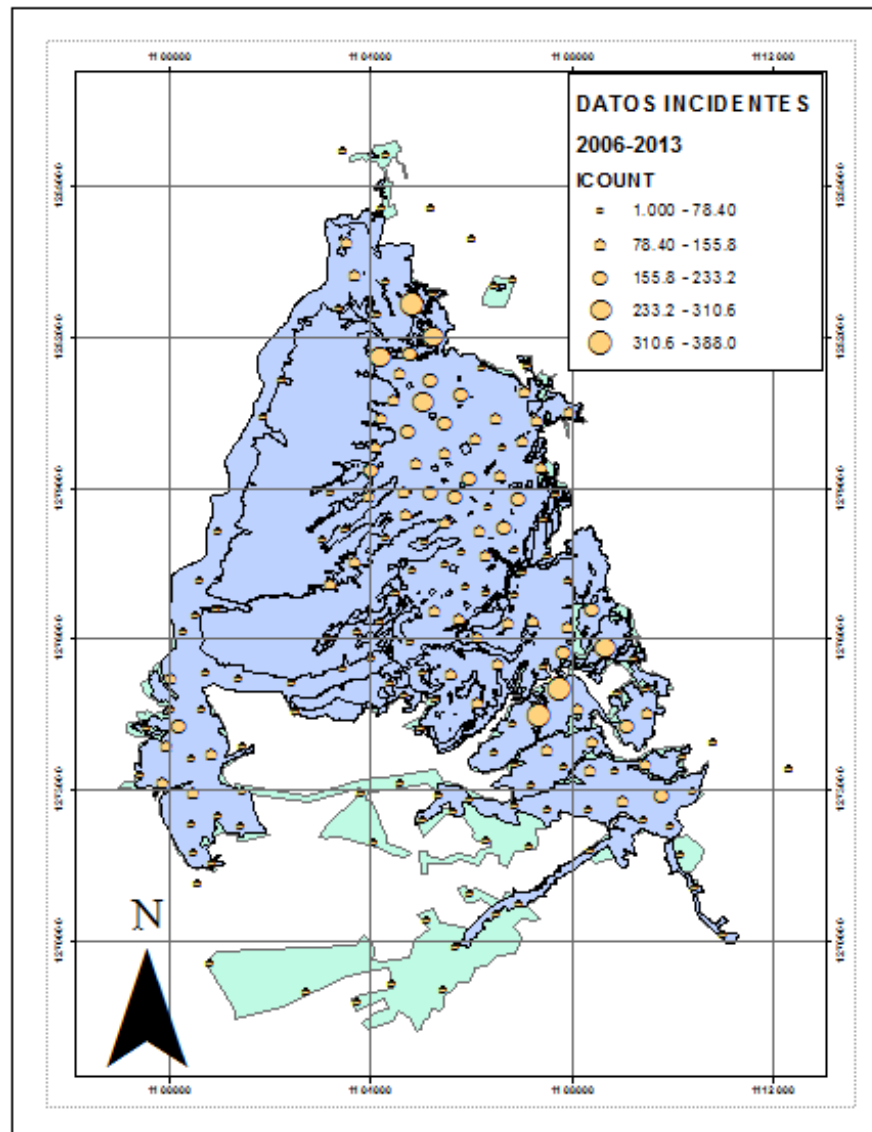
**Fuente:** ESRI. Recopilar Eventos (Estadística espacial) [en línea] Disponible en Internet: <http://resources.arcgis.com/es/help/main/10.1/index.html#//005p0000003s000000>

Los datos incidentes se obtuvieron como salida del proceso realizado con las herramientas *Integrate* y *Collect Events*. En la Figura 4 se observan los datos incidentes para los daños desde el año 2006 al año 2013.

<sup>8</sup> ESRI. Integrate (Data Management) [en línea]. ESRI, 1995-2013. [New York, E.U.] ESRI, 2013. [Citado: 22 de junio 2015] Disponible en Internet: <http://resources.arcgis.com/en/help/main/10.1/index.html#//00170000002s000000>

<sup>9</sup> ESRI. Collect Events (Spatial Statistics) [en línea]. ESRI, 1995-2013. [New York, E.U.] ESRI, 2013. [Citado: 22 de junio 2015] Disponible en Internet: <http://resources.arcgis.com/en/help/main/10.1/index.html#//005p0000003s000000>

Figura 4. Datos incidentes agrupados desde 2006 al año 2013.



**2.1.2 Incremental Spatial Autocorrelation:** En el momento de realizar el análisis de puntos calientes es necesario determinar una distancia de banda de la cual dependerá la calidad del análisis. Esta distancia de banda o distancia crítica

permitirá decidir qué puntos se encuentran “cercaños” y cuáles no, y con el análisis generar puntos calientes y puntos fríos<sup>10</sup>.

Esta herramienta permite estimar una distancia de banda que tenga sentido para el análisis cuando no se tiene una distancia base de los procesos espaciales que promueven el agrupamiento de los fenómenos que se están estudiando.

*Incremental Spatial Autocorrelation* mide la autocorrelación espacial para una serie de distancias de banda asumidas. También crea un gráfico de puntuaciones Z que reflejan la intensidad de la agrupación espacial correspondiente a una distancia. Los picos en este gráfico representan las distancias donde los procesos espaciales que promueven la agrupación son más pronunciados. Por lo cual estas distancias correspondientes a los picos son más apropiadas para ser utilizadas como distancia de banda<sup>11</sup>.

Para el caso de los datos incidentes desde el año 2006 al año 2013 se estimó la distancia de banda mediante esta herramienta obteniendo el gráfico mostrado en la Figura 5.

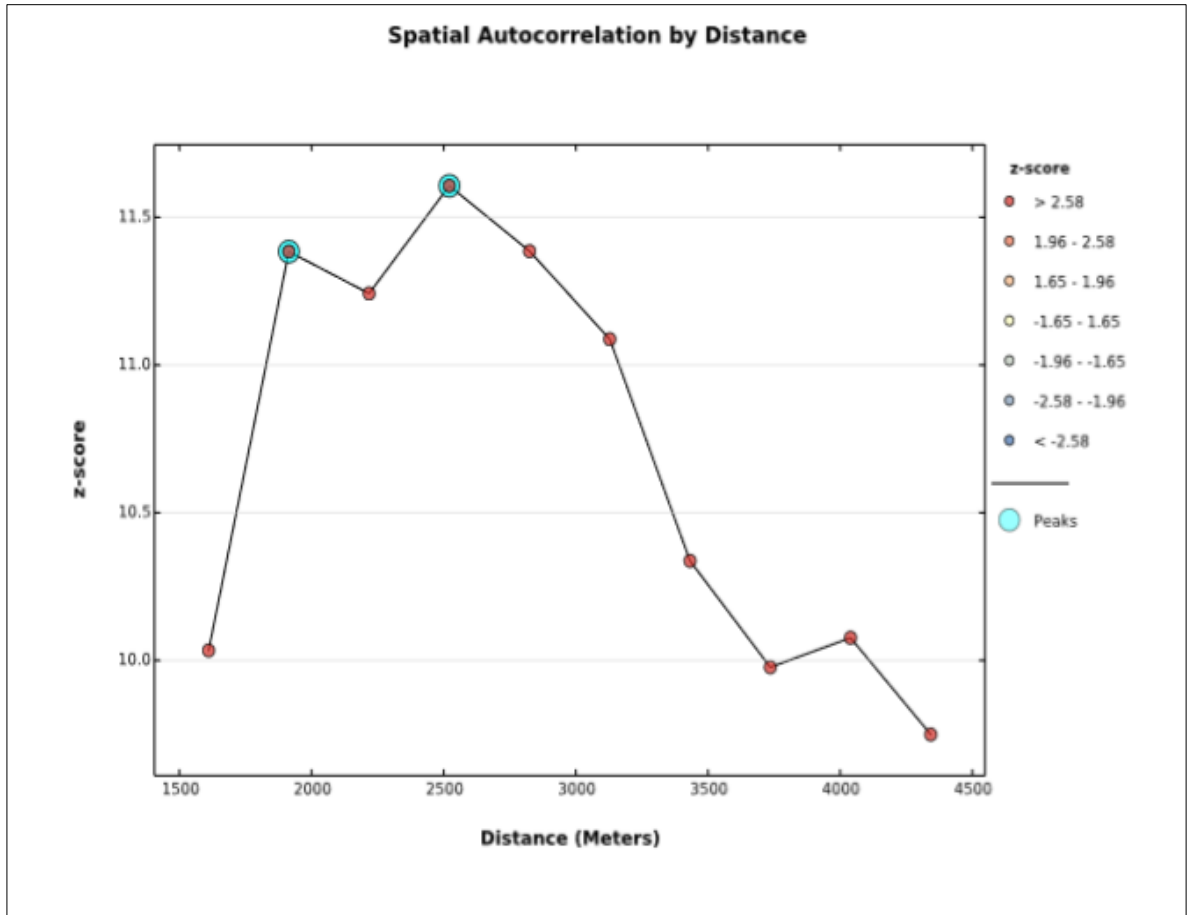
La herramienta *Incremental Spatial Autocorrelation* genera una gráfica en la cual se pueden visualizar dos picos, el primero correspondiente a una distancia de 1914.24 metros y el segundo a una distancia de 2521.34 metros. Esto indica que estas dos distancias promueven el agrupamiento de los datos.

---

<sup>10</sup> MITCHELL, Andy. Introducing Spatial Measurements and Statistics. En: The ESRI Guide to GIS Analysis, Volume 2: Spatial Measurements and Statistics, Redlans, 2005. p. 12.

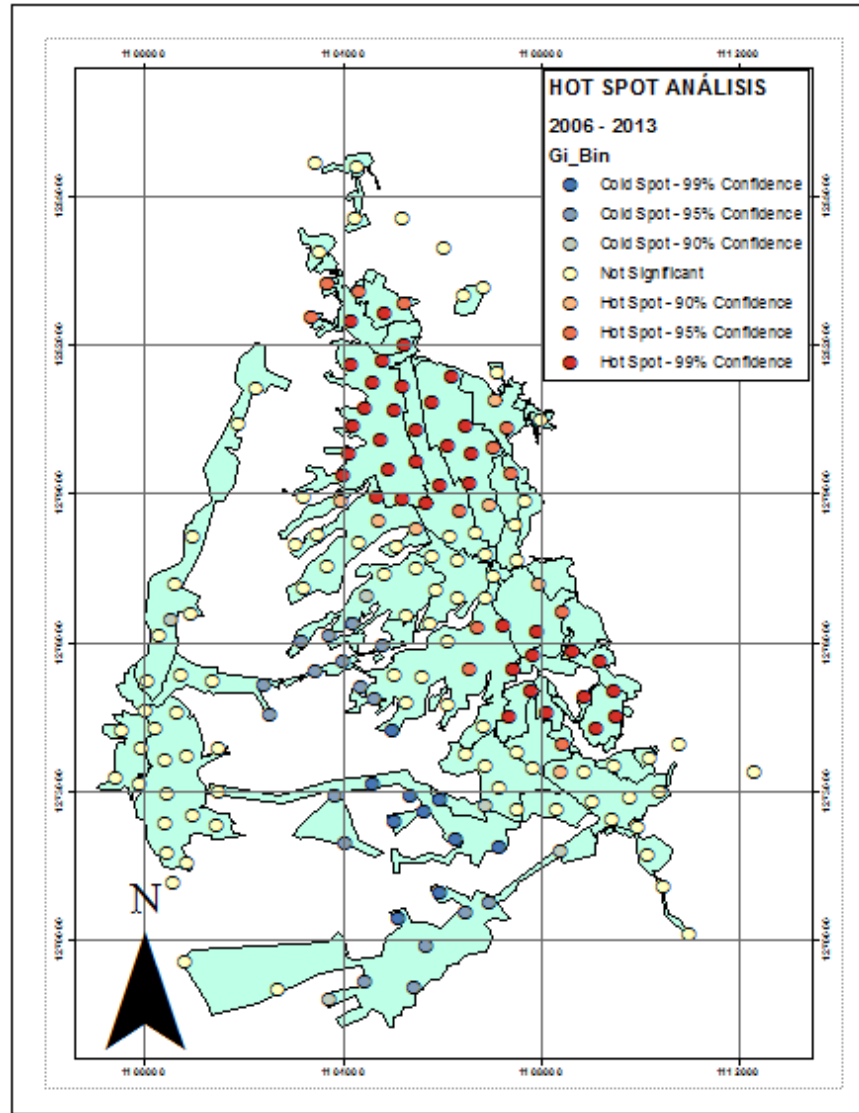
<sup>11</sup> ESRI. Incremental Spatial Autocorrelation (Spatial Statistics) [en línea]. ESRI, 1995-2013. [New York, E.U.] ESRI, 2013. [Citado: 22 de junio 2015] Disponible en Internet: <http://resources.arcgis.com/en/help/main/10.1/index.html#//005p0000004z000000>

Figura 5. Reporte generado por Incremental Spatial Autocorrelation para los datos agrupados desde el año 2006 hasta el año 2013



**2.1.3 Resultados del análisis de puntos calientes:** Una vez obtenida la distancia de banda se procedió a realizar el análisis de puntos calientes. En la Figura 6 se presenta el resultado del análisis realizado para los daños presentados desde el año 2006 hasta el año 2013. En la Figura 6 se aprecia el grado de confianza con el cual se determinaron tanto los puntos fríos como los puntos calientes, expresado en porcentaje.

Figura 6. Análisis de puntos calientes de los daños agrupados desde el año 2006 hasta el año 2013.

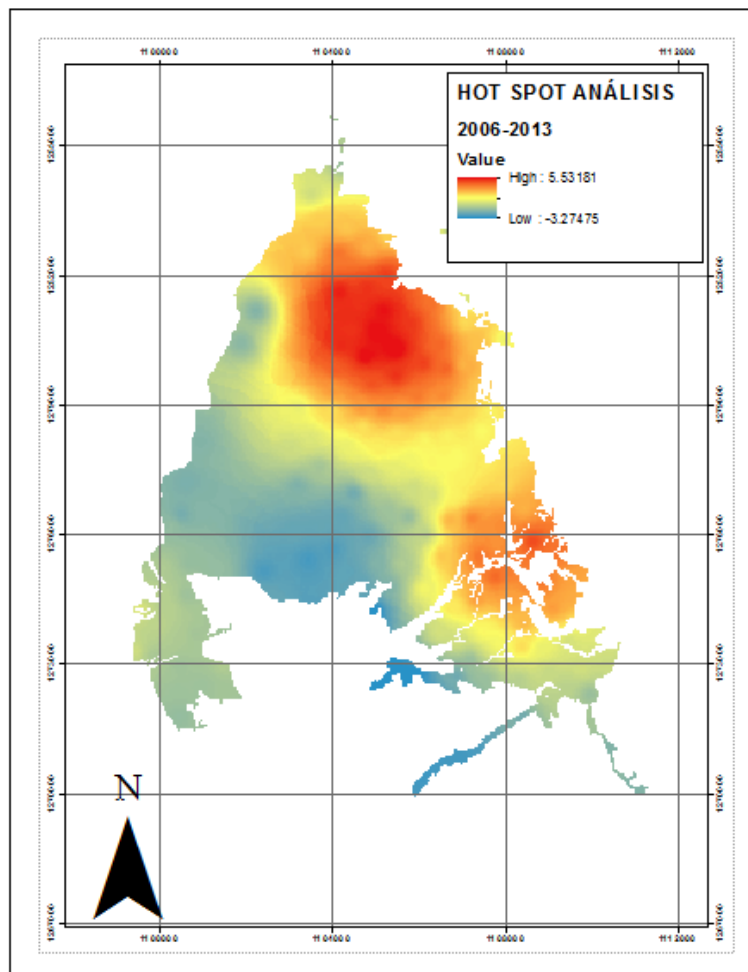


Con el fin de poder visualizar de mejor manera los resultados generados por la herramienta *Hot Spot Analysis* se hace necesario emplear un método de interpolación. Se empleó la herramienta IDW (Distancia Inversa Ponderada) la cual permite crear una superficie interpolada a partir de una serie de puntos usando la

técnica de distancia inversa ponderada<sup>12</sup>. También se realiza un *Mask* con base en el perímetro del área metropolitana de Bucaramanga. Un *Mask* es un recorte perimetral de manera que nos muestre sólo el área de interés.

En la Figura 7 se presenta la visualización del análisis realizado para los daños presentados desde el año 2006 hasta el año 2013.

Figura 7. Visualización del análisis de puntos calientes de los todos daños presentados desde el año 2006 hasta el año 2013.



<sup>12</sup> LONDOÑO C; LIBARDO A.; RESTREPO C. y MARULANDA E. Distribución espacial del dengue basado en herramientas del Sistema de Información Geográfica. En: Revista Facultad Nacional de Salud Pública. Enero - Abril 2014, vol. 32, p. 9.

## 2.2 Distritos con mayor prioridad de rehabilitación

El área metropolitana de Bucaramanga se encuentra dividida en 34 distritos hidráulicos, los cuales son áreas o sectores de la red de tuberías establecidos por el amb S.A. E.S.P. que pueden operar de manera aislada.

Un análisis visual del IDW no permite identificar con certeza los distritos con mayor prioridad de rehabilitación, por lo cual se hace necesario el empleo de la Fórmula 4 la cual permite calcular el puntaje Z total para cada distrito.

$$f(\bar{Z}(Gi^*)) = \bar{Z}(Gi^*) * n \quad (4)$$

Donde  $\bar{Z}(Gi^*)$  representa el promedio de las puntuaciones Z y n el número total de puntos calientes o fríos para cada distrito<sup>13</sup>.

Se estableció una mayor prioridad para aquellos distritos que presentan un mayor valor de  $f(\bar{Z}(Gi^*))$ . En la Tabla 9 se muestra la prioridad para cada distrito.

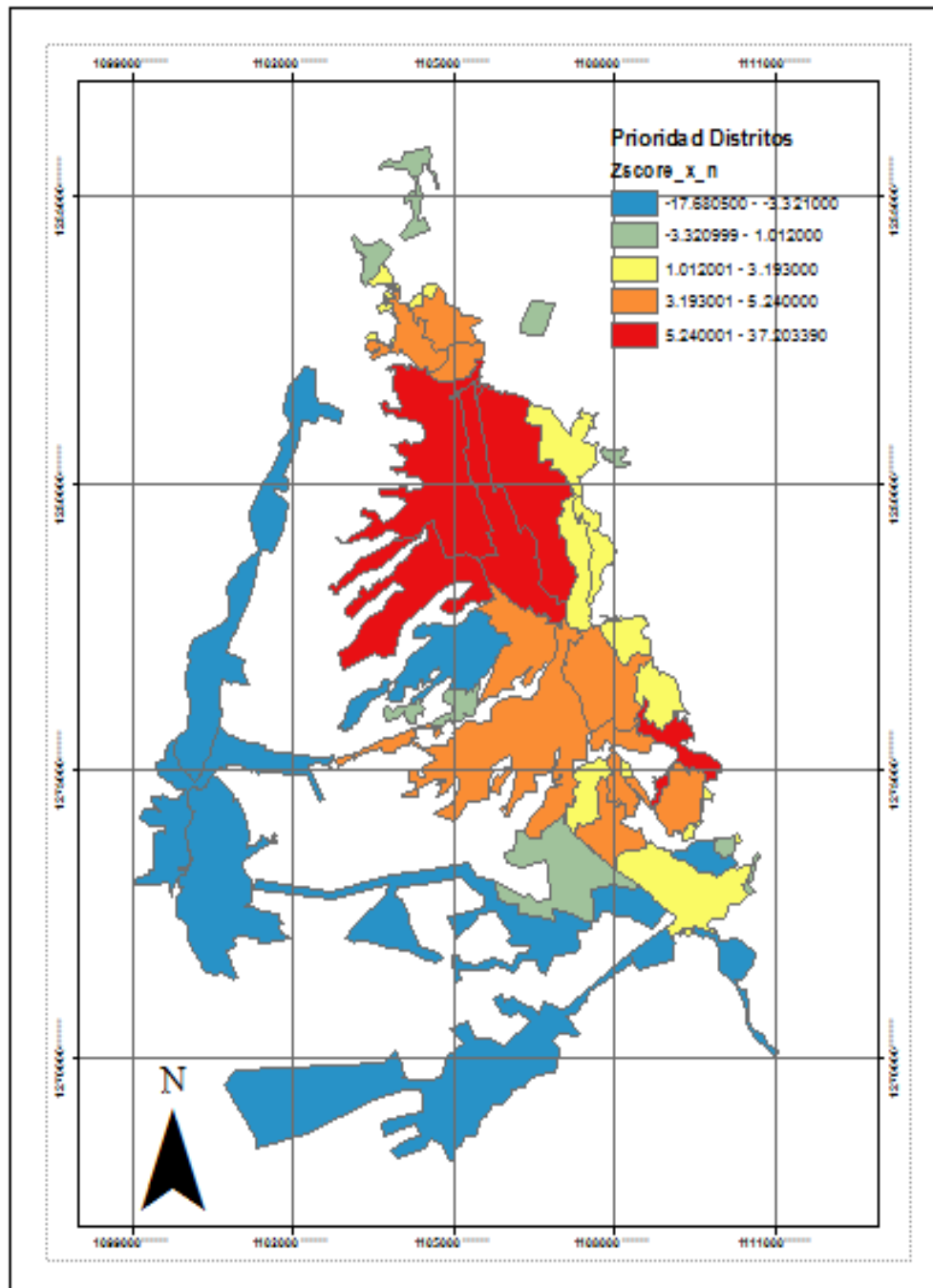
---

<sup>13</sup> CASTAÑEDA GALVIS, Luis Fernando. Modelo de gestión para el mantenimiento de tuberías en redes de distribución de agua potable. Trabajo de Grado presentado como opción parcial para optar al título de Magister en Tecnologías de la Información Geográfica. Manizales. Universidad de Manizales. Facultad de Ingeniería, 2015. 49 p.

Tabla 9. Prioridad de rehabilitación por distrito.

<b>Distrito</b>	<b>n</b>	<b><math>f(\bar{Z}(Gi^*))</math></b>	<b><math>\bar{Z}(Gi^*)</math></b>	<b>P</b>
<b>Estadio</b>	9	4.1337	37.20330	1
<b>Norte</b>	7	2.9534	20.67380	2
<b>Morro Alto</b>	4	2.8626	11.45040	3
<b>El Carmen</b>	3	3.1948	9.58440	4
<b>Centro</b>	6	0.8757	5.25420	5
<b>La Cumbre</b>	2	2.2956	4.59120	6
<b>Tejar</b>	2	2.2816	4.56320	7
<b>Bienestar</b>	2	2.1519	4.30380	8
<b>Villabel Alto</b>	2	1.9954	3.99080	9
<b>Morro Bajo</b>	3	1.3263	3.97890	10
<b>Regadero</b>	1	3.9606	3.96060	11
<b>Zapamanga</b>	1	3.7305	3.73050	12
<b>Malpaso</b>	14	0.2565	3.59100	13
<b>La Esperanza</b>	2	1.6967	3.39340	14
<b>Cabecera</b>	2	1.5960	3.19200	15
<b>Trinidad</b>	1	2.1329	2.13290	16
<b>Villabel Bajo</b>	1	1.8041	1.80410	17
<b>Florida</b>	4	0.4368	1.74720	18
<b>Ferrovias</b>	1	1.5076	1.50760	19
<b>Pan de Azúcar</b>	1	1.4190	1.41900	20
<b>La Flora</b>	1	1.2721	1.27210	21
<b>Angelinos</b>	1	1.0113	1.01130	22
<b>Buenos Aires</b>	1	0.8441	0.84410	23
<b>Caracolí</b>	1	0.3110	0.31100	24
<b>Café Madrid</b>	1	-0.2273	-0.22730	25
<b>Cañaveral</b>	3	-0.1812	-0.54360	26
<b>La Iglesia</b>	1	-1.4507	-1.45070	27
<b>Colorados</b>	2	-1.1960	-2.39200	28
<b>San Juan</b>	5	-0.6644	-3.32200	29
<b>Puerta_del_Sol</b>	3	-1.2000	-3.60000	30
<b>Bucarica</b>	5	-0.9709	-4.85450	31
<b>Sin distrito</b>	13	-0.4095	-5.32350	32
<b>Girón</b>	15	-0.7254	-10.8810	33
<b>Ruitoque</b>	9	-1.9207	-17.2863	34
<b>Río Frio</b>	9	-1.9645	-17.6805	35

Figura 8. Visualización prioridad de rehabilitación de los distritos.



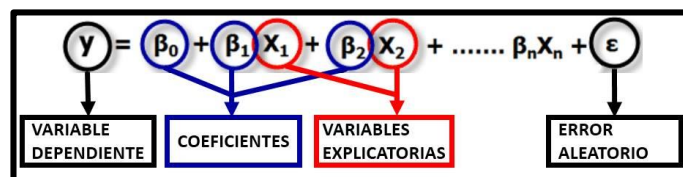
### 3. ESTIMACIÓN DE LOS FACTORES QUE INFLUYEN EN LA PRESENCIA DE DAÑOS CON BASE EN ANÁLISIS DE REGRESIÓN

#### 3.1 ANÁLISIS DE REGRESIÓN

Un análisis de regresión permite estimar, modelar y explorar las relaciones espaciales existentes entre un grupo de variables, generando un modelo con el cual se puedan representar patrones espaciales observados y predecir resultados<sup>14</sup>.

El análisis de regresión múltiple es una técnica de análisis multivariable en el que se establece una relación funcional entre una variable dependiente o a explicar y una serie de variables independientes o explicativas, en la que se estiman los coeficientes de regresión que determinan el efecto que las variaciones de las variables independientes tienen sobre el comportamiento de la variable dependiente<sup>15</sup>. La forma general de representar este modelo es:

Figura 9. Ecuación análisis de regresión.



**Fuente:** ESRI. Regression analysis basics [en línea], Disponible en Internet:

[http://resources.arcgis.com/en/help/main/10.2/index.html#/Regression\\_analysis\\_basics/005p0000023000000/](http://resources.arcgis.com/en/help/main/10.2/index.html#/Regression_analysis_basics/005p0000023000000/)

<sup>14</sup> ABDELWAHAB M. Bubiya; AHMED H. y OTHMAN Jaafar. Performance improvement for pipe breakage prediction modeling using regression method. En: International Journal of the Physical Sciences, Octubre, 2011, p. 6027.

<sup>15</sup> REQUENA LAVIÑA, María. El desarrollo de las marcas gestionadas por la distribución. Análisis de las variables relevantes. Tesis de Grado Doctor. Madrid. Universidad Pontificia Comillas de Madrid. Facultad de Ciencias Económicas y Empresariales (ICADE). 2015. 126-127 p.

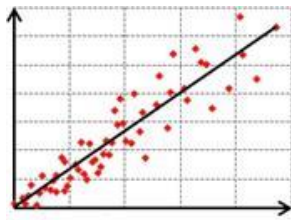
- **Variable dependiente:** Variable cuyo valor depende de una u otras variables independientes. Representa el fenómeno o proceso el cual se busca modelar.
- **Variable independiente:** Variable que determina el comportamiento de la variable dependiente. Representa las causas del fenómeno que se busca modelar. También es conocida como variable explicativa.
- **Coefficiente de regresión:** valor que representa el tipo y peso de la relación entre la variable dependiente y la variable explicativa.
- **Residuo ( $\epsilon$ ):** Porción de la variable dependiente que el modelo no puede explicar, es denominado normalmente como un término de error aleatorio.

Las relaciones lineales establecidas mediante un análisis de regresión pueden ser positivas o negativas. Cuando el valor de la variable explicativa aumenta y ésta genera un aumento a su vez del valor de la variable dependiente, se dice que existe una correlación positiva. Por el contrario si genera una disminución en el valor de la variable dependiente hablamos de una relación negativa. El signo asociado al coeficiente que acompaña cada variable explicativa (+/-) identifica el tipo de relación<sup>16</sup>.

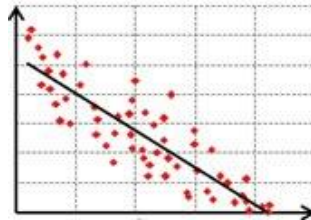
---

<sup>16</sup> ESRI HEALTH GIS CONFERENCE. Modeling Spatial Relationships Using Regression Analysis (20 -22 de Agosto, 2012: Scottsdale Az, USA). pp. 1-6.

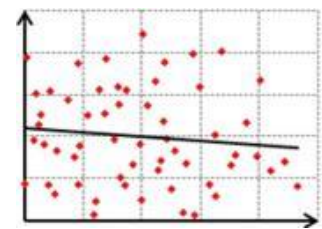
Figura 10. Tipos de relación



a) Relación positiva



b) Relación negativa



c) Sin relación

La estimación de los factores influyentes en la presencia de daños en las redes de distribución de agua potable se llevó a cabo mediante un análisis de regresión OLS “*Ordinary Least Squares*”.

El método de mínimos cuadrados es una técnica de análisis numérico cuyo objetivo se basa en minimizar la suma variaciones cuadradas entre el valor real de la variable dependiente y el valor determinado por el modelo. Normalmente se encuentran diferencias entre el valor real y el determinado por el modelo, a esta variación o error se le conoce como residuo.

**3.1.1 Factores influyentes considerados para el análisis de regresión:** La presencia de daños en las redes de distribución obedece comúnmente a múltiples factores.

Los mecanismos físicos que conducen a la rotura de una tubería implican 3 aspectos principales.

- a) Propiedades estructurales de la tubería, tipo de material, interacción tubería suelo y calidad de la instalación.
- b) Cargas internas debidas a la presión operativa, cargas externas debido a la sobrecarga del suelo, cargas de tráfico y la interferencia de terceros.

c) El deterioro del material debido a la interacción química interna y externa<sup>17</sup>.

Por tal motivo y por disponibilidad de la información se consideró el material, edad, diámetro y presión de los elementos de conducción para llevar a cabo el análisis de regresión.

Con el fin de determinar qué tipo de material, qué grupo de edad, diámetro y presión son los más influyentes, se realizó la siguiente subdivisión de cada factor apreciada en la Tabla 10.

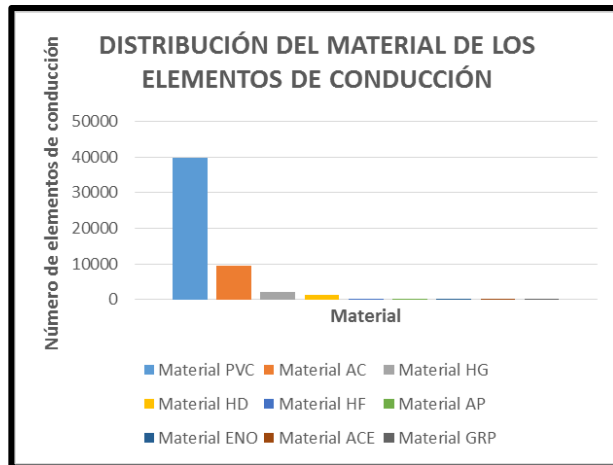
Tabla 10. Factores considerados para el análisis de regresión.

<b>Material</b>	Policloruro de vinilo	PVC
	Asbesto cemento	AC
	Hierro galvanizado	HG
	Hierro dúctil	HD
<b>Edad</b>	$e \leq 10$	[Años]
	$10 < e \leq 20$	
	$20 < e \leq 30$	
	$30 < e \leq 40$	
	$e \geq 40$	
<b>Diámetro</b>	$d \leq 3$	[Pulg]
	$3 < d \leq 6$	
	$6 < d \leq 10$	
	$d \geq 10$	
<b>Presión</b>	$p \leq 15$	[m.c.a.]
	$15 < p \leq 25$	
	$25 < p \leq 35$	
	$35 < p \leq 50$	
	$p \geq 50$	

<sup>17</sup> RAJANI B.B. and KLEINER Y. Comprehensive review of structural deterioration of water mains: physically based models. En: NRCC-43722. NRC-CNRC. 2001. p. 2.

La distribución de los elementos de conducción en base a los grupos de material, edad, diámetros y presión se representa en las Figuras 11, 12, 13 y 14 respectivamente.

Figura 11. Distribución del factor material.



Los materiales seleccionados para llevar a cabo el análisis representan el 98,83% del número total de tubos que conforman la red de distribución de agua potable.

Figura 12. Distribución del factor edad.

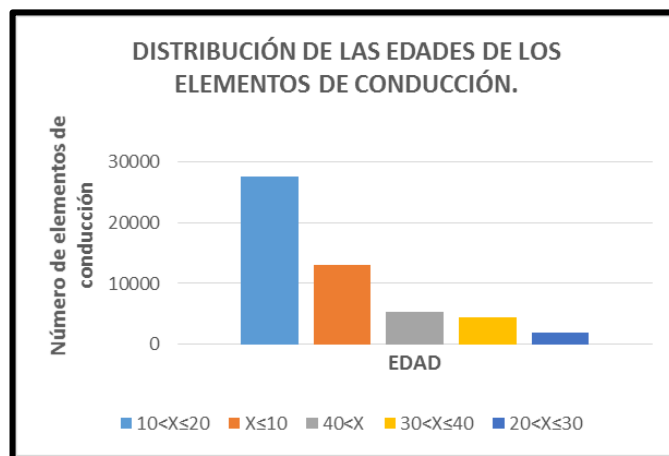


Figura 13. Distribución del factor diámetros.

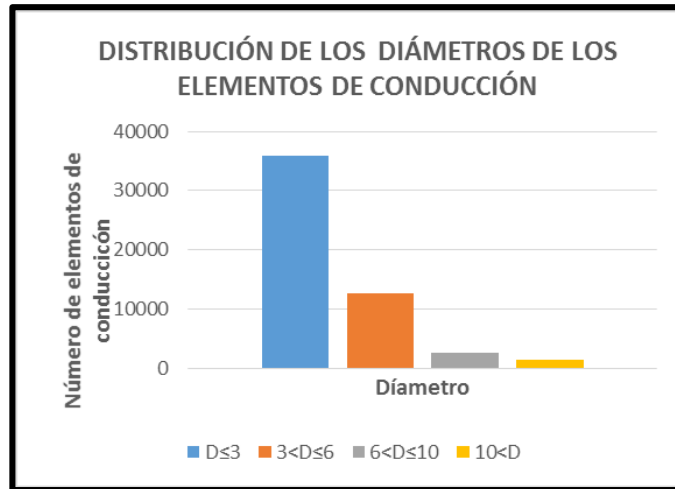
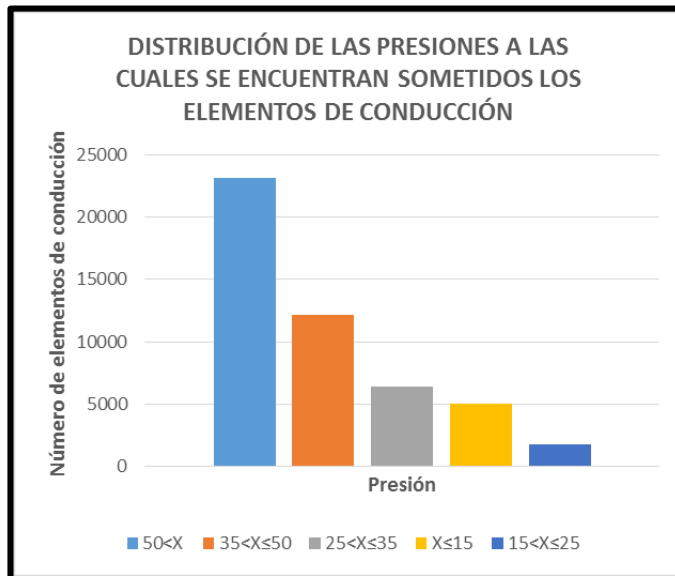


Figura 14. Distribución del factor presiones.

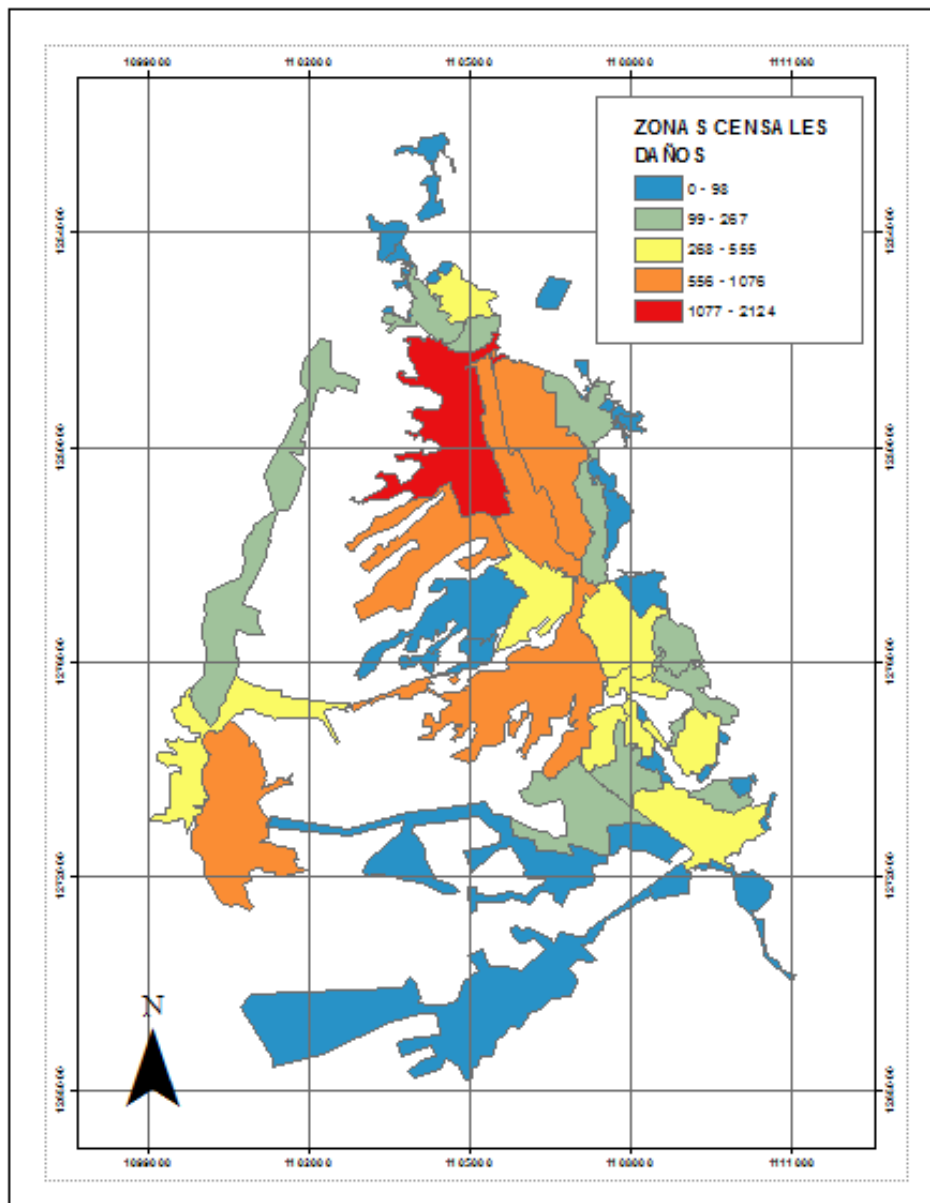


**3.1.2 Zonas censales:** Con el objetivo de realizar el análisis de regresión se establecieron zonas censales a las cuales se les asignó una variable contable llamada “DAÑOS”. Esta variable representa el número de daños ocurridos en las

redes de distribución de agua potable dentro de las zonas censales. La variable “DAÑOS” se identificó como la variable dependiente.

Las zonas censales se definieron como los distritos hidráulicos del acueducto metropolitano de Bucaramanga.

Figura 15. Visualización zonas censales.



## 3.2 VALIDACIÓN DE LOS RESULTADOS

Es necesario realizar las siguientes validaciones de los modelos generados por los análisis de regresión con el fin de confiar y utilizar sus resultados.

- Revisar si el coeficiente que acompaña cada variable explicativa tiene el signo esperado, el cual indica la relación positiva o negativa con la variable dependiente.
- Revisar si las variables explicativas son estadísticamente significativas. Si la prueba Koenker es estadísticamente significativa se deben utilizar las probabilidades robustas para evaluar la significancia estadística. (Las probabilidades significativas tienen un asterisco al lado de ellas (\*)). Se deben remover del modelo aquellas variables no estadísticamente significativas.
- Revisar que la prueba Jarque-Bera no sea estadísticamente significativa. Si la prueba Jarque-Bera es estadísticamente significativa significa que le falta una o más variables explicativas claves al modelo.
- Revisar el ajuste del modelo mediante el coeficiente  $R^2$ . Se buscan valores entre 0.5 y 1.
- Verificar que los residuos del modelo no estén espacialmente correlacionados con el uso de la herramienta *Spatial Autocorrelation (Morgen's I)*.

## 3.3 ANÁLISIS DE RESULTADOS DE REGRESIÓN OLS

**3.3.1 Factor material.** En el Anexo A se observan los resultados numéricos del análisis de regresión OLS para el factor material. La ecuación 5 representa el modelo generado.

$$DAÑOS = 15.1672 - 1.3289(HD) + 3.9125(AC) - 3.0706(HG) - 0.4105(PVC) \quad (5)$$

### Validación factor material

- El material con mayor coeficiente es el asbesto cemento (AC) con un valor de 3.9125. Las demás variables explicativas presentaron una relación negativa.
- La prueba Koenker es estadísticamente significativa sin embargo a partir de probabilidad robusta no se identificó ninguna variable explicativa estadísticamente significativa.
- La prueba Jarque-Bera es estadísticamente significativa.
- El modelo presento un valor  $R^2$  ajustado de 0.8041, lo cual indica que el modelo explica el 80,41% del número de daños.
- En el Anexo E se muestra que los residuos de la regresión OLS se encuentran aleatoriamente distribuidos.

**3.3.2 Factor edad.** En el Anexo B se observan los resultados numéricos del análisis de regresión OLS para el factor edades. La ecuación 6 representa el modelo generado.

$$DAÑOS = 3.1928 + 0.7611(e \leq 10) + 0.0422(10 < e \leq 20) + 0.1724(20 < e \leq 30) + 0.1904(30 < e \leq 40) + 0.17777 (p \geq 50) \quad (6)$$

### Validación factor edad.

- El grupo de edades con mayor coeficiente es el comprendido por edades menores o iguales a 10 años con un valor de 0.7611. Las demás variables explicativas presentaron una relación positiva.
- La prueba Koenker es estadísticamente significativa sin embargo a partir de probabilidad robusta solo se identificó como variable estadísticamente significativa el grupo de edades menores o iguales a 10 años.
- La prueba Jarque-Bera es estadísticamente significativa.

- El modelo presento un valor  $R^2$  ajustado de 0.8296, lo cual indica que el modelo explica el 82,96 % del número de daños.
- En el Anexo F se muestra que los residuos de la regresión OLS se encuentran aleatoriamente distribuidos.

**3.3.3 Factor diámetro.** En el Anexo C se observan los resultados numéricos del análisis de regresión OLS para el factor diámetro. La ecuación 7 representa el modelo generado.

$$DAÑOS = 2.0927 + 0.1329(d \leq 3) + 0.7059 (3 < d \leq 6) + 1.0456(6 < d \leq 10) - 2.6919(d \geq 10) \quad (7)$$

**Validación factor diámetro.**

- El grupo de diámetros con mayor coeficiente es el comprendido por diámetros mayores que 6 y menores o iguales a 10 pulgadas con un valor de 1.04563. Los diámetros mayores a 10 pulgadas presentar una relación negativa, las demás variables explicativas presentaron una relación positiva.
- La prueba Koenker es estadísticamente significativa sin embargo a partir de probabilidad robusta solo se identificó como variable estadísticamente significativa el grupo comprendido por diámetros mayores que 6 y menores o iguales a 10 pulgadas.
- La prueba Jarque-Bera es estadísticamente significativa.
- El modelo presento un valor  $R^2$  ajustado de 0.7852, lo cual indica que el modelo explica el 78,52 % del número de daños.
- En el Anexo G se muestra que los residuos de la regresión OLS se encuentran aleatoriamente distribuidos.

**3.3.4 Factor presión:** En el Anexo D se observan los resultados numéricos del análisis de regresión OLS para el factor presión. La ecuación 8 representa el modelo generado.

$$\begin{aligned} \text{DAÑOS} = & -15.0671 + 0.5615(p \leq 15) + 0.2110(15 < p \leq 25) \\ & -0.2871(25 < p \leq 35) + 0.5378(35 < p \leq 50) + 0.2221(p \geq 50) \end{aligned} \quad (8)$$

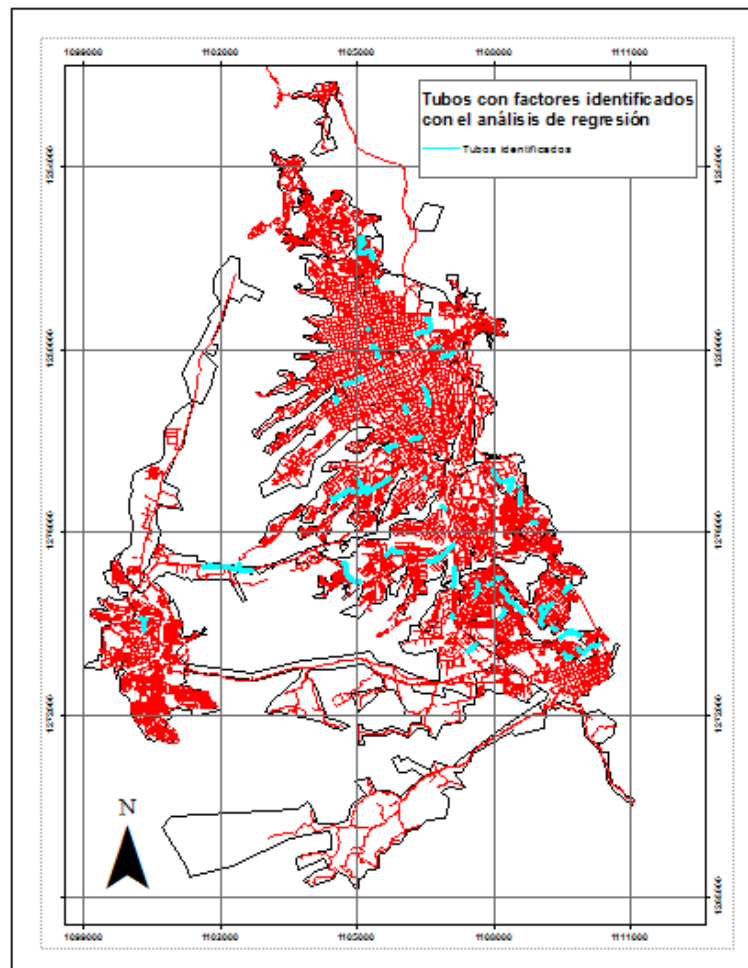
#### **Validación factor presión.**

- Los dos grupos de presiones con mayor coeficiente son el grupo conformado por presiones menores que 15 m.c.a. y el grupo conformado por presiones mayores que 35 y menores o iguales a 50 m.c.a. con un valores de 0.56149 y 0.53777 respectivamente. Las presiones mayores a 25 y menores o iguales a 35 m.c.a. presenta una relación negativa, las demás variables explicativas presentaron una relación positiva.
- La prueba Koenker es estadísticamente significativa sin embargo a partir de probabilidad robusta solo se identificó como variable estadísticamente significativa el grupo comprendido por presiones mayores que 50 m.c.a.
- La prueba Jarque-Bera es estadísticamente significativa.
- El modelo presento un valor  $R^2$  ajustado de 0.7563, lo cual indica que el modelo explica el 75,63 % del número de daños
- En el Anexo H se muestra que los residuos de la regresión OLS se encuentran aleatoriamente distribuidos.

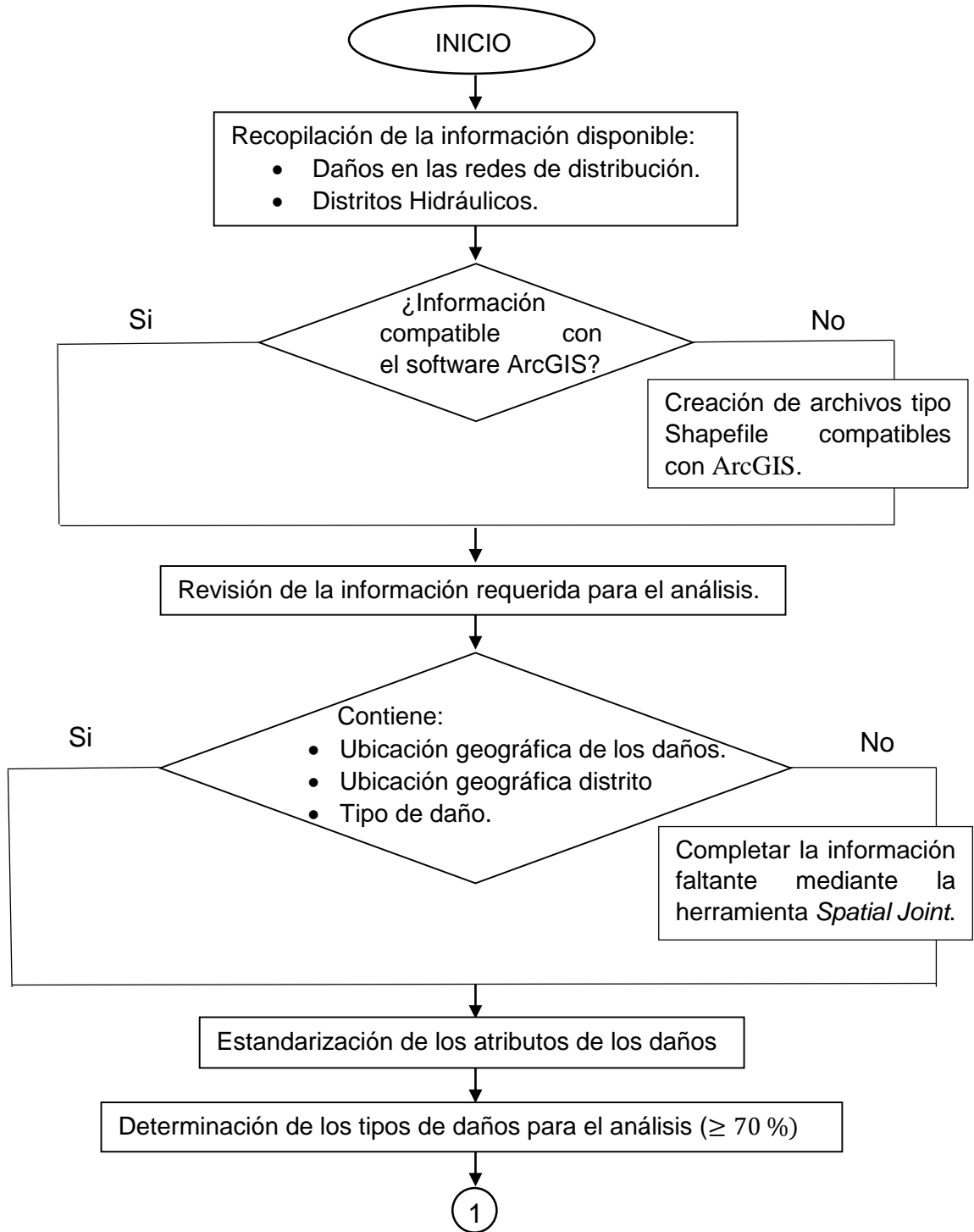
### 3.4 IDENTIFICACIÓN DE TUBOS CON FACTORES IDENTIFICADOS MEDIANTE EL ANÁLISIS DE REGRESIÓN

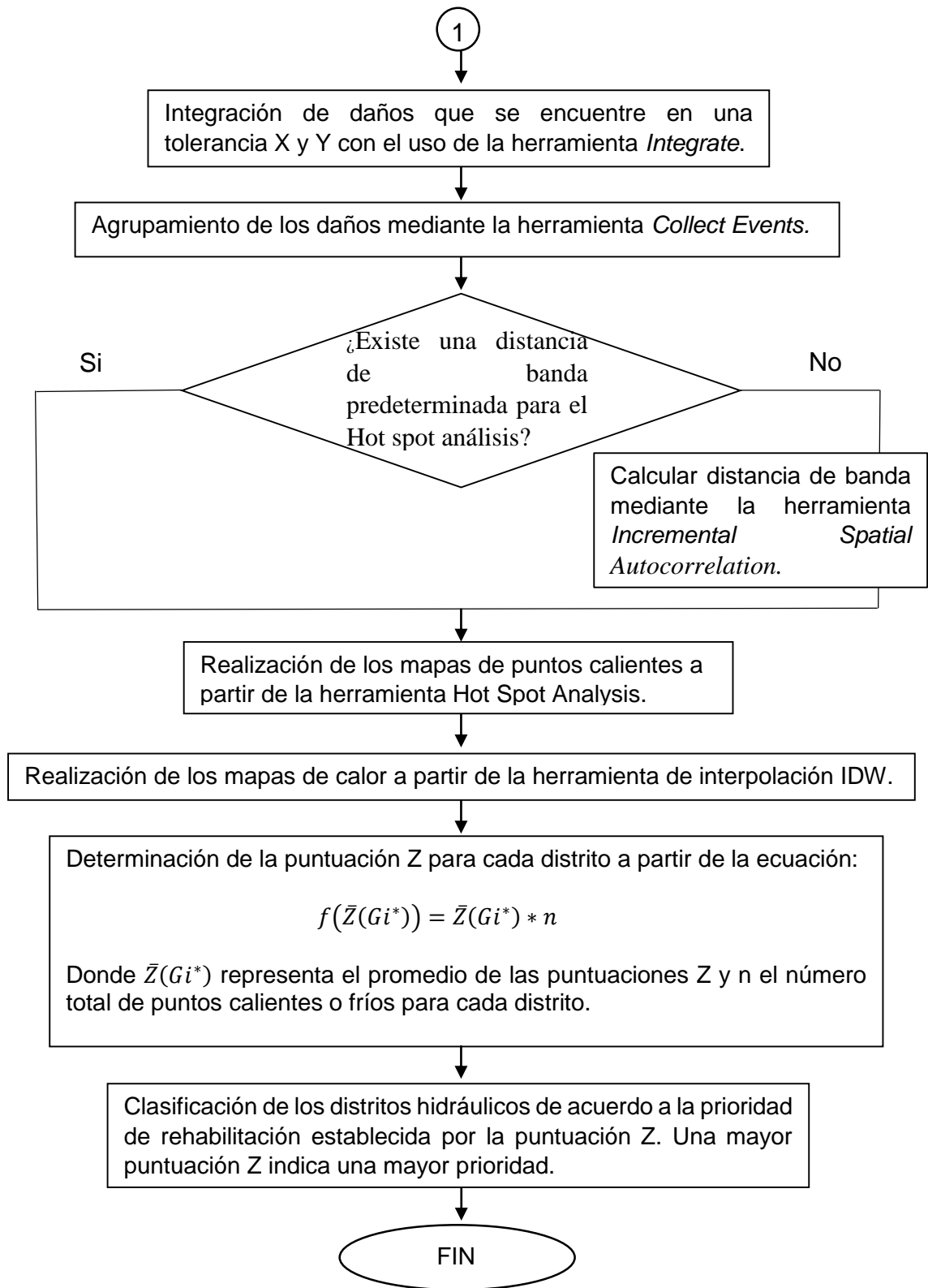
En la Figura 16 se presentan los elementos de conducción de la red de distribución del área metropolitana de Bucaramanga, los cuales presentan los factores de mayor influencia en la presencia de daños, identificados mediante el análisis de regresión. Se observó que los resultados de este análisis se relacionan con los resultados obtenidos mediante el análisis de puntos calientes puesto que los elementos de conducción identificados se encuentran ubicados en su mayoría en los distritos con mayor prioridad de rehabilitación.

Figura 16. Tubos con factores identificados con el análisis de regresión.



#### 4. METODOLOGÍA PARA LA ESTIMACIÓN DE ZONAS CON NECESIDAD REHABILITACIÓN DE LAS REDES DE DISTRIBUCIÓN DE AGUA POTABLE





## 5. CONCLUSIONES

Los 5 distritos identificados con mayor prioridad de rehabilitación a partir del análisis de puntos calientes nombrados de manera descendente son: Distrito el Estadio, Distrito Norte, Distrito Morro Alto, Distrito el Carmen y Distrito centro.

Se plantea una metodología para la estimación de zonas con necesidad de rehabilitación adaptable a cualquier sistema de distribución de agua potable a partir de información histórica mediante la implementación del análisis de puntos calientes.

Los factores más influyentes en la presencia de daños determinados a partir de los análisis de regresión fueron:

- Material: asbesto cemento (AC)
- Edad: menores o iguales a 10 años.
- Diámetro: mayores que 6 y menores o iguales a 10 pulgadas
- Presión: Menores que 15 m.c.a. y presiones mayores que 35 y menores o iguales a 50 m.c.a.

Se identificó el grupo de edades menores o iguales a 10 años como el grupo más influyente del factor edad. Se plantean dos hipótesis:

- Según RAJANI Y KLEINER (2001) el ciclo de vida de las tuberías puede finalizar en una de 3 fases. La primera de estas se denomina ajuste y corresponde al periodo después de la instalación, los daños se presentan debido a la baja calidad de la instalación y a problemas propios de la tubería. Los daños en la fase de ajuste podría ser la explicación a los resultados arrojados por este análisis.

- Cierta información correspondiente a las edades de los elementos de conducción en los cuales se presentaron los daños fue completada mediante el uso de relaciones espaciales. Lo cual pudo incidir en los resultados de este modelo, afectando la calidad del mismo.

En cada uno de los 4 análisis de regresión realizados la prueba Jarque-Bera fue estadísticamente significativa. Esto indica que para cada uno de los modelos hacen falta variables claves explicativas, este es un resultado esperado debido a que como se mencionó anteriormente en el subtítulo 3.1.1, la rotura de tuberías obedecen a múltiples factores y los modelos fueron realizados a partir de un único factor.

La calidad de los análisis de regresión y por ende de sus resultados depende de la información histórica recopilada y de los estándares de almacenamiento con los que se encuentre.

La utilización de herramientas de análisis espacial enfocadas al desarrollo de estrategias para la aplicación actividades de mantenimiento predictivo/preventivo en las redes de distribución permite la optimización del servicio prestado.

## BIBLIOGRAFIA

ABDELWAHAB M. Bibtiena; AHMED H. y OTHMAN Jaafar. Performance improvement for pipe breakage prediction modeling using regression method, En: International Journal of the Physical Sciences, Octubre, 2011, p. 6025-6035.

ANSELIN, Luc. y REY, Sergio J. Perspectives on Spatial Data Analysis. USA, Springer, 2010. 290 p.

ARAQUE, Diego Alejandro. Optimización de Redes de Acueducto con el fin de Uniformizar el Estado de presiones. Trabajo de grado Maestría. Bogotá D.C: Universidad de los Andes. Facultad de Ingeniería. Departamento de Ingeniería Civil y Ambiental, 2004.

CASTAÑEDA GALVIS, Luis Fernando. Modelo de gestión para el mantenimiento de tuberías en redes de distribución de agua potable. Trabajo de Grado presentado como opción parcial para optar al título de Magister en Tecnologías de la Información Geográfica. Manizales. Universidad de Manizales. Facultad de Ingeniería, 2015.

DÍAZ VIERA, Martín Alberto. Geostatística aplicada. Cuba, 2002. 131 p.

ESRI. ArcGIS Resources, ArcGIS Help 10.1. [en línea]. ESRI, 1995-2013 [citado 22 de Junio de 2015]. Disponible en: <http://resources.arcgis.com/en/help/main/10.1/>

ESRI HEALTH GIS CONFERENCE. Modeling Spatial Relationships Using Regression Analysis (20 -22 de Agosto, 2012: Scottsdale Az, USA). 36 p.

ESRI. ESRI Shapefile Technical Description. ESRI, 1995-2013. New York, E.U, July 1998, p. 30.

GETIS, Arthur y ORD J. Keith, The analysis of spatial association by use of distance statistics. En: Geographical analysis, 1992, vol. 24, no 3, p. 189-206.

LONDOÑO C; LIBARDO A.; RESTREPO C. y MARULANDA E. Distribución espacial del dengue basado en herramientas del Sistema de Información Geográfica. En: Revista Facultad Nacional de Salud Pública. Enero - Abril 2014, vol. 32, p. 7-15.

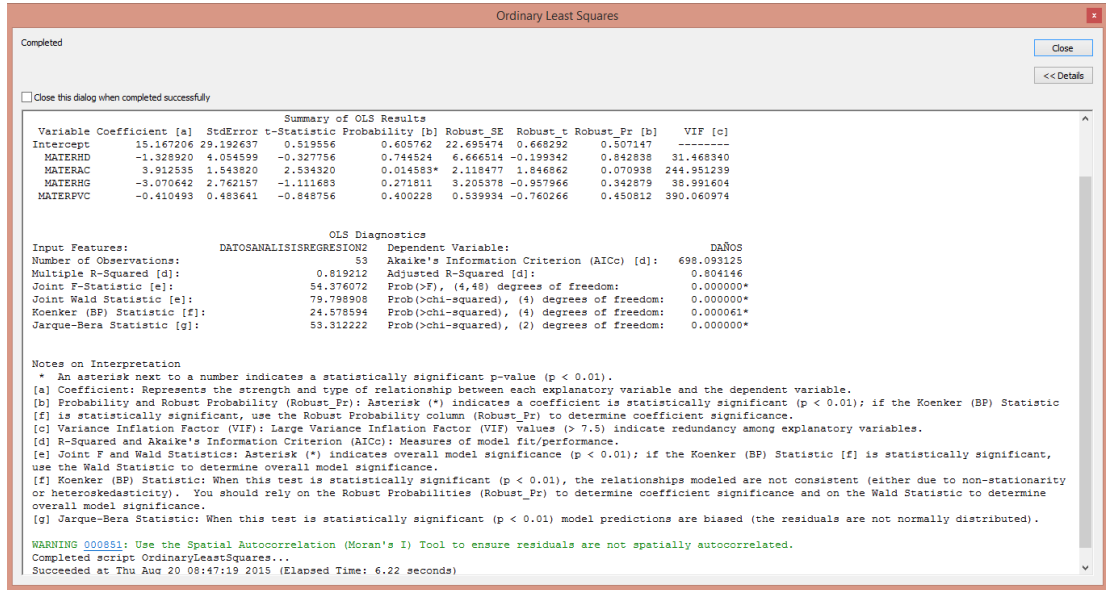
MITCHELL, Andy. The ESRI Guide to GIS Analysis, Volume 2: Spatial Measurements and Statistics, Redlans, 2005.

RAJANI B.B. and KLEINER Y. Comprehensive review of structural deterioration of water mains: physically based models. En: NRCC-43722. NRC-CNRC. 2001. 32 p.

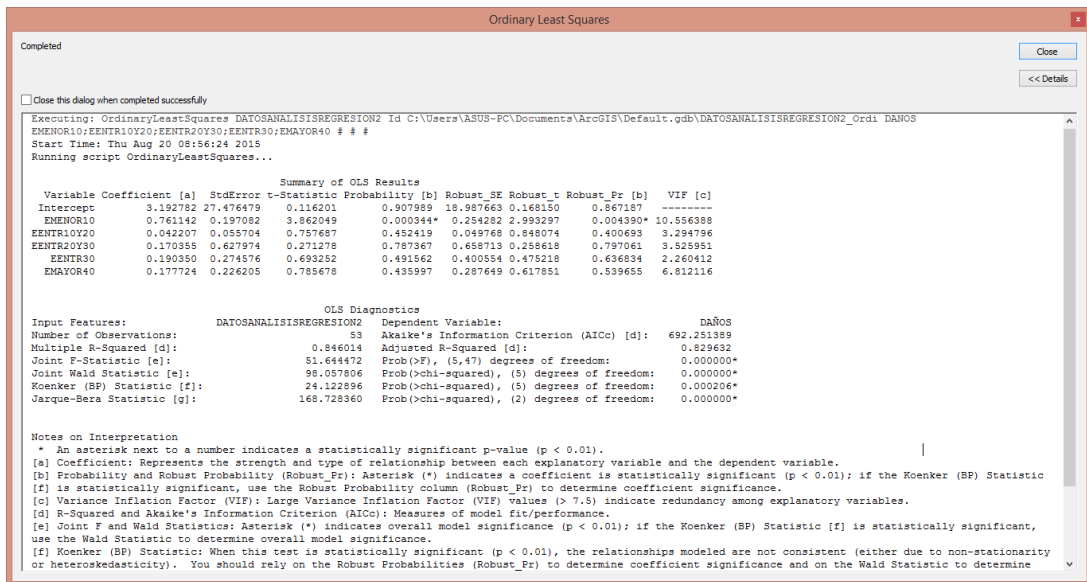
REQUENA LAVIÑA, María. El desarrollo de las marcas gestionadas por la distribución. Análisis de las variables relevantes. Tesis de Grado Doctor. Madrid. Universidad Pontificia Comillas de Madrid. Facultad de Ciencias Económicas y Empresariales (ICADE). 2015.

## ANEXOS

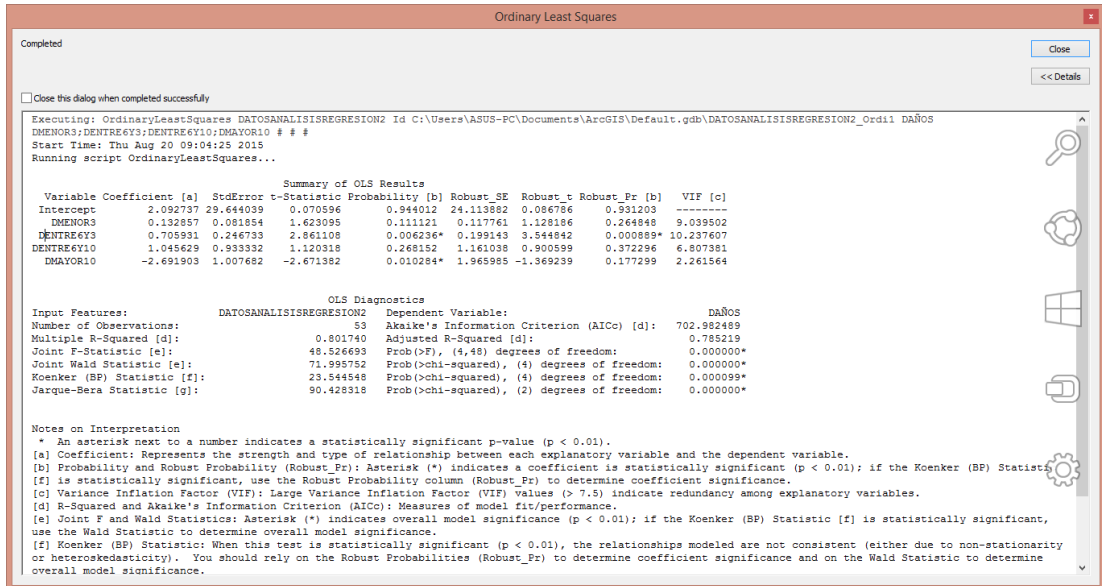
### Anexo A. Resultados numéricos análisis de regresión factor material.



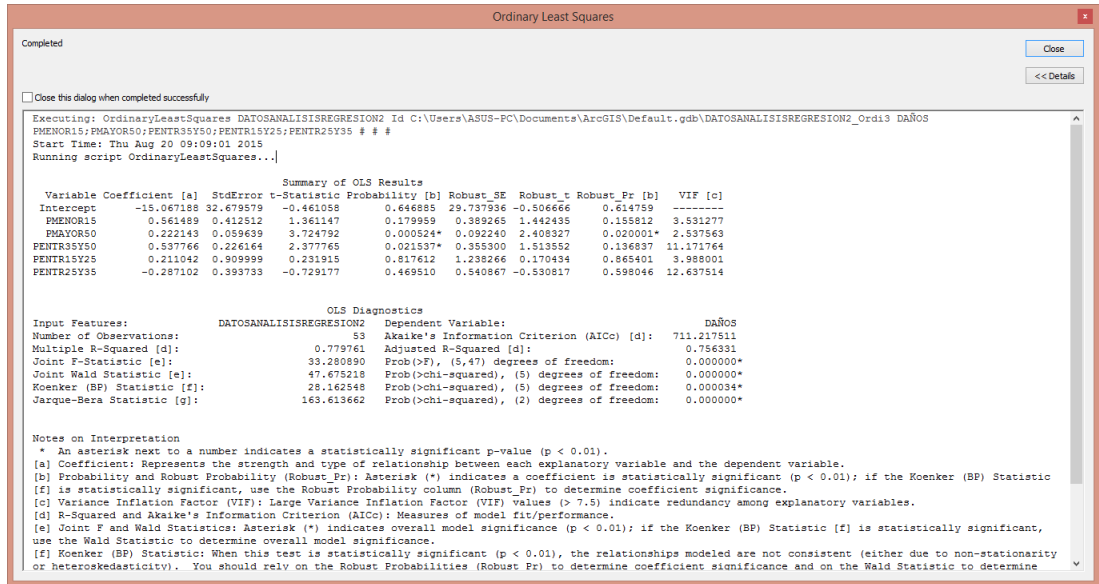
### Anexo B. Resultados numéricos análisis de regresión factor edad.



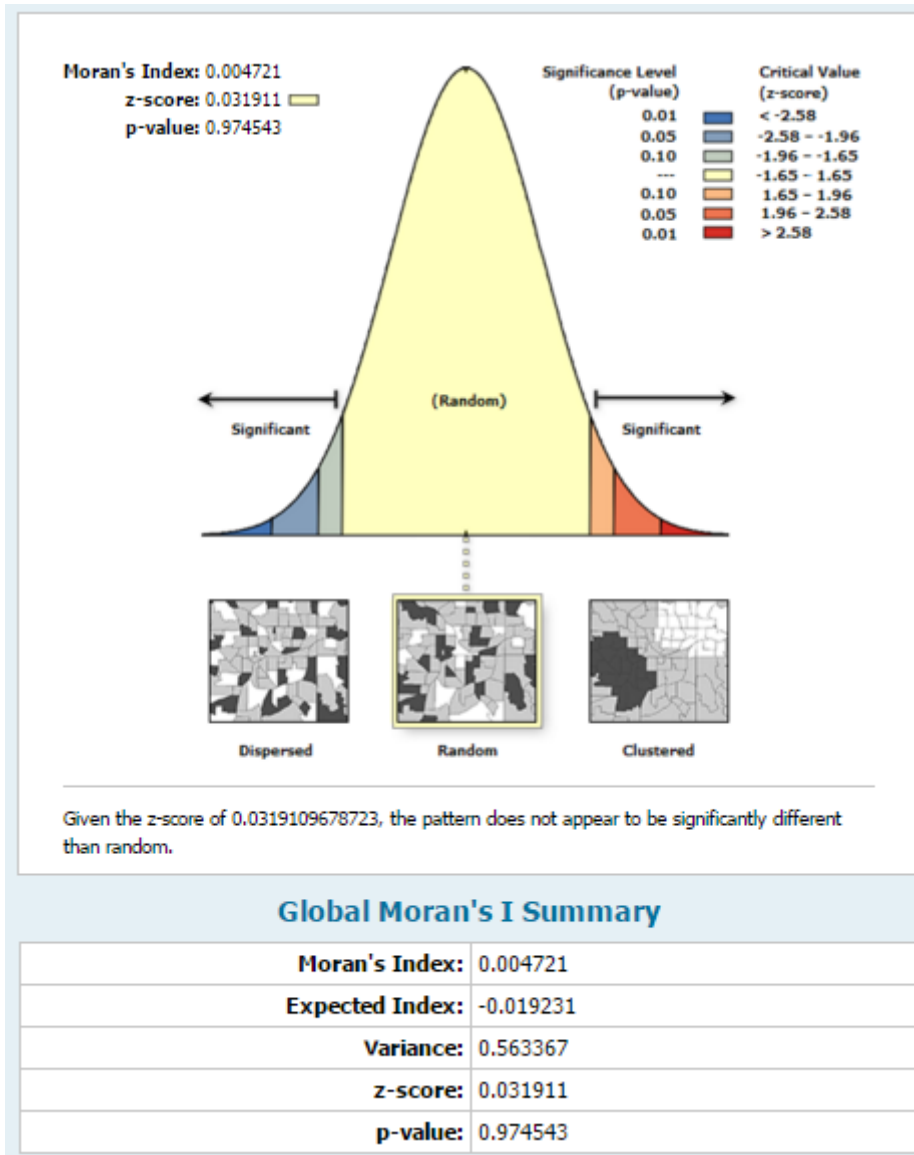
## Anexo C. Resultados numéricos análisis de regresión factor diámetro.



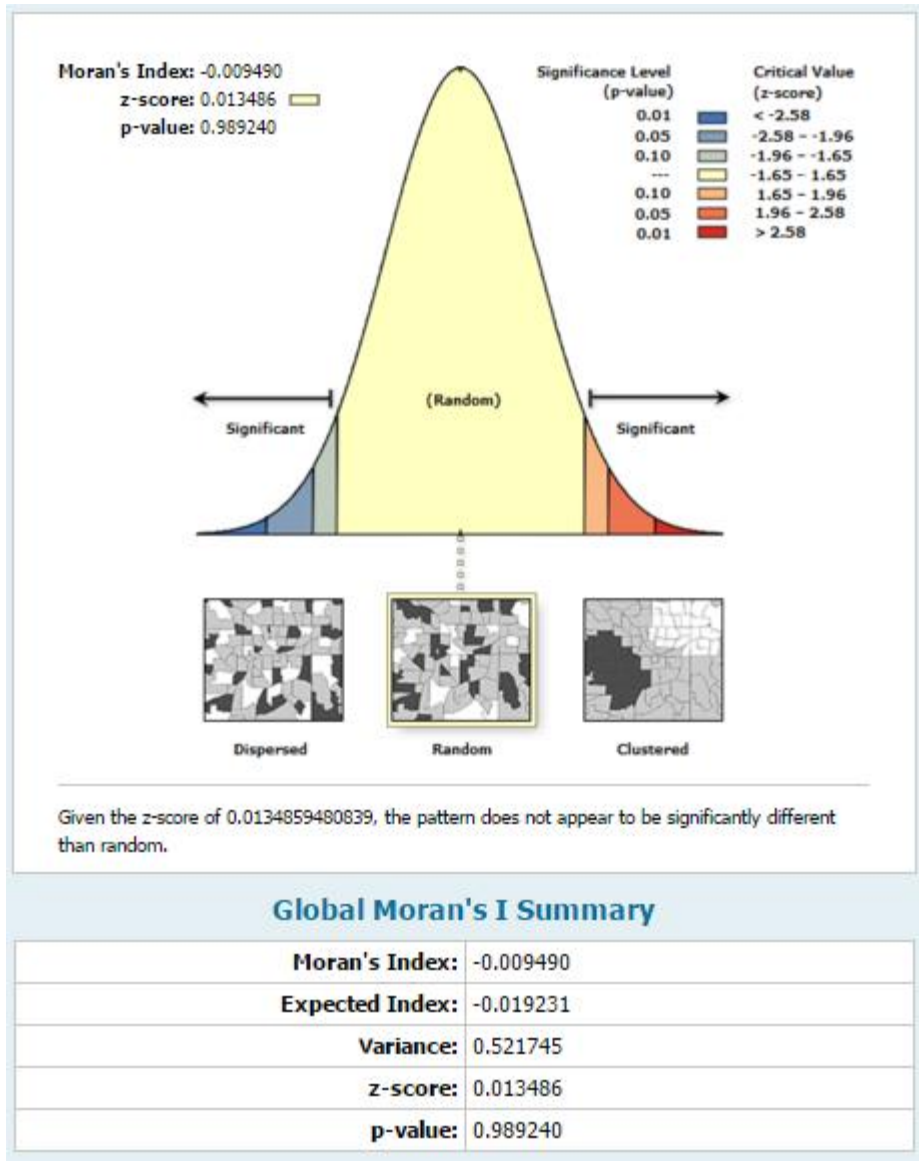
## Anexo D. Resultados numéricos análisis de regresión factor presión.



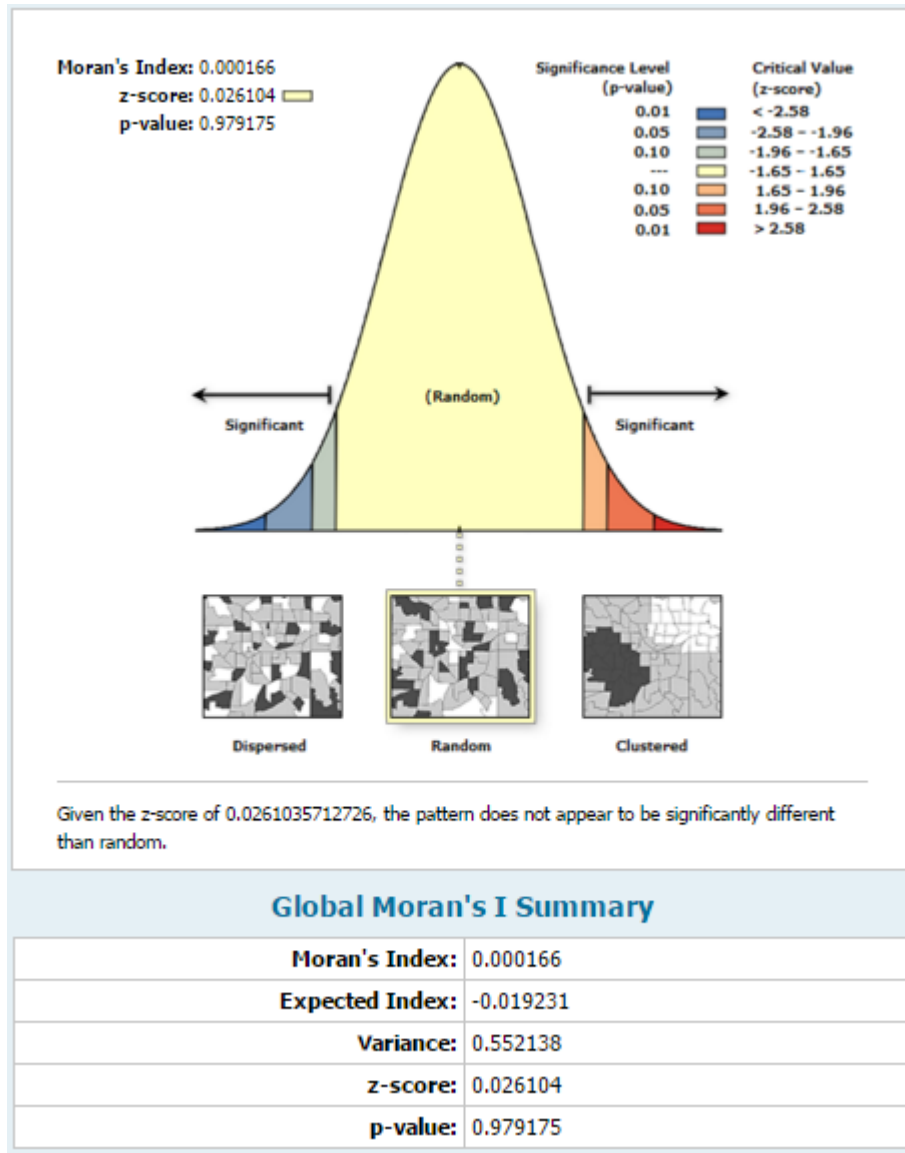
Anexo E. Distribución residuos análisis de regresión factor material.



Anexo F. Distribución residuos análisis de regresión factor edad.



Anexo G. Distribución residuos análisis de regresión factor diámetro.



Anexo H. Distribución residuos análisis de regresión factor presión.

