

***EVALUACIÓN DE DIFERENTES MÉTODOS DE INTERPOLACIÓN PARA LA
VARIABLE PRECIPITACIÓN EN EL DEPARTAMENTO DE CALDAS –
COLOMBIA***

ADRIANA CIFUENTES CARVAJAL



**UNIVERSIDAD DE
MANIZALES**

**UNIVERSIDAD DE MANIZALES
FACULTAD DE CIENCIAS E INGENIERÍA
PROGRAMA ESPECIALIZACIÓN EN SISTEMAS DE INFORMACIÓN
GEOGRÁFICA
MANIZALES
2016**

***INTERPOLACIÓN DE LLUVIA ANUAL PARA EL DEPARTAMENTO DE
CALDAS – COLOMBIA***

ADRIANA CIFUENTES CARVAJAL

Trabajo de Grado presentado para optar
al título de Especialista en Información Geográfica

**UNIVERSIDAD DE MANIZALES
FACULTAD DE CIENCIAS E INGENIERÍA
PROGRAMA ESPECIALIZACIÓN EN SISTEMAS DE INFORMACIÓN
GEOGRÁFICA
MANIZALES
2016**

AGRADECIMIENTOS

La autora expresa sus agradecimientos a:

La Federación Nacional de Cafeteros de Colombia, por la información suministrada sobre la red de estaciones presentes en el departamento.

La Universidad de Manizales por el acompañamiento en el desarrollo del proceso académico.

Andrés J. Peña Q. Msc. en Meteorología y Grand student in PhD of Biological and Agricultural Engineering, por sus aportes y acompañamiento como asesor del trabajo.

Julián Andrés Valencia Arbeláez, Msc. en Medio Ambiente y Desarrollo Sostenible, estudiante de Doctorado en Ciencias Agrarias, por sus aportes y acompañamiento como asesor.

Luis Carlos Correa Ortiz Docente Universidad de Manizales por la asesoría en el desarrollo de este trabajo.

Alberto Boada, Docente del módulo de Estadística Espacial, por la asesoría.

CONTENIDO

1. ÁREA PROBLEMÁTICA	19
2. OBJETIVOS	20
2.1 OBJETIVO GENERAL	20
2.2 OBJETIVOS ESPECÍFICOS	20
3. JUSTIFICACIÓN	21
4. MARCO TEÓRICO	22
5. METODOLOGÍA	27
5.1 TIPO DE TRABAJO	27
5.2 INFORMACIÓN UTILIZADA	27
5.3 PROCEDIMIENTO	28
6. RESULTADOS	33
6.1 DESCRIPCIÓN DE RESULTADOS	33
7. DISCUSIÓN DE RESULTADOS	56
8. CONCLUSIONES	60
9. RECOMENDACIONES	61
BIBLIOGRAFÍA	62
ANEXOS	64
ANEXO A	64
ANEXO B	74
ANEXO C	85
ANEXO D	95
ANEXO E	105
ANEXO F.....	123

LISTA DE GRÁFICOS

	Pág.
Gráfico 1. Estaciones meteorológicas ubicadas en el departamento de Caldas.	31
Gráfico 2. Histograma año 1992.	36
Gráfico 3. Histograma año 1989.	37
Gráfico 4. Gráfica QQPlot, año 1991.	38
Gráfico 5. Análisis de tendencia del año 1992.	39
Gráfico 6. Semivariogramas, fase exploratoria de datos, año 1992.	40
Gráfico 7. Semivariograma fase exploratoria de datos, año 2001.	41
Gráfico 8. Método IDW, año 1995.	42
Gráfico 9. Método IDW, año 1991.	42
Gráfico 10. Global Polynomial Interpolation, año 1991.	43
Gráfico 11. Global Polynomial Interpolation, año 1995.	43
Gráfico 12. Paralelo entre métodos IDW vs. Global Polynomial Interpolation, año 1989.	44
Gráfico 13. Validación Cruzada, método IDW vs. Kriging Universal, año 1989.	45
Gráfico 14. Validación Cruzada, método IDW vs. Kriging Ordinario, año 1989.	45
Gráfico 15. Comparación entre los valores de predicción y los valores observados en las estaciones.	46
Gráfico 16. Superficie de precipitación generada con el método IDW año 1989.	47

Gráfico 17. Superficie de precipitación generada con el método Global Polynomial Interpolation año 1989.	48
Gráfico 18. Superficie de precipitación generada con el método Geoestadístico Kriging Ordinario año 1989.	48
Gráfico 19. Superficie de precipitación generada con el método Geoestadístico Kriging Universal año 1989.	49

LISTA DE TABLAS

	Pág.
Tabla 1. Estaciones meteorológicas pertenecientes a la red climática de la Federación Nacional de Cafeteros de Colombia, ubicadas entre los 1026 y 2088 metros de altitud.	25
Tabla 2. Georreferenciación de las estaciones meteorológicas ubicadas en el departamento de Caldas.	30
Tabla 3. . Análisis de varianza de los valores de precipitación generados en cada uno de los métodos de interpolación.	53

LISTA DE CUADROS

	Pág.
Cuadro 1. Tabla dinámica de los datos de las estaciones de la red meteorológica en Caldas.	33
Cuadro 2. Base de datos anual, escogidos aleatoriamente.	34
Cuadro 3. Valores retirados por estación.	35
Cuadro 4. Valores de precipitación inicial vs. Valores generados con los diferentes métodos de interpolación utilizados.	50
Cuadro 5. Valores de error medio cuadrático generado con los diferentes métodos de interpolación determinísticos y geoestadísticos.	52

LISTA DE ANEXOS

ANEXO A. Análisis Exploratorio de Datos, Histogramas anuales.

ANEXO B. Gráficos anuales, Normal QQ Plot

ANEXO C. Análisis Exploratorio de datos: Trend Analysis

ANEXO D. Análisis Exploratorio de datos: Semivariogramas

ANEXO E. Superficies de Interpolación para datos de Precipitación anual

ANEXO F. Validaciones cruzadas, métodos de Interpolación

GLOSARIO

Agrometeorología: Es la ciencia que estudia las condiciones meteorológicas, climáticas e hidrológicas y su interrelación en los procesos de producción agrícola.

Anisotropía: Un material es anisótropo cuando sus propiedades dependen de la orientación según la cual se hace la medición de ellas.

Asimetría: Deformación horizontal de las distribuciones de frecuencia.

Análisis exploratorio de datos: Proceso para examinar los datos previamente a la aplicación de cualquier técnica estadística. De esta forma se consigue un entendimiento básico de sus datos y de las relaciones existentes entre las variables analizadas.

ArcGIS: Conjunto de productos de software en el campo de los Sistemas de Información Geográfica. Aplicaciones para la captura, edición, análisis, tratamiento, diseño, publicación e impresión de información geográfica

Autocorrelación espacial: Dado un conjunto de entidades y un atributo asociado, evalúa si el patrón expresado está agrupado, disperso o es aleatorio.

Correlación: La correlación indica la fuerza y la dirección de una relación lineal y proporcionalidad entre dos variables estadísticas.

Distribución espacial: Localización, comprensión y explicación de determinados elementos sobre el espacio.

Estadística espacial: Comprende las técnicas formales que estudian las entidades que utilizan sus propiedades topológicas, geométricas o geográficas.

Elevación: Distancia vertical sobre (o por debajo) del geoide o del nivel medio del mar.

Error cuadrático medio: Estimador mide el promedio de los errores al cuadrado, es decir, la diferencia entre el estimador y lo que se estima

Extrapolación: Método científico lógico que consiste en suponer que el curso de los acontecimientos continuará en el futuro, convirtiéndose en las reglas que se utilizarán para llegar a una nueva conclusión.

Geoestadística: Rama de la Geografía matemática que se centra en los conjuntos de datos de la superficie terrestre, conocidos también como datos

espaciales o espacio-temporales. La Geoestadística puede considerarse como una disciplina que se ocupa del análisis estadístico de variables espacialmente distribuidas.

Geografía: Ciencia que estudia los hechos y fenómenos físicos, biológicos y humanos de la superficie terrestre atendiendo las causas que lo originan y a su relación con otros hechos y fenómenos.

Geoestatistical analyst: Extensión que presenta diferentes herramientas para determinar la normalidad de los datos y los diferentes métodos de interpolación.

Geomorfología: Ciencia que tiene como objeto el estudio de las formas de la superficie terrestre enfocado a describir, entender su génesis y su actual comportamiento

GPI: Técnica de interpolación determinística inexacta.

Hidrología: Rama de las ciencias de la Tierra que estudia las propiedades físicas, químicas y mecánicas del agua continental y marítima, su distribución y circulación en la superficie de la Tierra, en la corteza terrestre y en la atmósfera

IDW: Predicción que asume implícitamente la continuidad de la variable, pues se basa en un conocido principio de las ciencias de la tierra que postula que “los valores cercanos (valores obtenidos en nuestras cercanas) son probablemente más similares que los que se encuentran más separados”.

Interpolación: La interpolación se puede definir como un procedimiento que permite calcular el valor de una variable en una posición del espacio (punto no muestral, con un valor estimado) conociendo los valores de esa variable en otras posiciones del espacio (puntos muestrales con valores reales).

Isotrópico: Es decir, cuando una propiedad tiene el mismo valor independiente de la dirección según la cual se hace la medida.

Kriging: Conjunto de métodos que generan superficies que incorporan las propiedades estadísticas de los datos muestrales y que proporcionan una medida de error de las mismas.

Latitud: Es la distancia angular entre la línea ecuatorial (el ecuador), y un punto determinado de la Tierra, medida a lo largo del meridiano en el que se encuentra dicho punto.

Longitud: Expresa la distancia angular entre un punto dado de la superficie terrestre y el meridiano que se tome como 0° (es decir el meridiano de base) medida a lo largo del paralelo en el que se encuentra dicho punto, una

circunferencia cuyo centro es la intersección del eje de la Tierra con el plano del citado paralelo.

Medición continua: Es aquella que puede tomar un valor fijo dentro de un intervalo determinado.

Medición discreta: Sólo puede tomar algunos valores dentro de un mínimo conjunto numerable.

Meteorología: Es la ciencia interdisciplinaria, de la física de la atmósfera, que estudia el estado del tiempo, el medio atmosférico, los fenómenos producidos y las leyes que lo rigen.

Métodos determinísticos: Formulan unas suposiciones generales y habitualmente no contrastadas, sobre el carácter general de la superficie a interpolar, y en función de ellas establecen la función matemática de interpolación (métodos directos).

Normalidad: En geoestadística son aquellas distribuciones comunes que presentan histogramas en forma de campana y más o menos simétricos con respecto a la media.

Precipitación: Se conoce como precipitación a la cantidad de agua que cae a la superficie terrestre y proviene de la humedad atmosférica, ya sea en estado líquido (llovizna y lluvia) o en estado sólido (escarcha, nieve, granizo).

Probabilidad: Método por el cual se obtiene la frecuencia de un acontecimiento determinado mediante la realización de un experimento aleatorio, del que se conocen todos los resultados posibles, bajo condiciones suficientemente estables.

SIG: Sistemas de Información Geográfica, conjunto de programas informáticos que captan, almacenan, recuperan, transforman, muestran y analizan diversos tipos de datos geográficamente referenciados.

Tendencia: Es un patrón de comportamiento de los elementos de un entorno particular durante un período.

Topografía: Ciencia que estudia el conjunto de principios y procedimientos que tienen por objeto la representación gráfica de la superficie terrestre, con sus formas y detalles; tanto naturales como artificiales.

Validación cruzada: Técnica utilizada para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y prueba.

Variable: Una variable es una propiedad que puede fluctuar y cuya variación es susceptible de adoptar diferentes valores, los cuales pueden medirse u observarse.

Variable aleatoria: Es una variable estadística cuyos valores se obtienen de mediciones en experimento aleatorio.

Variable regionalizada: Variación espacial de ciertas magnitudes.

Zona andina: Región tiene este nombre porque está formada por las tres cordilleras de los Andes.

RESUMEN

Con el fin de determinar la utilidad de diferentes métodos de interpolación tanto determinísticos como geoestadísticos en eventos climáticos en zonas andinas ecuatoriales, se utilizaron los registros de precipitación diaria provenientes de las estaciones meteorológicas pertenecientes a la Red Climática de la Federación Nacional de Cafeteros de Colombia, ubicados en el departamento de Caldas. Diferentes herramientas geoestadísticas fueron empleadas para el análisis exploratorio y estructural de datos hasta generar superficies de datos de lluvia a escala de tiempo anual. La validación cruzada permitió comprobar la validez del modelo e indicar cuál era el mejor, al utilizar todos los datos para estimar el modelo de autocorrelación. Al quitar 2 de las 21 estaciones predijo con el resto de la muestra los valores ausentes, para después compararlos con los valores observados. La red de 21 estaciones influyó directamente en la estimación de los valores, siendo un parámetro determinante en la precisión de la predicción. Es así como el resultado óptimo está basado en estaciones pluviométricas distribuidas dentro de un radio de influencia. La aplicación del método determinístico IDW, mostró como resultado mayor precisión en la predicción, con valores menores de error cuadrático. Es necesario evaluar si en aquellos lugares en los que la red climática es limitada, los estudios exigen un aumento de las estaciones o una redistribución, y si las herramientas de análisis geoestadístico son las indicadas.

PALABRAS CLAVES: Precipitación, interpolación, determinístico, geoestadístico, IDW.

ABSTRACT

With the purpose of determining the utility of different interpolation methods, deterministic, as well as, geostatistic, on climatic events in equatorial Andean zones, daily rainfall records were used, which were taken from meteorological stations, each one belonging to Red Climatic of Federación Nacional de Cafeteros de Colombia, located the department of Caldas. Different geostatistic tools were employed for the analysis of exploratory and structural data, until generating rainfall data surfaces, over a year time scale. Crossed validation permitted to prove the validity of the model right, and to indicate which was the best one, by utilizing all the data to estimate autocorrelation model. By removing 2 of the 21 stations, it all was predicted based on the rest of the simple, from the absent values, to be compared afterwards with the observed values. The 21 station networks influenced directly on the value estimation, being a determining parameter over the precision of the prediction. That's how the optimum result is based on the rainfall stations spread out in the influence radius. The application of the IDW, deterministic method, showed as a result, a higher predicting precision with values under the quadratic error. It's necessary to evaluate if those places where the climatic network is limited, the studies demand in increase in the number of stations or a new distribution, and if tools for geostatistic analysis are the suitable ones.

KEY WORDS: Rainfall, Interpolation, geostatistic, deterministic, IDW.

INTRODUCCIÓN

De acuerdo con Eslava (1993), la topografía es un factor determinante en el clima de un lugar, básicamente por la combinación de la latitud, pendiente y dirección de la vertiente, así como por la orientación de las cordilleras cercanas. Los estudios realizados en las regiones montañosas demuestran que la intensidad y la duración de la precipitación cambian con la altitud y con las variaciones de la fisiografía. Aspectos como el crecimiento y desarrollo de cultivos y plantas en general dependen en gran medida del exceso o la deficiencia hídrica, resaltando la importancia de analizar esta variable.

En este trabajo se pretenden utilizar las ventajas de los Sistemas de Información Geográficos (SIG) como herramienta de análisis para la optimización y la planificación del entorno. Específicamente, se fusionan los conocimientos agro-meteorológicos con los SIG para proveer a los tomadores de decisiones de un análisis detallado de las características espaciales particulares, o distribución espacial de la lluvia anual (Ranses et al., 2012). La precipitación es un elemento del clima, cuya distribución espacio-temporal es extremadamente variable. Sin embargo, debido a la escasez de instrumentos de medida que permitan tener una certidumbre con respecto a las zonas más o menos lluviosas, como ocurre en el departamento de Caldas, la mejor opción es interpolar los valores de lluvia a una escala dada, en este caso a escala anual.

Esencialmente, la interpolación es una técnica estadística (matemática) que se usa para obtener información en puntos no muestreados. La ventaja de utilizar las herramientas de interpolación dentro de los SIG radica en que bajo estas Plataformas, los factores geográficos y topográficos pueden ser incluidos como variables secundarias en procedimientos de interpolación. Sin embargo no ha sido determinado un método general, debido a factores como el tamaño de la zona de estudio, variables topográficas y escala temporal considerada (diaria, mensual y anual) (Portalés et al., 2008). Eventos de tipo espacial que presentan correlación, permiten predecir valores de variables en sitios donde no han sido muestreadas (Giraldo, 2000).

La literatura indica que el kriging es uno de los métodos utilizados para interpolar datos de precipitación pero en la práctica este presenta ciertos inconvenientes en su utilización operativa ya que resulta muy laborioso y requiere de encontrar un variograma óptimo para cada conjunto de datos diferentes, además de generar extrapolaciones fuera de rango cuando aparecen en el espacio lagunas de información (Ranses et al, 2012). Situaciones como la existencia de valores atípicos dentro del conjunto de datos a analizar, pueden presentarse, sujetos a errores humanos o por causa de falla de los instrumentos de medición. Es posible eliminarlos, aunque no se pueda generalizar esta solución a todos los casos, ya que puede repercutir en el potencial con el que cuenta un área determinada. La

importancia de identificar y evaluar la distribución, la ubicación geográfica, los valores extremos, las medidas de localización, la variabilidad y correlación facilitan establecer si la aplicación de la geoestadística es válida o si se hace necesario implementar otros procedimientos (Giraldo, 2000).

1. ÁREA PROBLEMÁTICA

La alta variabilidad espacio-temporal de la precipitación constituye una fuente importante de incertidumbre en la modelación hidrológica y del comportamiento de la precipitación sobre diferentes sectores de la geografía colombiana, por lo que se hace indispensable la fusión de los conocimientos agrometeorológicos y la utilización de los sistemas de información geográfica para proveer a los agricultores análisis detallados de las características particulares y la distribución espacial de los elementos o índices que se necesiten (Ranses et al, 2012). En la actualidad, es preciso definir el método de interpolación que más se ajusta a variables climáticas como la precipitación, en zonas donde la disponibilidad de estaciones es limitada.

2. OBJETIVOS

2.1 OBJETIVO GENERAL

Generar superficies de datos de lluvia a escala de tiempo anual para determinar la idoneidad de cuatro métodos de interpolación.

2.2 OBJETIVOS ESPECÍFICOS

- Establecer el método de interpolación determinístico o geoestadístico idóneo para datos de precipitación en el departamento de Caldas.
- Desarrollar un mapa, basado en la combinación de datos pluviométricos de Caldas, a escala anual, y un modelo de interpolación.
- Utilizar los resultados para determinar si los análisis geoestadísticos son adecuados para interpolar datos en regiones andinas ecuatoriales.

3. JUSTIFICACIÓN

Diversos estudios, en diferentes regiones del mundo han propuesto múltiples metodologías de interpolación para identificar aquella que se ajuste mejor a las condiciones climáticas de una región, con condiciones específicas tanto topográficas como geográficas. De manera tradicional se han empleado métodos simples, como los lineales o cuadráticos, en los cuales los valores interpolados se derivan exclusivamente a partir de los valores de precipitación dados en los observatorios (Portalés et al, 2008). Actualmente el uso de herramientas más avanzadas de SIG y geoestadística incorporan metodologías que explican estos eventos, al tomar en cuenta variables secundarias. Sin embargo, no se ha estandarizado un método que explique el objeto de estudio, ya que depende de la escala temporal considerada (diaria, mensual, anual etc.) y de variables adicionales propias de cada zona.

El conocimiento de la distribución espacial y temporal de la precipitación a escala regional está encaminada a la aplicación científica, económica, social, ambiental, diagnósticos climáticos, detección temprana de amenazas, administración del recurso hídrico, mapeo de zonas con susceptibilidad a inundaciones, afectación de cultivos entre otros. La ausencia de estaciones meteorológicas en algunas zonas del departamento de Caldas, dificultan el análisis de la información, haciendo necesario integrar todos los recursos disponibles para aplicar el modelo de interpolación que mejor se ajuste a los datos suministrados por una red climática, en la zona andina.

Este trabajo está orientado a generar una herramienta que facilite la toma de decisiones en procesos investigativos, en centros como Cenicafé, entes gubernamentales como el IDEAM, y en el desarrollo de los planes de ordenamiento territorial POT, entre otras aplicaciones.

4. MARCO TEÓRICO

Viera en el 2002, relaciona a Matheron, quien en los años 60, acuñó el término de Geoestadística. Reconocido como el padre de esta disciplina, Matheron formalizó y generalizó matemáticamente un conjunto de técnicas desarrolladas por D. G. Krige (1941) que explotaban la correlación espacial para hacer predicciones en la evaluación de reservas de las minas de oro en Sudáfrica. Él definió a la Geoestadística como "la aplicación del formalismo de las funciones aleatorias al reconocimiento y estimación de fenómenos naturales" (Matheron, 1962 citado por Viera 2002). Si bien más adelante nos detendremos en el concepto de una función aleatoria, baste decir de momento que puede visualizarse como la asociación de una variable aleatoria a cada punto del espacio. El estudio de fenómenos con correlación espacial, por medio de métodos geoestadísticos, surgió a partir de los años sesenta, especialmente con el propósito de predecir valores de las variables en sitios no muestreados.

Giraldo, 2000, cita los trabajos de Sichel (1947; 1949) y Krige (1951) como antecedentes del estudio de fenómenos con correlación espacial, por medio de métodos geoestadísticos. El primero observó la naturaleza asimétrica de la distribución del contenido de oro en las minas surafricanas, la equiparó a una distribución de probabilidad log normal y desarrolló las fórmulas básicas para esta distribución. Ello permitió una primera estimación de las reservas, pero bajo el supuesto de que las mediciones eran independientes, en clara contradicción con la experiencia de que existen "zonas" más ricas que otras. Una primera aproximación a la solución de este problema fue dada por geólogo G. Krige que propuso una variante del método de medias móviles, el cual puede considerarse como el equivalente al krigeado simple que, como se verá más adelante, es uno de los métodos de estimación lineal en el espacio con mayores cualidades teóricas. La formulación rigurosa y la solución al problema de predicción (estimación en muchos textos geoestadísticos) vino de la mano de Matheron (1962) en la escuela de minas de París. En los años sucesivos la teoría se fue depurando, ampliando su campo de validez y reduciendo las hipótesis necesarias (Samper y Carrera, 1990). De la minería las técnicas geoestadísticas, se han "exportado" a muchos otros campos como hidrología, física del suelo, ciencias de la tierra y más recientemente al monitoreo ambiental y al procesamiento de imágenes de satélite.

Con el avance de la tecnología en sistemas computarizados, es posible conocer aplicaciones en áreas como la cartografía, meteorología, procesos investigativos en recursos naturales, desarrollo de planes de ordenamiento territorial, agricultura, minería, y múltiples aplicaciones en detección de zonas vulnerables a eventos naturales o antrópicos. Grandes cantidades de datos que anteriormente no podían

ser procesados de forma efectiva, ahora son analizados en poco tiempo; la automatización de diversas tareas ha mejorado la calidad de los procesos.

Los estudios ambientales que han utilizado técnicas geoestadísticas para predecir diferentes situaciones son innumerables, a pesar de lo reciente de estas aplicaciones. La estadística espacial reúne las metodologías apropiadas para el análisis de datos que corresponden a la medición de variables aleatorias en diversos sitios (puntos del espacio o agregaciones espaciales) de una región. De forma análoga a otros procesos estadísticos, la primera etapa en el estudio de la información georeferenciada, debe cumplir con el análisis exploratorio de datos (AED), que busca determinar las observaciones extremas, la localización y variabilidad. El eje central del análisis geoestadístico es la estructura de autocorrelación entre los datos y su uso en la predicción a través de las técnicas conocidas como kriging y cokriging (Giraldo, 2000)

Ciudades como Bogotá no han sido ajenas a la utilización de estos métodos. Gracias a la información puntual que proveen instrumentos de medición discreta o continua, se realizó un estudio sobre la distribución e interpolación espacial de las lluvias, resaltando la importancia de construir campos de esta variable utilizando métodos como Kriging, IDW (Inverse Distance Weighted) y el método V4 (Desarrollado en MATLAB® basado en las funciones de Green). La dificultad de estas técnicas se concentra en las medidas previas a considerar, es por ello que estudios como este generan un impacto en aquellos modeladores que buscan descifrar el comportamiento de este evento meteorológico. El análisis estructural es un paso importante antes de abordar los métodos de interpolación, por ello es definido en la literatura como el primer paso dentro del análisis geoestadístico y considera el análisis exploratorio de datos; fundamentado en las técnicas estadísticas convencionales que reconoce la información inexplorada sobre la muestra.

Una de las grandes limitantes que se encuentran en este tipo de investigaciones, se basa en la ausencia de datos o número de estaciones en el área de interés. Este cubrimiento limitado aumenta la incertidumbre de los valores interpolados a partir de registros de superficie. Teniendo en cuenta lo anterior y utilizando la información de tipo geográfica (longitud, latitud y elevación) para correlacionarla con datos climáticos, autores como Chen et al, (2009) elijen métodos como el regression plus kriging, co-kriging, gradient plus inverse distance square (GIDS), y Spline.

Resalta la necesidad de enfocar los estudios climáticos a eventos extremos, siendo los datos diarios los que toman más fuerza, y cuya interpolación se hace más complicada que a escala mensual, por efecto de la presencia de patrones discretos en el tiempo y en espacio. Tanto la frecuencia como la distribución, son factores determinantes y de gran influencia en la calidad de la interpolación de variables como la precipitación, así lo explica este autor en su artículo Spatial

Interpolation of Daily Precipitation in China: 1951- 2005; concluyendo además que los métodos IDW y KO (co-kriging) son los que mejor explican este evento climático a escala diaria, en China. Es indispensable resaltar que aunque se apoyó en la validación cruzada para concretar los mejores métodos, estos resultados no pueden ser generalizados para datos proporcionados en otra investigación, ya se ha dicho que las condiciones de cada lugar evitan que sea uno solo el mejor modelo. Describió que las ventajas del método IDW, radican en que tiene en cuenta la interacción del terreno montañoso y la variable precipitación.

En países como España, en la ciudad de Valencia, se determinó que el uso de la regresión múltiple permitió llegar a los modelos geoestadísticos mediante la técnica del cokrigado (KO) siendo este uno de los más fiables a escala estacional y anual.

Como muchos textos lo indican el análisis geoestadístico se compone de pasos o etapas ante el desarrollo de procesos investigativos en los que se sugiere se anticipe tres pasos: El análisis exploratorio de los datos, el análisis estructura y las predicciones (kriging o simulaciones), pero para proceder es necesario conocer conceptos básicos como el planteado por Viera, 2002, el cual habla de variable regionalizada $z(x)$ como aquella distribuida en el espacio de manera tal que presenta una estructura espacial de correlación. Así mismo propone una definición más rigurosa matemáticamente equivalente consistiría en decir que una variable regionalizada es una variable aleatoria z definida en un punto del espacio \underline{x} . Donde en el caso más general \underline{x} es un punto en el espacio tridimensional, es decir $\underline{x}: (x_1, x_2, x_3)$. La variable aleatoria, fue incluida en este conjunto de conceptos, descrita de la siguiente manera: Si a cada punto \underline{x} que pertenece a un dominio en el espacio le hacemos corresponder una variable aleatoria $Z(\underline{x})$, que en sentido general pueden ser dependientes, entonces el conjunto de variables aleatorias espacialmente distribuidas $\{Z(\underline{x}), \underline{x} \in \Omega\}$ será una función aleatoria $Z(\underline{x})$.

Adicional a estos significados, Giraldo, 2000, en su libro Introducción a la Estadística, Teoría y Aplicación, nombra aspectos como los Lattices (enmallados), expuestos como así: Las ubicaciones s pertenecen a un conjunto D discreto y son seleccionadas por el investigador (D fijo). Estas pueden estar regular o irregularmente espaciadas. Acompaña al lector con algunos ejemplo sobre la tasa de morbilidad de hepatitis en Colombia medida por departamentos, tasa de accidentalidad en sitios de una ciudad, producción de caña de azúcar en el departamento del Valle del Cauca según municipio, colores de los pixeles en interpretación de imágenes de satélite. En los ejemplos anteriores pretendía que se observara que el conjunto de ubicaciones de interés es discreto y que estas corresponden a agregaciones espaciales más que a un conjunto de puntos del espacio. Así definía la carencia de sentido de la interpolación espacial con este tipo de datos. Los patrones espaciales los nombra así: Las ubicaciones pertenecen a un conjunto D que puede ser discreto o continuo y su selección no depende del investigador (D aleatorio), los ejemplos de datos dentro de esta área

son: Localización de nidos de pájaros en una región dada, puntos de imperfectos dentro de una placa metálica, ubicación de los sitios de terremoto en Colombia o cuadrantes de una región con presencia de una especie particular. Debe notarse que en los ejemplos anteriores hay aleatoriedad en la selección de los sitios, puesto que la ubicación de los nidos de los pájaros, de los imperfectos dentro de la placa metálica, de los sitios de terremoto o de los cuadrantes con presencia de la especie, no depende del criterio del investigador. Una vez se ha hecho la selección de sitios es posible hacer medidas de variables aleatorias en cada uno de ellos. Por ejemplo si en primera instancia se establece la ubicación de árboles de pino dentro de un bosque, es posible que sea de interés medir en cada uno de los árboles el diámetro o la altura. En general el propósito de análisis en estos casos es el de determinar si la distribución de los individuos dentro de la región es aleatoria, agregada o uniforme. Para cualquier tipo de estudio que involucre datos espaciales es vital incluir la georreferenciación de los datos; las mediciones de las características de interés en un estudio regionalizado tienen implícitamente asociadas las coordenadas de los sitios en donde estas fueron tomadas. Cuando el área de estudio es considerablemente grande se usa un geoposicionador para establecer dichas coordenadas (Giraldo, 2000).

Viera, 2002, conceptualiza el semivariograma, conocido también como variograma, como la herramienta central de la geoestadística. El semivariograma es una función que relaciona la semivarianza con el vector h conocido como "lag", el cual denota la separación en distancia y dirección de cualquier par de valores $Z(\underline{x})$ y $Z(\underline{x} + \underline{h})$. Este autor determina, las consideraciones para que el cómputo del semivariograma muestral eleve la eficiencia y la calidad de la estimación, independientemente del tipo de estimador que se utilice. Esas reglas son:

- En la estimación del semivariograma los pares de las observaciones se agrupan según la distancia dentro de un intervalo $h = |\underline{h}|$ con una tolerancia $\pm \Delta h/2$ y dentro de una dirección θ con una tolerancia $\pm \Delta\theta/2$. El semivariograma así estimado es considerado suavizado o regularizado.
- El semivariograma muestral debe ser considerado solamente para pequeñas distancias por lo que generalmente, se estima para valores de $|\underline{h}|$ menores que la mitad de la distancia máxima ($|\underline{h}| < d_{max}/2$).
- La elección del número de intervalos es arbitraria. No obstante se considera que un número máximo de 25 intervalos es suficiente para cualquier propósito, y un mínimo de 10 debe ser usado para determinar con precisión el rango y la meseta del semivariograma.
- El largo de los intervalos debe ser elegido de forma tal que el número de pares en cada intervalo sea lo suficientemente grande para que el estimado

del semivariograma sea relativamente estable. Se considera que entre 30 y 50 pares satisfacen este requerimiento.

- Los valores estimados para cada intervalo se deben graficar contra la distancia promedio de todos los pares que se encuentran dentro de dicho intervalo.

Considera por su forma, dos tipos principales de semivariogramas. En el primer tipo, la semivarianza se incrementa con el incremento del valor absoluto del intervalo $|h|$ hasta alcanzar un valor máximo a partir del cual se mantiene relativamente constante y oscila alrededor del mismo. Estos semivariogramas son conocidos como de tipo transitivo. El valor del intervalo a partir del cual el semivariograma no se incrementa es conocido como alcance o rango (radio de correlación) y marca el límite de la dependencia espacial de la propiedad. La varianza máxima es conocida como "sill" o meseta del semivariograma y teóricamente debe coincidir con la varianza a priori σ^2 de la muestra de la función aleatoria $Z(\underline{x})$. Una variable con este tipo de semivariograma no sólo cumple la hipótesis intrínseca, sino también es estacionaria de segundo orden. Esto significa que su valor esperado es constante

Un aspecto del semivariograma, que menciona Viera, 2002, que es importante señalar, es que por definición $\gamma(0) = 0$ pero en la práctica el semivariograma muestral $\gamma^*(h)$ cuando h tiende a cero no necesariamente se anula. Esto es conocido como efecto "nugget" o pepita, y el valor del semivariograma en cero $\gamma(0)$ es conocido como la varianza "nugget" o microvarianza. En principio esto puede ocurrir solamente si existen discontinuidades en la función aleatoria. En la práctica su existencia se debe a la variación espacial que no puede explicar el variograma debido a la escala del muestreo. Para un material continuamente variable el efecto nugget se produce a partir de la contribución de los errores de medición y la variación a distancias mucho menores que el intervalo de muestreo más pequeño.

Ahora bien, es importante resaltar que en geoestadística el propósito esencial es la interpolación y si no hay continuidad espacial pueden hacerse predicciones carentes de sentido.

5. METODOLOGÍA

5.1 TIPO DE TRABAJO

Este trabajo se encuentra enfocado a la investigación aplicada, de carácter exploratorio, con el fin determinar la utilidad de los métodos de interpolación tanto determinísticos como geoestadísticos en eventos climáticos en zonas andinas ecuatoriales.

5.2 INFORMACIÓN UTILIZADA

Se utilizaron los registros de precipitación diaria provenientes de 21 estaciones meteorológicas pertenecientes a la Red Climática de la Federación Nacional de Cafeteros de Colombia, ubicados en el departamento de Caldas (Tabla 1). Los datos corresponden a los últimos 25 años de observación (1988-2013).

Tabla 1. Estaciones meteorológicas pertenecientes a la red climática de la Federación Nacional de Cafeteros de Colombia, ubicadas entre los 1026 y 2088 metros de altitud.

Estación	Dpto	Mpio	Altitud	Latitud (N)		Longitud (W)	
			m	Grados	Min	Grados	Min
CUATRO ESQUINAS	CALDAS	AGUADAS	1900	5	40	75	25
GUAYMARAL	CALDAS	AGUADAS	1600	5	39	75	27
CENICAFE	CALDAS	CHINCHINA	1310	5	0	75	36
LA SIERRA	CALDAS	CHINCHINA	1440	4	59	75	38
NARANJAL	CALDAS	CHINCHINA	1381	4	58	75	39
LA JULIA	CALDAS	FILADELFIA	1650	5	18	75	34
AGRONOMIA	CALDAS	MANIZALES	2088	5	3	75	30
JAVA	CALDAS	MANIZALES	1778	5	1	75	32
LA SELVA	CALDAS	MANIZALES	1312	5	5	75	36
EL DESCANSO	CALDAS	MARMATO	1803	5	30	75	37
SANTA HELENA	CALDAS	MARQUETALIA	1395	5	19	75	0
LA LINDA	CALDAS	PACORA	1750	5	33	75	32
EL RECREO	CALDAS	PALESTINA	1430	5	2	75	39
GRANJA LUKER	CALDAS	PALESTINA	1031	5	4	75	41
LA ARGENTINA	CALDAS	PALESTINA	1354	5	2	75	41
LA PALMA	CALDAS	PALESTINA	1165	5	1	75	41
SANTAGUEDA	CALDAS	PALESTINA	1026	5	4	75	40
LA ARGENTINA	CALDAS	RIOSUCIO	1420	5	28	75	42
LA MANUELITA	CALDAS	RIOSUCIO	1460	5	22	75	41
RAFAEL ESCOBAR	CALDAS	SUPIA	1307	5	27	75	38
LA PASTORITA	CALDAS	VICTORIA	1122	5	19	74	58

5.3 PROCEDIMIENTO

5.3.1 Fase 1. Calidad de la información

La elección de las estaciones se hizo teniendo en cuenta aquellas que presenten mayor número de años con registros confiables o mensualmente más completos, dado que deberían ser más exactos. Un requisito importante en una preselección de estaciones es que éstas tengan un registro de datos de a lo menos 10 años (Pizarro, 2003). Durante la primera etapa, se aplicó el modelo adaptativo al dato, el cual comprende las siguientes actividades:

- **Actividad 1.**

Análisis de datos: Se efectuó el examen preliminar de datos. Durante este procedimiento se analizó la base datos que se tenía en crudo, determinando la calidad de la información. Los datos faltantes fueron obtenidos con la herramienta SueMulador, software que genera datos diarios de precipitación, brillo solar, humedad relativa y temperaturas (mínima y máxima) a través de un proceso estocástico markoviano, para suplir los datos faltantes de las series climáticas de las estaciones meteorológicas ubicadas en zonas ecuatoriales, (Chica et al, 2014).

- **Actividad 2.**

Análisis de datos confirmatorios: Una vez se determinó la calidad de la información, se seleccionaron aleatoriamente, 20 años de la base de datos anual, usando la función Random de Microsoft Excel ® 2010 siguiendo la metodología utilizada por Valencia, et al, 2015.

- **Actividad 3.**

Se retiró aleatoriamente 2 estaciones de las 21 nombradas anteriormente, con el fin de validar la efectividad de los 4 métodos de interpolación. El número de estaciones se determinó de acuerdo a la totalidad de las mismas.

5.2.2 Fase 2. Análisis Exploratorio y Estructural de datos espaciales

Se realizó con el fin de determinar las propiedades estadísticas espaciales de los datos. Se examinó la distribución de los datos, se buscó las tendencias globales, se examinó la autocorrelación espacial y la variación direccional de los datos. Se creó un shape de puntos a partir de datos de coordenadas geográficas, validando la ubicación de las estaciones en la zona de estudio dentro del sistema de coordenadas geográfica GCS_Magna. Esta ubicación describe el comportamiento de la variable en el mismo sentido a la geomorfología de la zona evitando concluir que las estaciones se encuentran mal ubicadas.

- **Actividad 1.**

Software utilizado: Fue empleada la herramienta Geostatistical Analyst dentro del programa ArcGis versión 10.2.2. El Análisis Geoestadístico es un software que usa un conjunto de puntos de muestreo para crear una superficie continua mediante un proceso de interpolación (Superficie Interpolada). Su base es la autocorrelación espacial y hace parte del Análisis Espacial en general (Boada, 2015).

- **Actividad 2.**

Análisis del Histograma: Se examinó la distribución de los datos con medidas de tendencia central, dispersión y de simetría (centralización, Variabilidad y forma). Se determinó el cumplimiento de diferentes condiciones: Que la curva normal

fuese asintótica al eje de abscisas, el coeficiente de sesgo sea igual a cero (0), y la Curtosis igual a tres (3).

Si el sesgo es igual a cero, las distribuciones se presentarían simétricas; si el sesgo es positivo la distribución tendrá una cola asimétrica extendida hacia los valores positivos. Un sesgo negativo indica una distribución con una cola asimétrica extendida hacia los valores negativos. Ante estas características es necesario implementar las transformaciones de los datos.

QQ Plot: Esta herramienta fue utilizada para determinar la normalidad de los datos. Permite relacionar la distribución acumulada obtenida a partir de los datos ordenados vs. la curva de distribución Normal. Si las distribuciones son idénticas, QQPlot general será una línea recta. Si los datos son asimétricos (lejos de lo normal), los puntos se desviarán de la línea recta.

Mapa de Voronoi: Herramienta que permitió construir polígonos alrededor de la localización de los puntos de muestreo, en este caso de las estaciones meteorológicas.

Trend Analysis: Esta ventana enseña la tendencia que siguen los datos. Se toma en la abscisa una coordenada geográfica y en la ordenada el atributo, para determinar si existe tendencia en cada una de las direcciones. El análisis de tendencia proporciona un gráfico en tres dimensiones permitiendo identificar tendencias globales en los datos, la localización de los puntos muestrales dibujados sobre un plano x, y. Se utilizó la barra de desplazamiento para las líneas de tendencias (verde y azul de la misma figura), se observó si estas siguen una línea recta, (caso de tendencia lineal); una curva (la tendencia es cuadrática) o si es una línea con más de una concavidad, tendencia de orden 3. Se buscó la presencia de anisotropía, isotropía o tendencia, que se asocia a la variabilidad y comportamiento de los datos en el espacio.

Para mirar la proyección de los datos, tipo de tendencia de los puntos proyectados en el plano, fue necesario analizar el plano X utilizando en la horizontal 90° y en la vertical un valor de -9, para generar el gráfico en sentido este- oeste. Para la Tendencia en Y, se modificó 180 en horizontal, -9 en vertical.

Modelamiento del semivariograma: Utilizado para la evaluación de la autocorrelación, el cual permite explorar, identificar y cuantificar la dependencia espacial de una variable regionalizada. A partir de las características del semivariograma se dedujo la autocorrelación espacial de la variable y se detectan los valores extremos (Boada, 2015). Se tuvo en cuenta en las gráficas, que los puntos que se alejen hacia la derecha estarán separados por mayores distancias y que a medida que estén situados más altos sobre el eje Y presentaran mayores diferencias entre ellos, presentando así dependencia espacial. Se empleó la opción attribute, herramienta con la que analizó intuitivamente la presencia de

anisotropía a 45 grados, de ángulo de tolerancia, con máximo de distancia de 0.07333 metros.

5.2.3 Fase 3. Cálculo y modelado de superficies

- **Actividad 1. Predicción de superficies**

Utilizando la herramienta Geostatistical Analyst, que incluye diferentes métodos determinísticos y geoestadísticos, fue posible generar las superficies de interpolación. Fue utilizada la precipitación como variable de predicción. No fue empleado el campo de ponderación en ninguno de los métodos (variables secundarias como la altitud). Para establecer el método de interpolación a utilizar, se confrontaron 4 procedimientos: Dos métodos determinísticos, IDW y Global Polynomial Interpolation (GPI); y dos métodos geoestadísticos Kriging Ordinario y Kriging Universal. De esta manera las técnicas de interpolación estimaron el valor de la variable en los puntos no muestrales.

Determinísticos

- IDW (Método exacto)
- Global Polynomial Interpolation (teniendo en cuenta todos los datos, incluyendo aquellos extremos)

Geoestadísticos

- Kriging Universal (cuando los datos presentan tendencia)
- Kriging Ordinario (cuando los datos no presentan tendencia)

Se usó el parámetro de distancia (P), optimizado (ArcGis utiliza por defecto el P: 2), teniendo en cuenta que los valores menores determinan si las observaciones se encuentran cerca. El método determinístico IDW, incluye un número máximo de 15 vecinos y un mínimo de 10, así mismo métodos geoestadísticos como el Kriging, utilizan un mínimo de 2 vecinos y un máximo de 5 vecinos. Estos modelos generan la predicción en el eje X y en Y, luego permiten la construcción del modelo matemático y la regresión, en este caso simple. El uso del error cuadrático medio, facilita la comparación de los valores de precipitación con respecto al promedio de todas las precipitaciones, alejadas en dicho valor.

Los métodos determinísticos formulan unas suposiciones generales y habitualmente no contrastadas, sobre el carácter general de la superficie a interpolar, y en función de ellas establecen la función matemática de interpolación

(métodos directos). Mientras que los métodos geoestadísticos estudian el carácter de la autocorrelación espacial de la variable a interpolar, usando para ello los puntos muestrales.

- **Actividad 2: Bondad del modelo**

Para valorar la bondad de los modelos se generó la validación cruzada, herramienta que permite comprobar la eficacia del modelo e indicar cuál es el mejor. La validación cruzada utiliza todos los datos muestrales para estimar el modelo de autocorrelación. Al quitar un dato muestral, lo predice con el resto de la muestra, después compara el valor predicho de ese punto con el valor observado. Este procedimiento se repite con cada uno de los puntos muestrales y al final de la validación cruzada compara los valores observados con los calculados (Boada, 2015). El estadístico determinado fue el error medio cuadrático (RMSE) (Vicente-Serrano et al. 2003), metodología citada por Portalés et al, 2008.

Se determinó el modelo que más mejor se acomoda, para la generación de las superficies de interés o mapas de precipitación del departamento de Caldas. Una vez comparados los resultados estadísticos se determinó si es necesario utilizar los métodos geoestadísticos para generar los mapas de lluvia para el departamento de Caldas.

- **Actividad 3. Análisis de varianza (Anova)**

Fue empleada la herramienta estadística Anova para determinar si, diferentes métodos utilizados para interpolar, muestran diferencias significativas o por el contrario puede suponerse que sus medias poblacionales no difieren.

6. RESULTADOS

6.1 DESCRIPCIÓN DE RESULTADOS

6.1.1 Ubicación de las estaciones

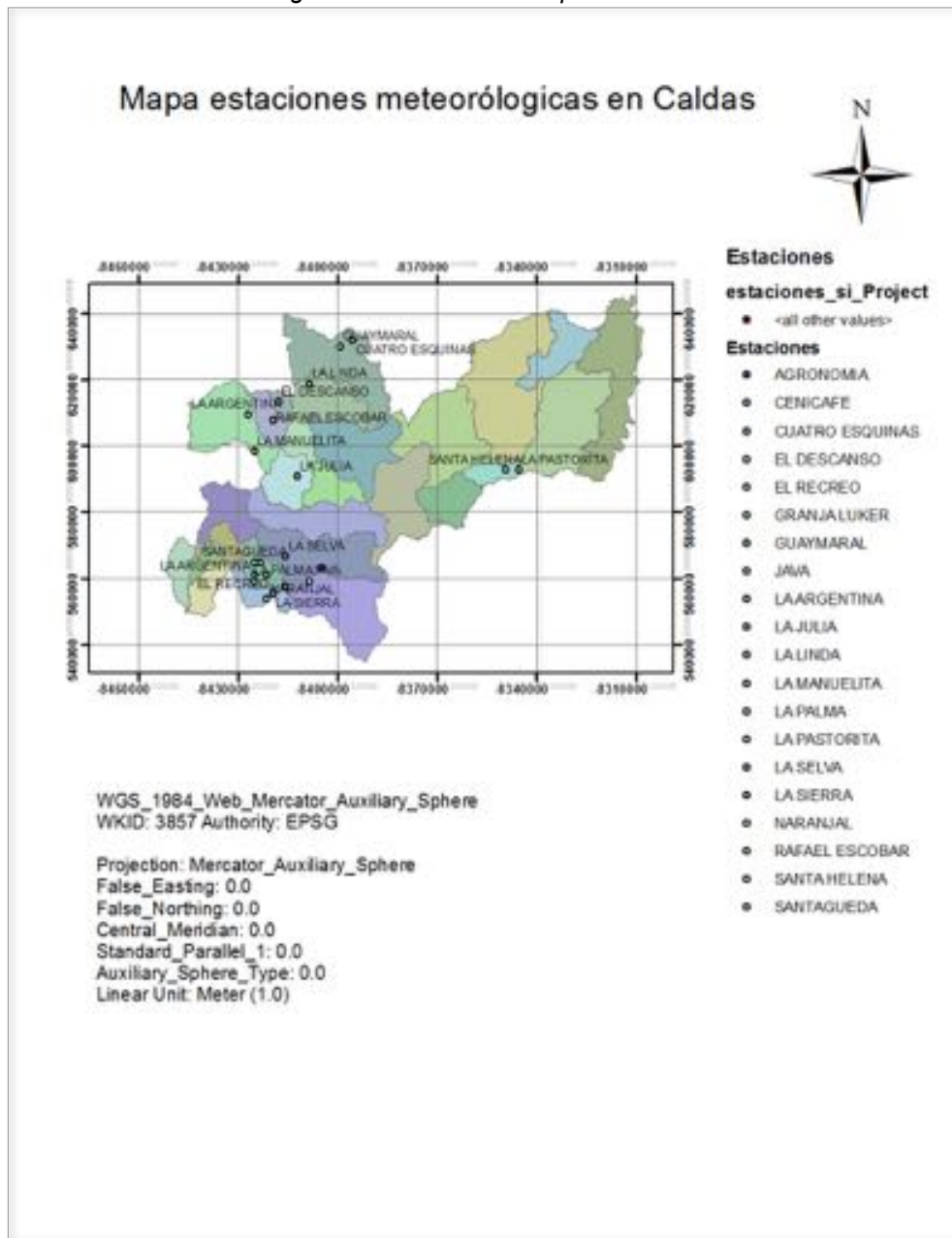
Fueron utilizados los datos de precipitación diaria, de cada una de las 21 estaciones pertenecientes a la red meteorológica de la Federación Nacional de Cafeteros de Caldas, cada una con su respectiva ubicación geográfica, y la altitud (Tabla 2)

Tabla 2. Georreferenciación de las estaciones meteorológicas ubicadas en el departamento de Caldas.

Estaciones	X	Y	Z
CUATRO ESQUINAS	-75.4167	5.6667	1900
GUAYMARAL	-75.4500	5.6500	1600
CENICAFE	-75.6000	5.0000	1310
LA SIERRA	-75.6333	4.9833	1440
NARANJAL	-75.6500	4.9667	1381
LA JULIA	-75.5667	5.3000	1650
AGRONOMIA	-75.5000	5.0500	2088
JAVA	-75.5333	5.0167	1778
LA SELVA	-75.6000	5.0833	1312
EL DESCANSO	-75.6167	5.5000	1803
SANTA HELENA	-75.0000	5.3167	1395
LA LINDA	-75.5333	5.5500	1750
EL RECREO	-75.6500	5.0333	1430
GRANJA LUKER	-75.6833	5.0667	1031
LA ARGENTINA	-75.6833	5.0333	1354
LA PALMA	-75.6833	5.0167	1165
SANTAGUEDA	-75.6667	5.0667	1026
LA ARGENTINA	-75.7000	5.4667	1420
LA MANUELITA	-75.6833	5.3667	1460
RAFAEL ESCOBAR	-75.6333	5.4500	1307
LA PASTORITA	-74.9667	5.3167	1122

Una vez organizados los datos y utilizando el software ArcGis 10.2.2, se creó el mapa de puntos para indicar la posición de cada una de las estaciones dentro del departamento (Gráfico 1).

Gráfico 1. Estaciones meteorológicas ubicadas en el departamento de Caldas.



6.1. 2 Calidad de la información

Actividad 1:

Fue necesario crear una tabla dinámica para facilitar el uso de los datos que históricamente, van desde el año 1988 a 2013, (Cuadro 1).

- **Actividad 2.**

Con base en los datos diarios se generaron las series de lluvia a escala anual para toda la red. De la base de datos se seleccionaron 20 años (Cuadro 2), usando la función Random de Microsoft Excel ® 2010, siguiendo la metodología utilizada por Valencia et al, (2015).

- **Actividad 3.**

Fueron retirados los valores de dos estaciones, en escala de tiempo anual, permitiendo definir los valores iniciales de precipitación para generar la validación cruzada en el que se definen los valores iniciales y los generados en la predicción. (Cuadro 3)

- Tabla dinámica

Cuadro 1. Tabla dinámica de los datos de las estaciones de la red meteorológica en Caldas.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
	= 1988	= 1989	= 1990	= 1991	= 1992	= 1993	= 1994	= 1995	= 1996	= 1997	= 1998	= 1999	= 2000	= 2001	= 2002	= 2003	= 2004	= 2005	= 2006	= 2007	= 2008	= 2009		
9	Etiquetas de fila																							
10	AGRICOLA	2008.8	1483.2	1693.4	1753.2	1326.7	2177.5	1793.9	2888.6	1874.1	1407.5	1832	2627.4	2038	1483.7	1630.5	2223.9	1794.1	2223.9	1487.4	3004.3	1482.7		
11	CDUCAFE	2816.2	2989	2299	2126.2	2027.6	2616.8	2238.5	2982.2	2883.3	2386.3	2862	3061.3	2883.8	2399.9	1971.8	2894.9	2293.2	3206.5	2889.8	3090.2	3267.9	2674.8	81
12	CLMTRQ-ESQUINAS	3090	1083	839	1026	1000	1826	1480	1761	1518	1320	1878	1425	2326	1628	1688	1781	1821	2120	2089.4	2371	2727	1788	
13	EL DESCANSO	2185	1882	2170	1889	2318	1551	2872	1879	1804	1780	2047	2819	1784	1711	2121	1285	1823	1888	1854	2007	2422	894	
14	EL RECREO	2828	2484	2809	2223	2136	3816.1	2025	3027	2483.3	1925	2344.3	1293		2131	2329	3069	2326	3077	2881	2302	3532	2832	
15	GRANJA LUISER	2294.8	2423.1	1734.9	1802	1887.3	2975.1	2099.7	2200.8	2074.1	2321.9	2889.3	2869.4	2886.3	2022.4	2884	2384.8	2362.9	2474.8	2884	2884.8	3495.2	2202.3	28
16	GURUMARAL	2330	1336	1347	1244	2336	1854	2493	1751	1820	1373	2220	2037	2301	1489	1471.5	1708	1871	2334	1708	2330	2118	1841	
17	JANA	2176	1727	2355	1323	2329	1902	1888	1807	2056	1666	1872	2242	1840	1589	1340	2290	1859.3	2502	2136	2668	2810	1883	
18	LA ARGENTINA	2909	2172	1740	1736	2179	2823	2300	2203	2809	2097	2582	3264	2322	2389	2342	2185	1981	3375	3480	3725	4567	2634	
19	LA JULIA	2482	1089	1754	1387	1339	1844	1843	1975	2256	1322	1968	2881	3440	1552	1880	1718	1988	2326	2403	2838	3340	2103	
20	LA LINDA	3020		3078	1424				1234	175	1832	2895	3712	2887	2364	1818	2168	2346	3610	1875	2887	3126	1771	
21	LA MANUELITA	2880	2086	3829	1711	1780	2183	2011	144		1927	3408	3034	2336	1983	1830	2140	2138	2839	2577	3039.9	3658	2403	90
22	LA PALMA	2870	3032	2187	1832	2830	2480	2069	2898	2978	2320	2874	2182	3008	2857	2208.5	2825.5	3401.1	3119.5	2658.5	2836.4	3372.9	2829.8	20
23	LA PASTORITA	3348	2753	2849	2348.3	2830	2021	2982.5	112	2587	3703	3894	4488	4106	4907	2629	3738	3735	3296	4671	4807	8308	3088	
24	LA SELVA	1724	2983	1734	839	941.8	1291	908	1424	2783	3008	2290	3001	3485	1136	1901	2127	3053	2301	2920	2694	3296	1887	
25	LA SIERRA	2529.5	2834	2445.5	2303	2084	2924	2312	2833	2284	1570	2883	3479	3221	2882	2320	2950	2254	3126	3086	3034	3234.8	2677.8	81
26	NARANJAL	2822.8	3075.9	2849.4	2258.7	2573.8	2877.3	2634.3	2812.1	3033.4	2258.6	2887.4	3420.7	3054.7	2857.8	2321.1	2887	2585.4	3493.5	3157.3	2943.8	3888.1	3292.5	81
27	RAFAEL ESCOBAR	2388.7	1703.4	1564.2	1887.3	1593.3	3979.4	1752	3000.3	1960.2	1882.7	2095.3	2622.4	2153.8	1842.8	1748.1	2085.8	1710.2	2343.3	2383.5	2890.5	2897.5	1507	
28	SANTA HELENA	4188	4542.8	3317.4	3907.3	3627.3	3718.6	4089.9	3794.8	4188	3280.2	4583.8	5483.1	4624.2	3879	4013.5	3344.4	3373.7	5030.3	3917.5	4760.3	4487	3608.2	42
29	SANTIQUEDA	1967.7	2180.3	1896.8	2413.6	1891.2	3094.8	2106.9	2190.1	2162.7	3710.4	2198.7	3030.3	2784.4	3088.1	2074.3	2248.7	2298.4	2700.7	2718.1	2803.7	3376.3	2108.9	28
30	LA ARGENTINA(P)	2461.5	2393	3945	2236.3	3940	2253	2387	1739	2483	2139	2881	3405	3430	2517	2219	2930	2338	3130	2933	2728	3430	2708	
31	Total general	36344.8	43361.5	39788.6	37888.7	35487.6	44707.4	41646.7	44814.9	46191.5	40322.4	54887	64465	50505	48066.7	42661.1	50445	47252.3	61308.3	54474.1	61362.8	74147	50257.3	850

- Base de datos anual:

Cuadro 2. Base de datos anual, escogidos aleatoriamente.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
	2011	1996	2009	2001	1989	2012	1993	2004	1987	2001	2005	2008	1990	1991	1998	1988	1995	2002	2007	1991	
1	ARAGONIA	2295.1	1895.6	2842.7	1483.7	3483.2	1859.9	2336.7	1794.1	2487.3	2223.9	2223.9	3004.3	1893.4	2177.5	2812	2008.8	2827.4	2830.5	1887.4	2230.2
2	CANTABRIA	2748.6	2982.2	2474.5	1899.9	2589	2457.5	2017.6	2193.1	2388.3	2484.9	3106.5	3267.9	2298	2614.8	2842	2818.1	3061.3	2971.8	3096.2	2125.2
3	CANTON DE QUINCE	1763	2761	1788	2828	1988	2890	3000	1821	2200	1781	2200	2727	888	1828	2378	2083	2428	1888	2371	2128
4	EL ESCORIAL	95	3479	894	1711	1680	1410	2208	1815	1788	1088	1888	1412	1170	1591	2047	1189	2819	1191	2097	1888
5	EL ESCORIAL	2388.9	3307	2882	2131	2484	2312	1134	2348	1815	3044	3077	3832	2828	2818.1	2884.5	2888	2285	2328	2300	2218
6	ISLA DE LISER	2324	2021.8	2130.3	2022.4	2415.1	1992.9	2847.3	2341.9	1811.3	1384.3	2474.8	3405.2	1734.3	1971.1	2889.3	2194.8	1983.4	1984	2848.8	1801
7	GUANABANA	1877	1781	1841	1489	1898	1994	1388	1871	1673	1738	1934	1618	1747	1884	2128	1888	2887	3477.5	2819	1184
8	JAVA	1788	1907	1883	2889	1727	2397	1828	1888.3	1888	2280	2800	2828	2888	1880	2872	2178	2280	1940	2888	1328
9	LA ARGENTINA	3049	2129	2854	2889	2171	3081	2174	2483	2087	1045	3378	4587	1740	1818	2890	2888	2184	2181	2128	1798
10	LA LUNA	2878	2878	2103	1882	2289	2389	1888	2272	1718	2128	1941	1734	1844	1888	2482	2882	2882	2882	2838	1887
11	LA LUNA	2389	2224	1771	2384	2388	2388	2348	1892	2384	2819	2828	2878	2895	2895	2895	2895	2895	2895	2895	1404
12	LA MANUELITA	2821	144	2489	1883	2088	1898	1780	2118	1907	2140	2808	2808	1828	1383	2488	2881	2881	2881	1881	1711
13	LA PALMA	2842	2888	2889.8	2887	3182	1808.4	1881	2401.1	2320	2828.9	3118.8	1871.9	1387	2480	2878	2978	3180	2208.5	2838.8	2802
14	LA PALMERA	4288	112	3088	4907	1713	3084	2851	2735	2709	2738	3188	4504	1848	2011	2896	1848	4488	2819	4887	2848.5
15	LA SELVA	2871	2424	1887	2138	1883	2028	841.8	2081	2008	2127	2810	1098	1714	2041	2230	1724	3001	1801	2884	808
16	LA SERA	3078.7	2829	2877.6	2882	2884	2748	2884	2284	1170	1991	3118	3124.8	2441.5	1914	2899	2813.3	3471	1500	3094	1818
17	NARRUAL	3100.8	1811.1	2290.3	1887.8	3078.8	1881.8	2179.8	2488.8	2288.8	2887	3488.8	3888.1	2888.8	2877.3	2887.8	2820.8	3400.7	2320.4	1848.8	2288.7
18	REVAL ESCOBAR	2087	2020.8	1807	1841.8	1700.8	2300	1880.8	1720.2	2882.7	2088.8	2840.3	2887.3	1888.2	1878.4	2088.8	2188.7	1821.8	1748.2	1880.3	2887.3
19	SANTA HELENA	4174.9	1784.8	2889.2	3479	4540.8	4087.3	3417.3	3571.7	3380.1	3144.4	1980.3	4887	3317.4	1718.4	4883.8	4188	1885.1	4813.5	4780.3	2887.7
20	SANTAGUEDA	2888.1	2380.1	2100.8	2084.1	2180.8	2028.7	2881.2	2108.4	1710.4	2148.7	2700.7	3378.3	1888.8	2084.8	1898.7	1847.7	3080.3	2074.1	2801.7	2818.8
21	LA ARGENTINA (C)	1841	2739	1709	2517	1883	2128	1883	1888	2159	1888	3180	1480	1848	2218	2880	2401.8	3485	2239	1738	2288.8
22																					
23																					
24																					
25																					
26																					
27																					
28																					

Cuadro 3. Valores retirados por estación.

	2013	1995	2009	2001	1989	2012	1992	2004	1997	2003	2005	2008	1990	1993	1998	1988
AGRONOMIA		1866.6	1642.7	1483.7	1463.2	1859.9	1316.7	1734.1	1407.3	2223.9	2223.9	3004.3	1655.4	2177.5	1812	2008.6
CENICAFE			2474.5	2399.9	2569	2457.5	2017.6	2293.2	2269.3	2494.9	3106.5	3267.9	2299	2614.8	2662	2818.2
CUATRO ESQUINA	1763			1628	1065	1930	1000	1801	1310	1781	2220	2727	859	1816	1378	2093
EL DESCANSO	15	1679			1682	1410	1318	1615	1780	1265	1886	2422	1271	1551	2047	2165
EL RECREO	2389.3	3307	2652			2313	1134	2316	1915	3044	3077	3532	1609	1616.1	2564.5	2838
GRANJA LUKER	2324	2201.8	2230.3	2022.4			1667.3	2342.9	1531.9	2384.8	2474.6	3495.2	1734.9	1971.1	2689.3	2194.8
GUAYMARAL	1977	1751	1841	1469	1336			1871	1573	1738	1534	3118	1747	1854	2220	2330
JAVA	1766	1907	1883	1589	1727	2197			1666	2290	2502	2810	1555	1950	1672	2176
LA ARGENTINA	3269	2233	2654	2389	2172	3081	2174			2165	3375	4567	1741	2823	2592	2909
LA JULIA		1975	2103	1552	1089		1339	1988			2326	3341	1754	1844	1968	2482
LA LINDA	2139	1224	1771	2164		2103		2146	1832			3526	2078		2995	2020
LA MANUELITA	2921	144	2403	1983	2066	1955	1780	2118	1927	2140			1829	2183	3458	2883
LA PALMA	2942	2698	2839.6	2657	3192	1908.4	1831	2401.1	2320	2815.5	3119.5			2460	2874	2973
LA PASTORITA	4255	112	5098	4907	2753	5054	2851	3735	3703	3738	5296	6506			3896	5348
LA SELVA	2471	2424	1997	2136	2363	2033	941.6	2053	2008	2127	2531	3296	1724			2724
LA SIERRA	3073.7	2633	2677.8	2682	2834	2743	2084	2254	1570	2953	3226	3224.8	2443.5	2924		
NARANJAL	3102.9	2812.1	3292.3	2857.8	3075.9	2861.6	2173.6	2585.4	2258.6	2887	3435.5	3868.1	2649.4	2677.3	2987.4	
RAFAEL ESCOBAR	2097	2020.3	1507	1842.8	1703.4	2102	1593.3	1710.2	1682.7	2085.8	2345.3	2697.5	1564.2	1979.4	2095.3	2168.7
SANTA HELENA	4274.5	3794.8	3659.2	3879	4542.6	4097.3	3627.3	3371.7	3280.2	3244.4	5150.3	4887	3317.4	3716.6	4583.8	4198
SANTAGUEDA	2359.1	2350.1	2100.9	2086.1	2180.3	2015.7	1691.2	2238.4	1710.4	2248.7	2700.7	3376.3	1996.8	2084.6	2398.7	1967.7
LA ARGENTINA (P)	2642	2739	2709	2517	2393	2226	1983	2338	2159	2935	3130	3450	1945	2213	2691	2401.5

6.1.3 Fase 2. Análisis Exploratorio y Estructural de datos

La aplicación de la geoestadística constituyo una herramienta de suma importancia para el análisis gráfico de los datos. La evaluación de la forma, la distribución y el cálculo de medidas de localización, variabilidad y correlación fueron importantes para establecer si algunos supuestos necesarios para la aplicación de la teoría geoestadística resultan válidos o para definir que procedimiento de predicción es el más conveniente. (Giraldo, 2000)

Histograma: Durante el análisis exploratorio de datos se generaron los histogramas para cada uno de los 20 años. Bajo la opción de atributo precipitación se obtuvo valores de medida de centralización (media y mediana) similares, con diferencias menores a 1, así mismo las medidas de forma como el sesgo o Skewness arrojaron valores entre los -0,8 hasta 0,03 como lo indica la gráfica 2, año 1992. Por su parte la medida de forma Kurtosis, registró valores entre los 3,06 como lo muestra la gráfica 3, año 1989 hasta 4,9. La mayoría de los métodos geoestadísticos sólo son óptimos si la variable de estudio (precipitación para este caso) sigue una distribución normal. Es por ello que fue necesario tener en cuenta el cumplimiento de condiciones como: Que tengan una única moda que coincida con su media y su mediana, que la curva normal sea asintótica al eje de abscisas, que el coeficiente de sesgo o Skewness sea igual a cero (0) y que la Kurtosis es igual a tres (3). La herramienta Geostatistical Analyst no permite realizar transformaciones cuando se tienen menos de 20 estaciones, por ello no fue posible aplicar la transformación Box-Cox. Se hizo necesario retirar dos datos (estaciones) por año para la Validación Cruzada. En el anexo A se observan los histogramas para cada uno de los años utilizados.

Gráfico 2. Histograma año 1992

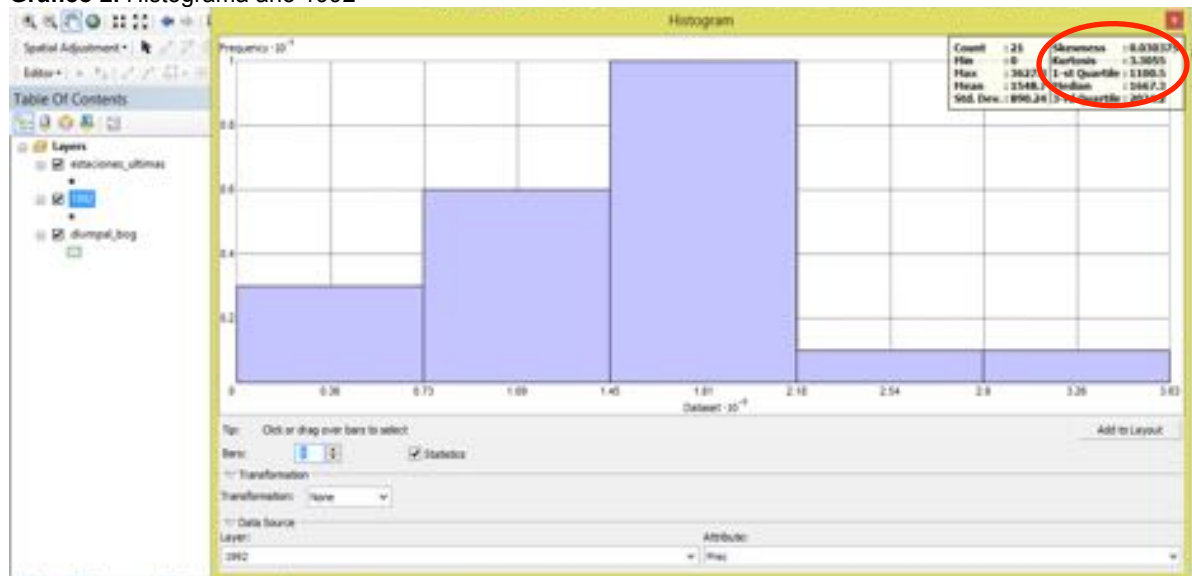
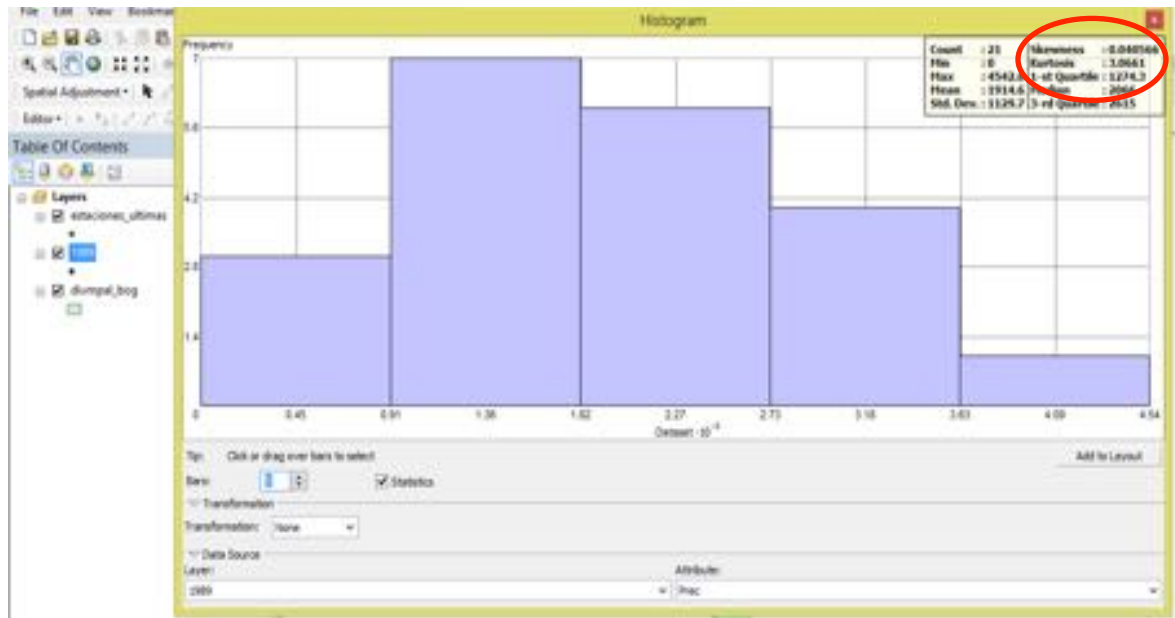
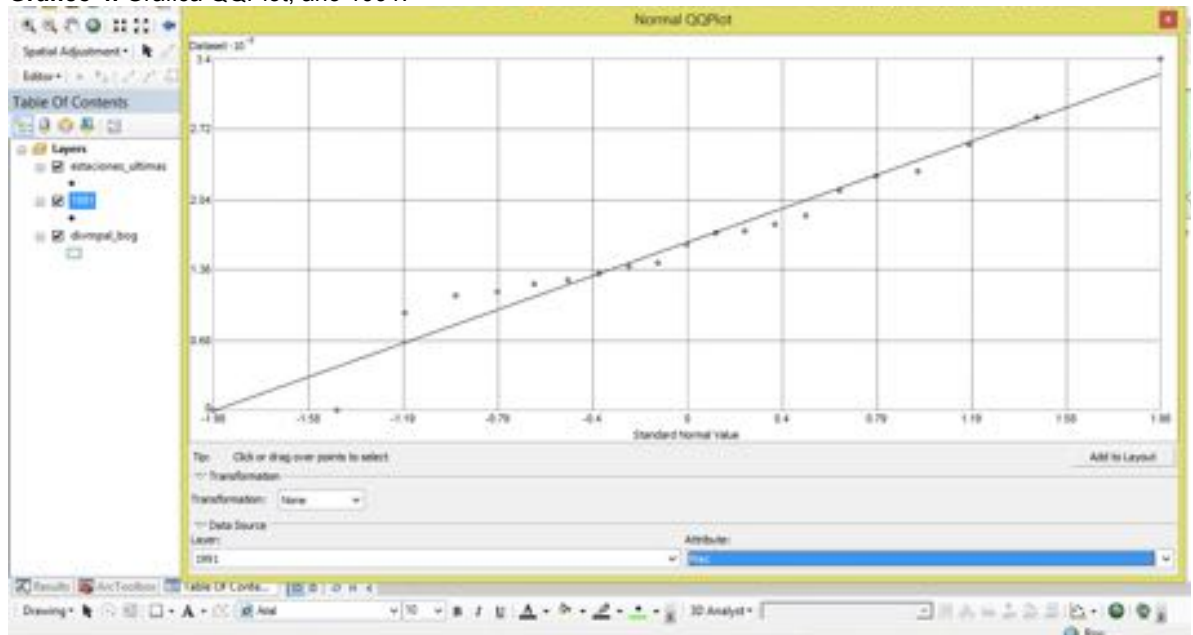


Gráfico 3. Histograma año 1989



QQ Plot: El gráfico 4 muestra los cuantiles de la distribución normal (eje X) y en el eje Y los cuantiles de la variable precipitación, proyectados en la recta, y los puntos que representan las estaciones. Se evidencia que, los valores observados frente a los esperados (recta) se aproximan, no son asimétricos, se ajustan a una distribución normal. En el anexo B se puede observar la distribución de los datos para cada uno de los años. En este caso solo se utilizó un conjunto de datos pertenecientes a la variable precipitación.

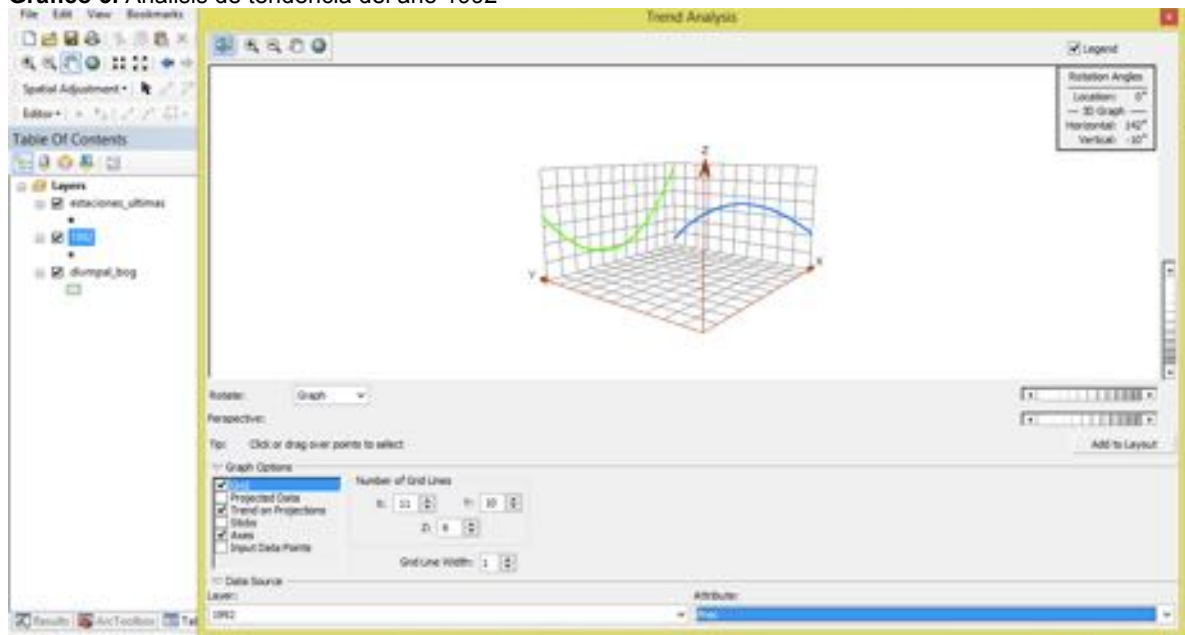
Gráfico 4. Gráfica QQPlot, año 1991.



Trend Analysis: En este gráfico se muestra cada línea representando el sitio y el valor de cada punto de muestreo. Los resultados muestran la tendencia de los datos, así la línea verde (tendencia de la variable desde el Este al Oeste), como la azul que representa la tendencia de la variable desde Norte hacia el Sur, son curvas; la tendencia es de segundo grado. Esto permite concluir que la tendencia más fuerte se presenta en dirección Norte- Sur en la que la línea de tendencia es más gruesa. La tendencia se puede dar por la presencia de las estaciones en zonas planas y de cordillera, sin representación significativa en las zonas bajas. En estadística se utilizan series de tiempo para eliminar las tendencias y generar un modelo. En este caso se realizó visualmente.

Se presentaron tendencias de segundo orden o cuadrática como se observa en el gráfico 5 año 1992. Ver anexo C Trend Analysis para los años utilizados.

Gráfico 5. Análisis de tendencia del año 1992



En el trazado de la curva en el plano YZ las precipitaciones aumentan desde el norte al centro, y disminuyen ligeramente en sentido sur. Según el plano XZ, la precipitación disminuye en sentido este, siendo pronunciado en el centro y aumentando de nuevo en sentido oeste.

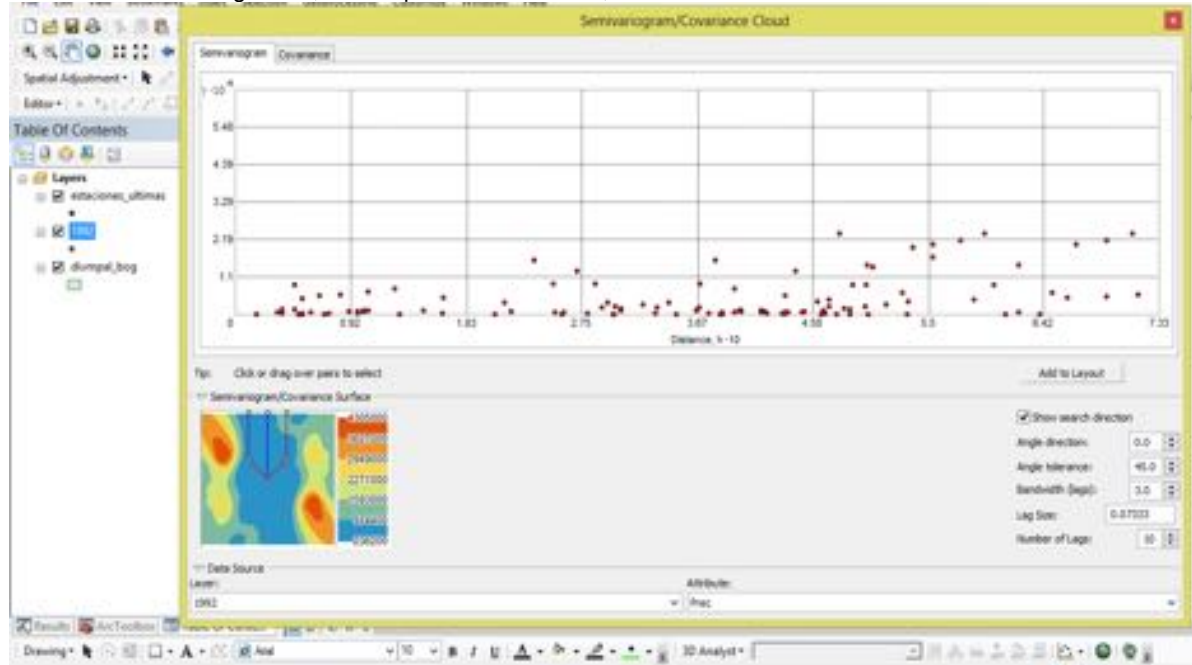
Modelamiento del semivariograma: La opción attribute, permitió elegir la variable a analizar, en este caso la precipitación. Con esta herramienta se analizó intuitivamente la presencia de anisotropía a 45 grados, de ángulo de tolerancia, con máximo de distancia de 0.07333 metros. Al desplazar la flecha de dirección se determinó pocas variaciones en los puntos analizados, evidenciando la posible isotropía (intuitivamente).

Los aspectos continuidad espacial y vecindad están íntimamente relacionados, ya que sí los datos son más continuos en una dirección que en otra, la forma de la búsqueda de vecindad es una elipse, cuyo eje mayor será paralelo a la dirección de máxima continuidad. Si no hay evidencia de anisotropía, la forma de búsqueda de vecindad será el círculo donde la orientación ya no es relevante (Boada, 2015).

El semivariograma permitió evaluar la autocorrelación espacial en el que se exploró identificó y cuantificó la dependencia de la variable precipitación (AnexoD). Seleccionando el cuadro de dialogo Show Search Direction (mostrar dirección de búsqueda) se comprueba de modo interactivo, que al variar el ángulo de dirección y anchura de banda, varia también la nube de puntos del semivariograma, detectando la existencia de una variable con distribución espacial isotrópica (gráfico 6). En el caso de anisotropía, la variable tendría una influencia espacial más extensa en un sentido que en otro, siendo la disposición de los puntos

diferente en cada dirección (Boada, 2015). Así mismo se identificaron valores extremos, aquellos que se alejan del eje X. En el gráfico es posible determinar los pares de lugares que se encuentran más próximos, apareciendo a la izquierda del eje X, cerca al origen del eje Y. Esto indica que los valores son similares,

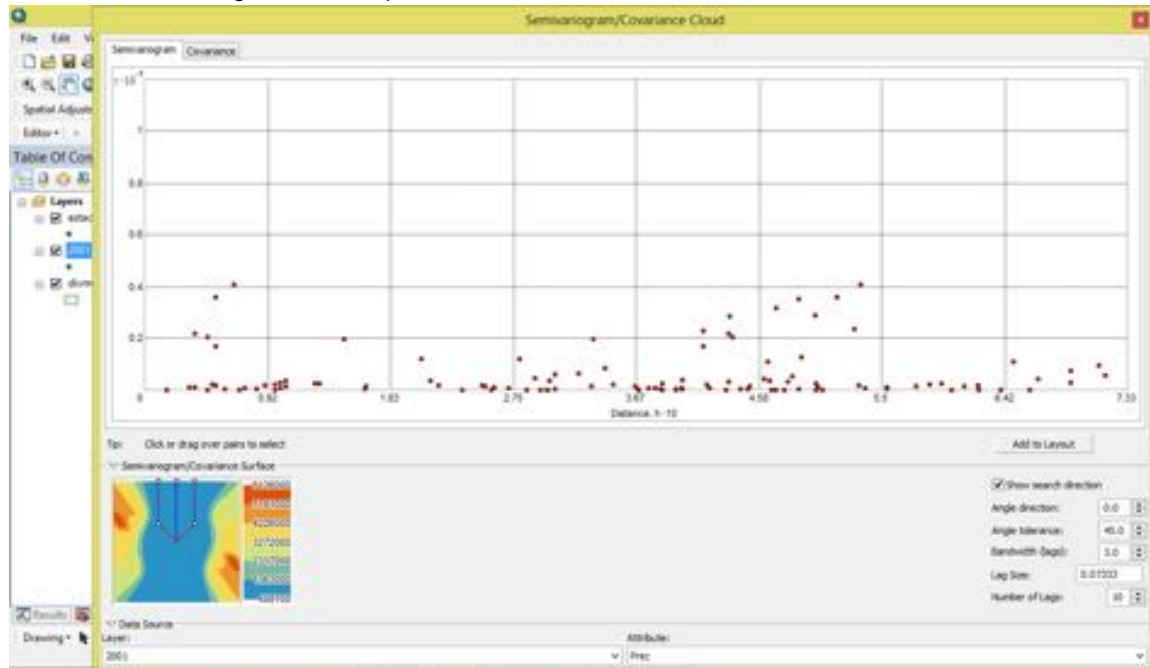
Grafico 6. Semivariogramas, fase exploratoria de datos, año 1992.



Los puntos observados hacia la derecha con respecto al eje X y más altos sobre el eje Y, indican que se encuentran separados por mayor distancia, indicando dependencia espacial. De esta forma se concluye que la autocorrelación espacial que depende de la distancia entre dos lugares genera isotropía. Si bien, la influencia es direccional se generará anisotropía.

Como es perceptible en la nube de puntos del gráfico 7, la forma depende del número de intervalos en que se agruparon las distancias y de la amplitud. Con una amplitud demasiado grande pueden desaparecer auto correlaciones de pequeño alcance. Ante la presencia de tamaños pequeños, es posible que se presenten vacíos o pocos pares de puntos.

Gráfico 7. Semivariograma fase exploratoria de datos, año 2001.



Si no hay influencia direccional en los pesos de los datos, se consideran los puntos en todas las direcciones y la forma de la búsqueda de la vecindad es un círculo (ejes mayor y menor igual longitud) como se observa en el gráfico 8 y 9 del método IDW, en donde los datos son aleatorios, no hay anisotropía. Si hay influencia direccional en los datos (influencia del viento predominante), se adopta la forma de una elipse con el eje mayor paralelo al fenómeno considerado, gráfico 10 y 11.

Gráfico 8. Método IDW, año 1995

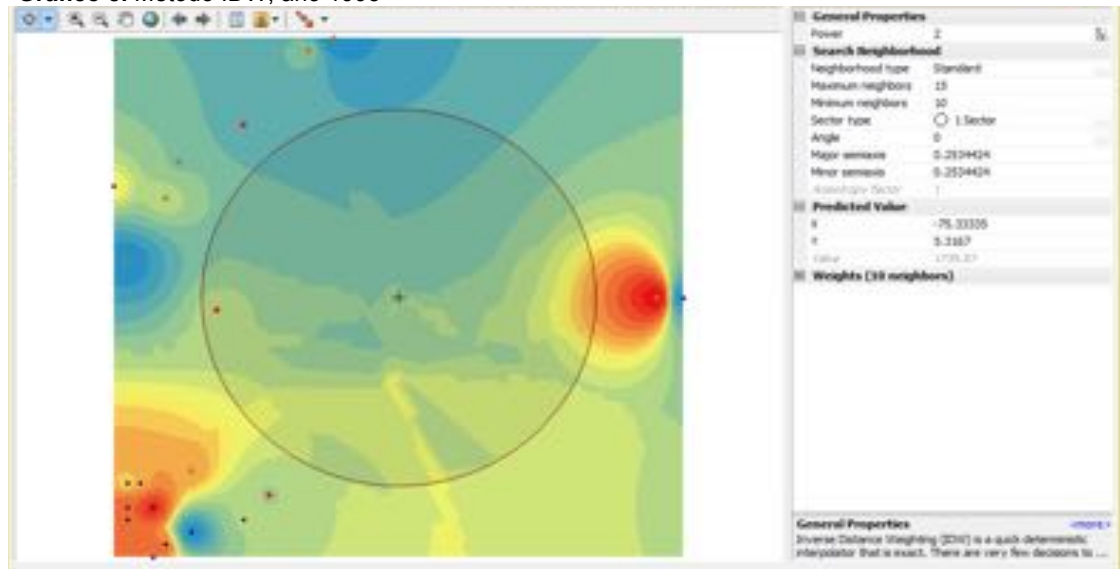


Gráfico 9. Método IDW, año 1991,

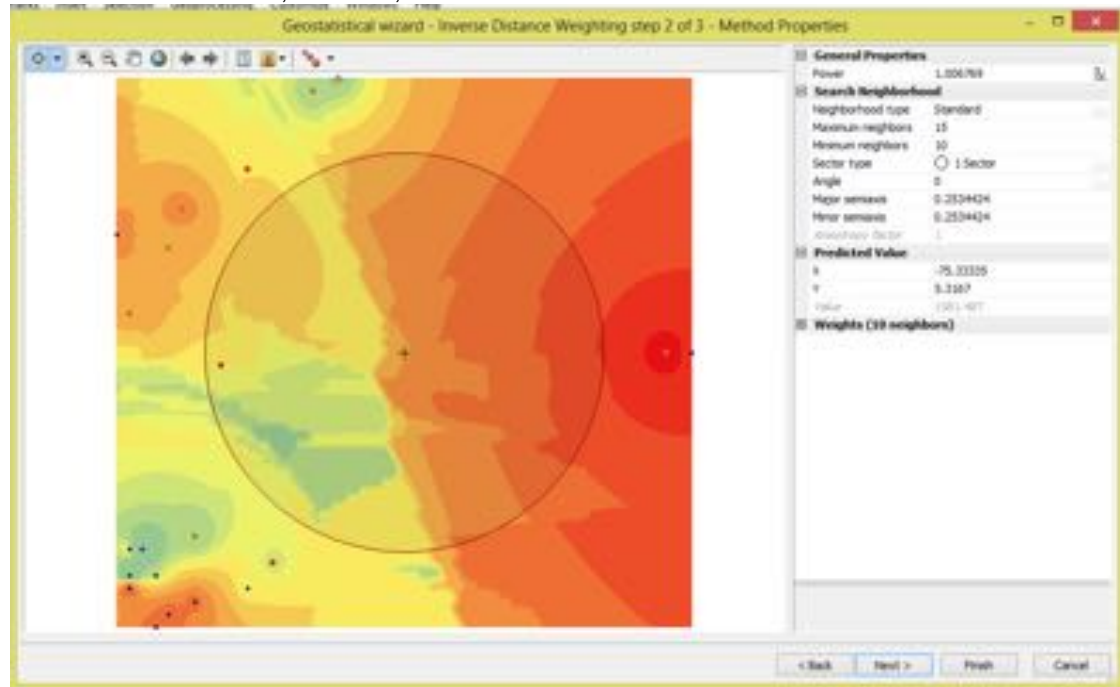


Gráfico 10. Global Polynomial Interpolation, año 1991.

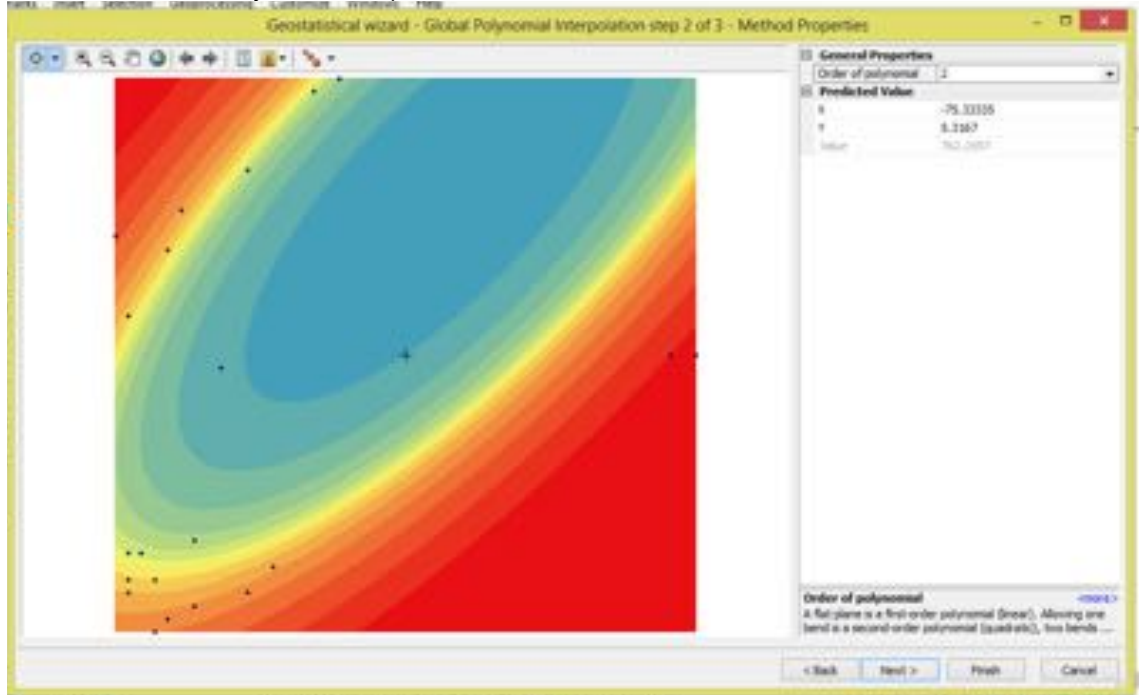
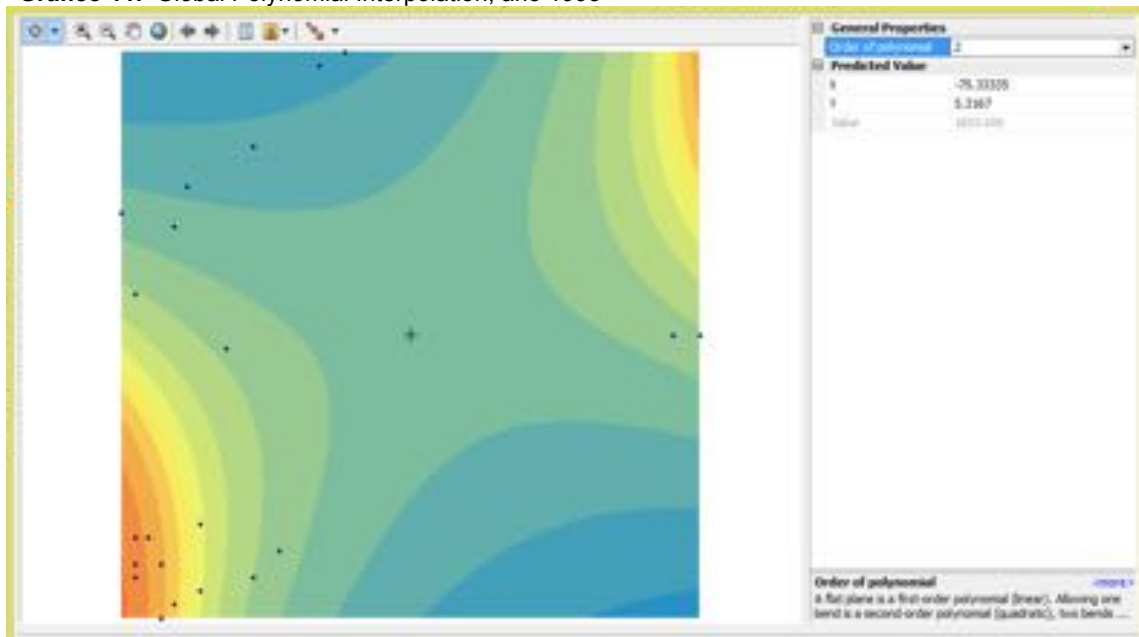


Gráfico 11. Global Polynomial Interpolation, año 1995



Validación Cruzada: Esta herramienta permitió comprobar la validez del modelo e indicar cuál es el mejor. La validación cruzada utiliza todos los datos para estimar el modelo de autocorrelación. Al quitar 2 de las 21 estaciones predijo con el resto de la muestra los valores ausentes, para después compararlos con los valores observados.

Se compararon los valores del error entre los métodos de interpolación (Anexo F), tomando como referente aquel que presentaba el menor valor, en este caso el IDW (a la izquierda del gráfico 12).

Los resultados en los métodos Kriging son satisfactorios solo si la media estandarizada de los errores es cero y error cuadrático es cercano a cero. En el gráfico 13 (K. Universal vs. IDW) se muestra un valor de 0.5464 para la media estandarizada y 1.206 como error cuadrático, alejándose de cero.

Al comparar los valores de Root Mean Square, se observa que los menores valores se presentan en el método IDW, y los mayores con el método Kriging Ordinario (gráfico 14). Es posible decir que el método que mejor se ajusta a este conjunto de datos de precipitación es el método IDW.

Gráfico 12. Paralelo entre métodos IDW vs. Global Polynomial Interpolation, año 1989.

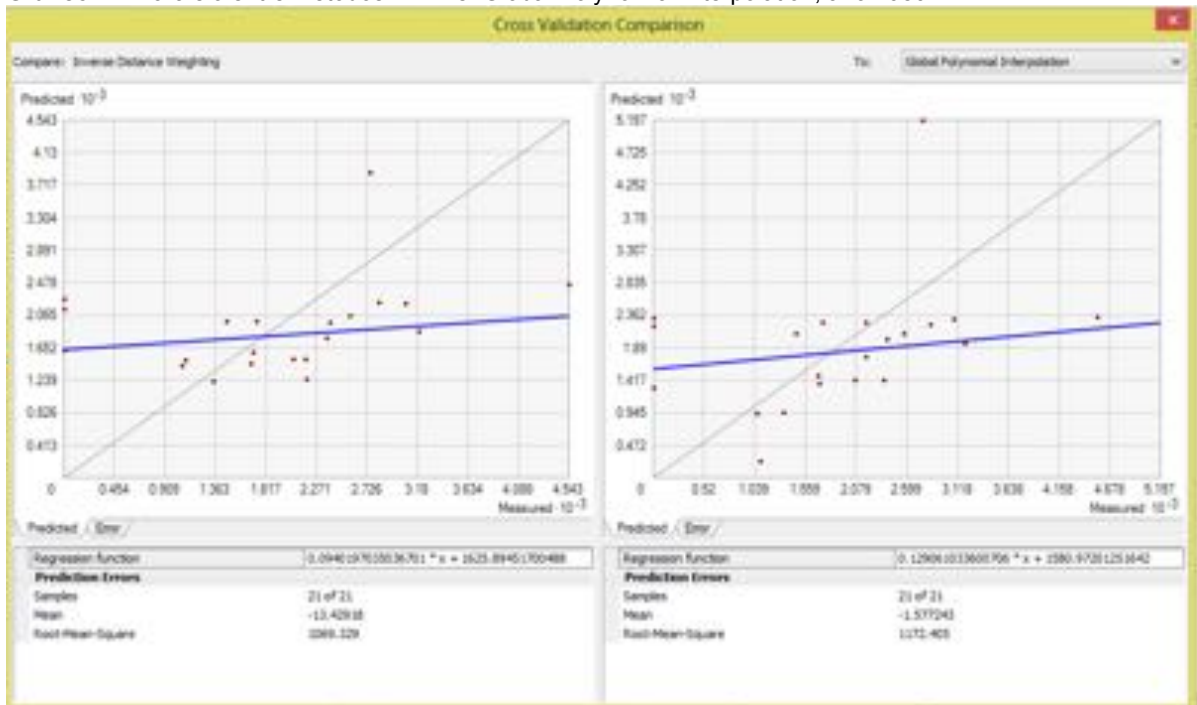


Gráfico 13. Validación Cruzada, método IDW vs. Kriging Universal, año 1989.

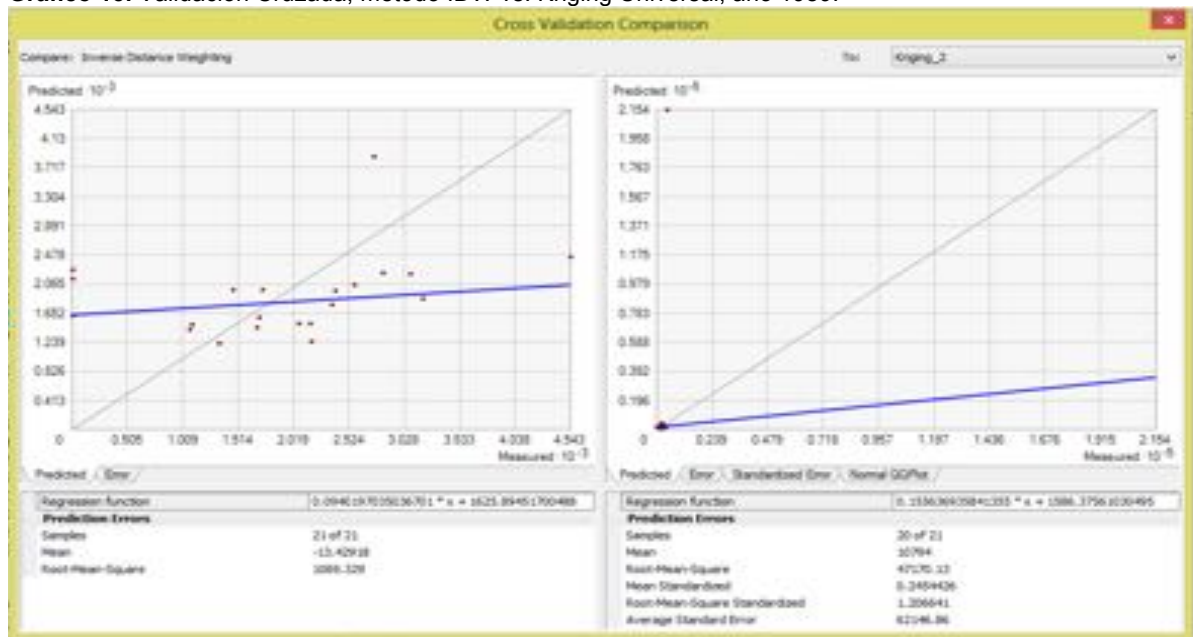
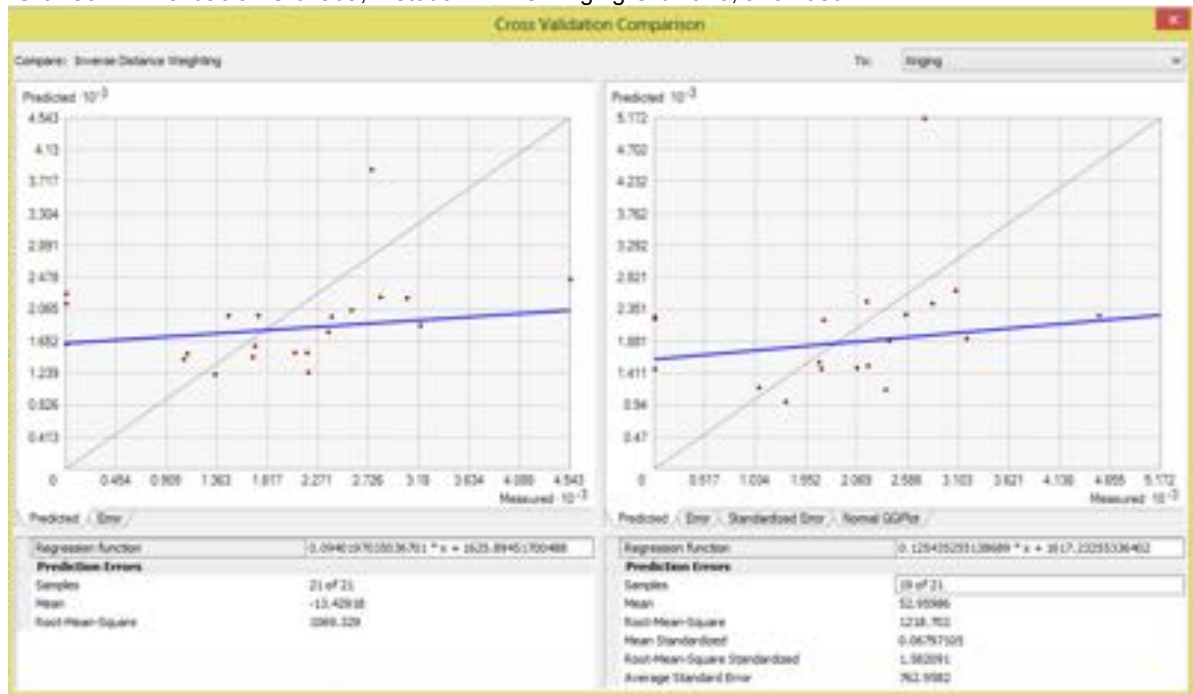


Gráfico 14. Validación Cruzada, método IDW vs. Kriging Ordinario, año 1989.

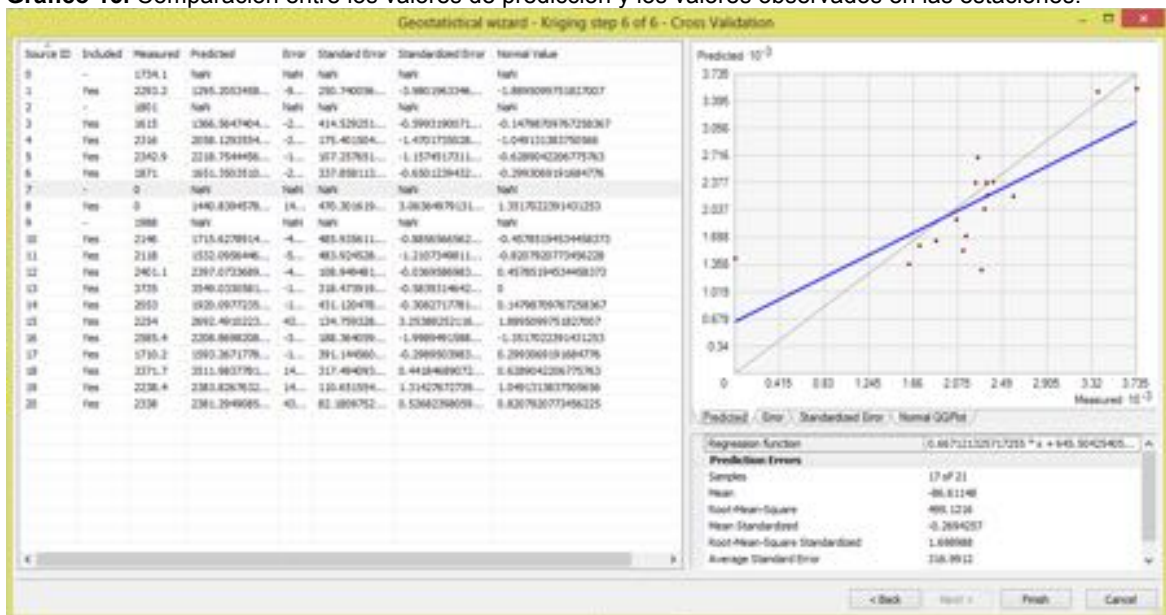


Para comparar las predicciones con los valores reales se elaboran varios gráficos y algunos estadísticos, iguales para cada caso. Un gráfico que muestra valores predichos frente a valores observados, una línea de ajuste aparece con trazado continuo y la ecuación aparece debajo del gráfico. En las diferentes opciones que

se observan en el gráfico de la validación cruzada, se encuentre el gráfico de los errores o residuales, que muestra la diferencia entre los valores observados y los predichos; en el eje horizontal los valores observados y en el vertical los residuales; el grafico de los errores estandarizados que se obtienen dividiendo los errores por la desviación estándar estimada del Kriging. Los tres gráficos ayudan a valorar los datos predichos del modelo. Mientras que el grafico de probabilidad normal QQPlot compara los cuantiles de los errores estandarizados con los cuantiles de una distribución normal. Los errores estarían normalmente distribuidos si coinciden con la línea de trazos. En caso de que se pretendiese elaborar mapas de cuantiles o mapas de probabilidad se requiere que los datos tengan distribución normal.

En algunas estaciones, en especial en los metodos Kriging Ordinario y Universal no se arrojo ningun valor como se observa en el gráfico 15 algunas estaciones no cuentan con valores de prediccion ya que el método no encontro vecinos cercanos. El IDW cuenta con uno máximo de 15 vecinos y un mínimo de 10, mientras que los geoestadísticos utilizan un máximo de 5 vecinos y un mínimo de 2.

Gráfico 15. Comparación entre los valores de predicción y los valores observados en las estaciones.



SUPERFICIES DE PRECIPITACIÓN GENERADAS CON CUATRO MÉTODOS DE INTERPOLACIÓN

En el Anexo E se presenta las diferentes superficies de interpolación generadas para cada uno de los años con cuatro métodos diferentes. El cálculo del error medio cuadrático y la generación de estas superficies facilita determinar el mejor ajuste del modelo y mejorar la calidad del mapa. En los gráficos 16,17,18 y 19 (en los que se presenta los mapas de interpolación con diferentes métodos) se observa que el el método IDW permite mantener las superficies de acuerdo con la ubicación de las estaciones, hay menos variabilidad en la predicción, mayor suavizado, presentando mejor ajuste a zonas donde la topografía no es plana, característico de la zona Andina.

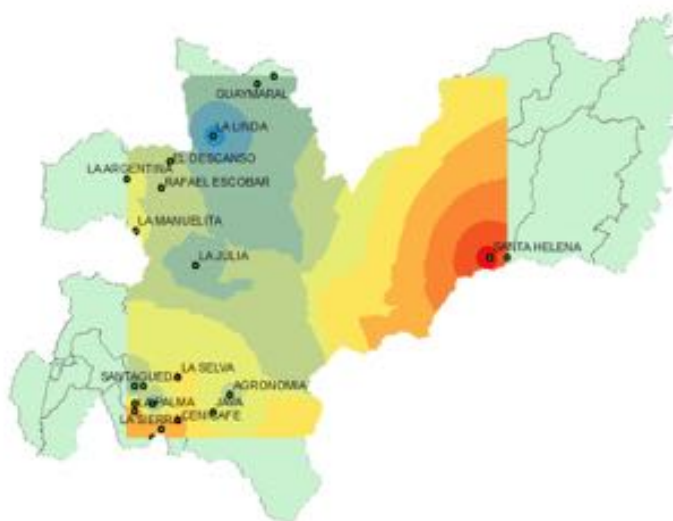


Grafico 16. Superficie de precipitación generada con el método IDW año 1989

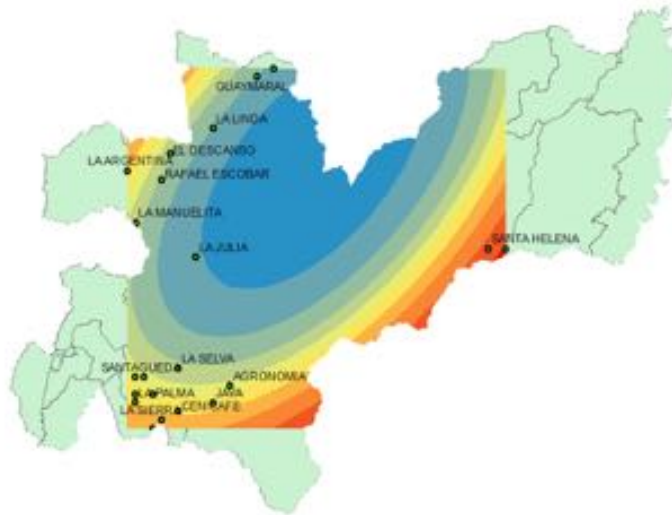


Gráfico 17. Superficie de precipitación generada con el método Global Polynomial Interpolation año 1989

Se debe considerar que los resultados generados con los métodos Kriging dependen en gran medida del adecuado análisis estructural de los datos y la previa elección del semivariograma.

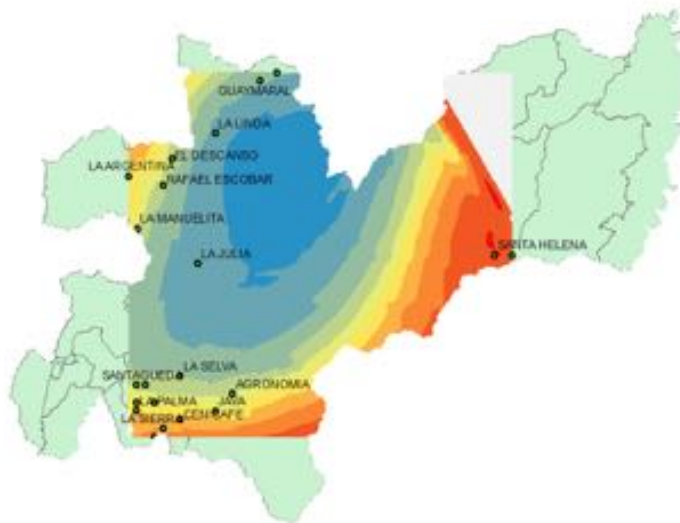


Gráfico 18. Superficie de precipitación generada con el método Geoestadístico Kriging Ordinario año 1989

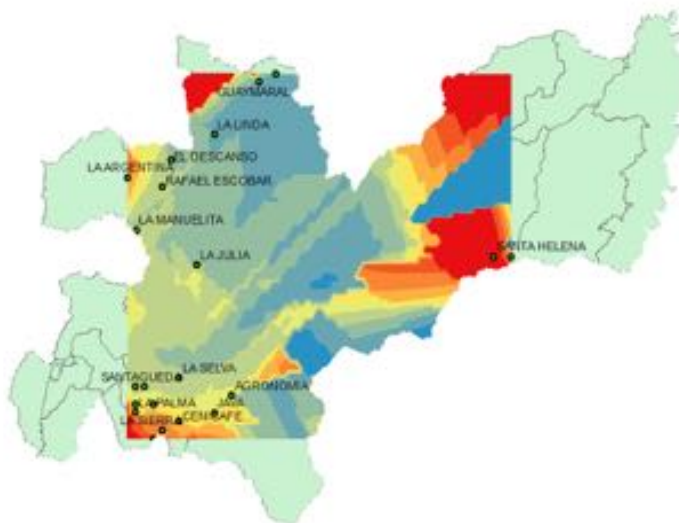


Gráfico 19. Superficie de precipitación generada con el método Geoestadístico Kriging Universal año 1989

El cuadro 4, presenta la comparación entre los valores originales de precipitación en cada una de las estaciones retiradas, y los valores generados por cada uno de los métodos de interpolación empleados. Se evidencia algunos años sin datos debido a que la herramienta no los calcula ante la ausencia de vecinos cercanos. Los datos negativos por su parte, hacen inferencia al número reducido de estaciones utilizadas en la interpolación.

De acuerdo a los resultados del análisis de varianza para los datos estimados, (Tabla 3), es posible determinar que no existe diferencias significativas entre ellos. Sin embargo los métodos IDW y el K. Ordinario muestran mejor ajuste, al tener la mayor cantidad de valores similares a los iniciales. Es importante tener en cuenta que el número de estaciones cercanas fue un condicionante para el método K. Ordinario, reflejando diferencias con respecto al IDW en cuanto a la capacidad para predecir valores en ausencia de vecinos cercanos, en zonas andinas ecuatoriales.

Cuadro 4. Valores de precipitación inicial vs. Valores generados con los diferentes métodos de interpolación utilizados.

Estacion	Año	Valor real precipitacion	IDW	GPI	K. Universal	K. Ordinario
Agronomia	2013	2295.1	2012,7255	720,4958	2911,5402	1270,9684
Cenicafe	2013	2748.6	2545,8639	2107,3534	2860,4031	2682,3153
Cenicafe	1995	2982.2	2554,9245	2519,3173	2544,5371	2555,9768
Cuatro esquinas	1995	1761	1700,9461	1999,8277	-58,1905	1845,5867
Cuatro esquinas	2009	1788	1884,5255	1814,2079	2495,7076	2059,488
El Descanso	2009	894	1860,9864	1896,7748	2014,5370	1844,7449
El Descanso	2001	1711	1976,3014	1931,5032	2065,3495	2035,0609
El Recreo	2001	2131	2373,2804	2356,0264	2330,4406	2335,7504
El Recreo	1989	2484	2264,6710	2209,2344	2162,8514	2245,3248
Granja Luker	1989	2411.1	2145,8585	2321,1644	2231,0835	2223,2545
Granja Luker	2012	1992.9	2151,5675	2030,6413	1902,0187	1992,7745
Guaymaral	2012	1954	2005,5987	2138,7239	2117,6453	2182,2981
Guaymaral	1992	1356	996,8272	754,9038	1816,1920	737,5444
Java	1992	1529	1581,6037	1525,9974	1633,5713	1610,0424
Java	2004	1859.5	2115, 1265	1925,1540	1947,4905	
La Argentina	2004	2481	1846,6517	1862,6933		1440,8394
La Argentina	1997	2097	1644,7307	2836,4674	2905,2095	
La Julia	1997	1322	1602,5312	2263,8950		972,8592
La Julia	2003	1718	2139,9999	1091,5006	839,1883	
La Linda	2003	2166	1265,0000	1328,2688	1540,5253	1191,0275
La Linda	2005	2610	1994,5243	2070,8058	-283,8984	99,8199
La Manuelita	2005	2839	2523,4174	1859,5393	3192,6714	2847,2471
La Manuelita	2008	3658	3367,7376	3404,4721	3518,6220	3610,1024
La Palma	2008	3372.9	3422,3559	3470,7570	3721,3527	3457,3083
La Palma	1990	1167	1974,63840	2024,7372	2119,4807	2040,0864

La Pastorita	1990	2849	2752,32573	3014,1156	3327,1903	2976,8125
La Pastorita	1993	2011	2990,0996	4337,7505		4451,3955
La Selva	1993	2241	2209,2987	2014,0553	1665,8213	2003,9766
La Selva	1998	2230	1143,1526	1279,6440	1442,1982	1222,9423
La Sierra	1998	2993	1115,4047	1310,5023	1173,9088	1262,8946
La Sierra	1988	2523.5	877.2	1933,726966	1369,46777	1662,72964
Naranjal	1988	2922.8	482.3	1966,178552	456,52003	1348, 33111
Naranjal	1999	3420.7	2967,8779	2863,1146	3704,4404	2973,1121
Rafael Escobar	1999	2622.6	3000,3146	3307,1172	3227,8010	3293,7928
Rafael Escobar	2002	1746.1	1725,8918	1731,8890	1756,7830	1742,6508
Santa Helena	2002	4013.5	2232,5818	2464,2123	- 43540,4382	2485,4812
Santa Helena	2007	4760.3	4033,1632	4392,2540	15958,1152	4354,1323
Santagueda	2007	2803.7	2645,0772	2756,2620	2872,4145	2754,2605
Santagueda	1991	1613.6	1540,8052	1572,3920	1467,5585	1426,2626
La Argentina (P)	1991	2236.5	1794,277	1784,4288	1952,8140	1773,7448

Tabla 3. Análisis de varianza de los valores de precipitación generados en cada uno de los métodos de interpolación.

ANÁLISIS DE VARIANZA

<i>Origen de las variaciones</i>	<i>Suma de cuadrados</i>	<i>Grados de libertad</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Probabilidad</i>	<i>Valor crítico para F</i>
Entre grupos	36166200,7	4	9041550,174	0,73531461	0,56889474	2,417962542
Dentro de los grupos	2397752268	195	12296165,48			
Total	2433918469	199				

Por su parte el cuadro 5 indica los diferentes valores de Error Medio Cuadrático arrojados en cada uno de los métodos de interpolación. Es importante definir que 13 de los 20 resultados fueron menores con el método determinístico IDW, 4 con Kriging Ordinario y 3 con el método GPI. Los valores más altos registrados se observaron con el método Kriging Universal. Aunque el método IDW generó menor cantidad de errores, estos son datos gruesos, que dan importancia a aspectos como la normalidad de los datos, densidad y distribución de estaciones.

Cuadro 5. Valores de error medio cuadrático generado con los diferentes métodos de interpolación determinísticos y geoestadísticos.

Año	Error con IDW	Error con GPI	Error con K. Universal	Error con K. Ordinario
2013	1120,799	980,0975	2547,318	1145,445
1995	1306,343	1559,185	2970,744	1508,366
2009	869,2017	883,5488	90459,73	906,0802
2001	849,4642	813,7881	24149,57	858, 5866
1989	1069,329	1172,405	47170,1268	1218,702
2012	979,3036	862,8327	149287,5	926,456
1992	651,0968	755,2318	1463,863	570,7649
2004	669,5671	743,6531	981,2035	499,1216
1997	627,4419	913,3919	832,27	785,267
2003	630,9336	549,8713	770,5905	451,886
2005	850, 9224	960,6277	3072,856	827,7479
2008	1330,136	1530,081	10971,02	1541,935
1990	1034,872	1088,01	1194,50445	1092,04620
1993	1214,089	1493,355	1817,121	1480,389
1998	1566,408	1789,002	16653,6	1875,633
1988	666,8998	914,7831	23545,66	802,8721
1999	1284,457	1347,648	22098,16	1387,087
2002	831,5553	984,9278	9764,791	987,5316
2007	1461,383	1702,445	3712,283	1700,994
1991	795,1433	869,1988	3576,621	837,0168

Los valores de error mayores no indicaron tendencias centrales y en general los errores extremos no mostraron ser pequeños, es importante definir que el rango de altitud para los valores de error menores fue similar al de los valores mayores (rango valores menores entre 1420 metros y 2080 metros, rango valores mayores entre los 1165 y 2088 metros).

7. DISCUSIÓN DE RESULTADOS

En el desarrollo del análisis estructural, se busca la presencia de anisotropía, isotropía o tendencia, que se asocia a la variabilidad y comportamiento de los datos en el espacio. La presencia de anisotropía, refleja que la correlación entre los datos depende de la dirección en la que esta se calcule, recomendándose el uso de transformación, para disminuir el efecto. El proceso de análisis de datos permite tener una visión detallada de la calidad y normalidad de la información, reflejando precisión en la interpolación y en la determinación del método indicado de acuerdo a la información disponible. De acuerdo con Rodríguez, 2014, la escala temporal, la naturaleza de la variable así como la densidad de los datos, y factores característicos de la zona de estudio (tamaño, forma, orientación y orografía) determinan la bondad de ajuste de los métodos de interpolación que se están utilizando. Determinar la tendencia de los datos, mediante el uso de un gráfico en tres dimensiones, restringe y determina el uso de métodos geoestadísticos como el Kriging. Esta selección previa de métodos induce a elegir aquellos que se acercan al ideal de predicción y que pueden describir variables como la precipitación. Es posible que la inclusión de variables secundarias de tipo topográfica y geográfica, permitan analizar la dependencia de las mismas. Aunque diferentes estudios han determinado que la altitud no está directamente relacionada con la precipitación, se hace interesante asociarla al análisis estructural, y concluir si puede existir correlación en regiones andinas ecuatoriales. Según Hong et al, (2005), tener en cuenta la elevación como un gradiente vertical, es particularmente importante para la interpolación climática precisa, especialmente en áreas con pocos datos y montañosas. Así mismo, resultados como los de Hao et al, 2013, demuestran y fortalecen la hipótesis que la elevación como segunda variable, fortalece las predicciones. En la mayoría de los casos no se hace posible conocer la dinámica de las variables atmosféricas en condiciones de montaña y de igual manera la mayoría de trabajos con los que se referencian lo estudios en el país provienen de sitios totalmente disimiles (Valencia et al, 2015).

Aspectos importantes como el número de estaciones a analizar compromete y limita el uso de herramientas como Geostatistical Analyst. Un número menor o igual a 19 estaciones impide la aplicación de transformaciones (familia Box- Cox, Log) cuando hay ausencia de normalidad en los datos. La red de 21 estaciones influyo directamente en la estimación de los valores, confirmando lo expuesto por autores como Feng-Wen Chen y Chen-Wuing Liu, 2012, quienes concluyeron que un número conocido de estaciones es un parámetro influyente, aumentando la precisión de la predicción. Así mismo determinaron que la exactitud de la interpolación de los datos de lluvia podría ser reducida por la interferencia del uso excesivo de estaciones pluviométricas. El resultado óptimo está basado en estaciones pluviométricas distribuidas dentro de un radio de influencia.

Valencia et, al (2015), asume que el radio de acción es cercano a cero cuando los semivariogramas presentan una discontinuidad puntual en el origen (efectos pepita), siendo indicativo de que parte de la estructura espacial se concentra a distancias inferiores a las observadas. Es un indicativo de carencia de correlación espacial entre las observaciones de una variable. Es claro que el radio de acción de una estación pluviométrica en la región andina es dependiente de la escala de tiempo analizada; es decir, el área cubierta es menor cuando se trata de datos diarios y es mayor cuando se están analizando datos anuales, de acuerdo al estudio Determinación del área de influencia de estaciones pluviométricas en la región Cafetera del departamento de Caldas, Colombia. El objetivo de la red pluviométrica definirá la distribución de las estaciones y el número de las mismas; generalizar y concluir que se deben usar siempre las escalas de tiempo mayores para dar cubrimiento al área no es adecuado.

La gran diferencia entre el área real cubierta y el área potencial cubierta a diferentes escalas de tiempo muestran que para resolver el problema de tiempo y optimizar las interpolaciones, más que instalar estaciones, se debería redistribuir la red pluviométrica actual, ya que hay información redundante en todas las escalas de tiempo, como lo menciona Valencia et al (2015). Menciona que incluso a escala anual, cuando se encuentran las mayores áreas de influencia de un pluviómetro en la región andina, esta área es inferior a 1000 km². Este argumento permite entender el suavizado presente en los mapas generados con el método IDW, en donde el área de interpolación o influencia depende de la proximidad de los vecinos.

De acuerdo a Chen (2012), la variable distancia puede generar un buen resultado al utilizar el método IDW, un valor de radio es evaluado para determinar los parámetros óptimos. Esto asegura la utilidad general del modelo IDW y permite una mejor predicción de datos que es comparable a los datos observados. Aunque la cantidad de valores menores de Error medio se presentaron con el método determinístico IDW, fueron gruesos, de lo que se estima, se deben tener en cuenta factores como el uso de variables secundarias, la calidad de la información, mejor distribución de estaciones o en su defecto aumento de las mismas. Este factor se vió reflejado en la capacidad de predicción de algunos métodos como el K. Ordinario, en el que no se presentaron valores de precipitación, al no contar con estaciones vecinas. La distancia entre las estaciones analizadas determina una relación directa con la exactitud de la predicción. Esta predicción asume implícitamente la continuidad de la variable.

Si no hay influencia direccional en los pesos de los datos, se consideran los puntos en todas las direcciones y la forma de la búsqueda de la vecindad un círculo (ejes mayor y menor de igual longitud). Si hay influencia direccional en los datos, se adopta la forma de una elipse con el eje mayor paralelo al fenómeno considerado. El círculo de vecindad se puede dividir en sectores, luego el número de vecinos será tenido en cuenta en cada sector, esto es, si se toman 5 vecinos y

se divide el círculo en 4 sectores, para la predicción se tendrán en cuenta 20 vecinos que se elegirán fuera del círculo si no hay los suficientes con los que están dentro del círculo (Boada, 2015).

La calidad del modelo predictivo evaluado con la validación cruzada como estimación, se hace sólida cuando el número de estaciones amplía la posibilidad de usarla K- veces. Chen (2012), en su estudio, seleccionó el 25% de las 45 estaciones utilizadas, facilitando el análisis y la comparación de diferentes estimaciones en la escala de tiempo. En el presente estudio se retiraron dos de las 21 estaciones utilizadas, con las que se reflejó restricciones con la herramienta Geostatistical Analyst en los casos en los que se debía aplicar transformaciones. Es necesario evaluar si en aquellos lugares en los que la red climática es limitada, los estudios exigen un aumento de las estaciones o una redistribución, y si esta herramienta de análisis geoestadístico es útil.

Según Chen, 2012, la aplicación de IDW para interpolación espacial de datos de lluvia, muestra como resultado mayor precisión de la predicción durante las estaciones secas (de octubre a abril) que en las temporadas de lluvia. Así mismo revelan que IDW tiene significativa capacidad de predicción en eventos de lluvia pequeñas que en eventos extremos. En resumen, concluyen que IDW es considerado como un método de interpolación espacial de lluvia adecuado cuando es tenido en cuenta el radio de influencia. Sin embargo autores como Taesombat y Sriwongsitanon, 2009, resaltan la importancia del uso de otros métodos convencionales como el Thin Plate Splines (TPS), que también puede ser usado para interpolar datos de precipitación de manera más exacta en especial en áreas montañosas. Esta técnica puede generar superficies de interpolación utilizando tres variables, latitud, longitud y altitud de las diferentes estaciones meteorológicas.

Los métodos geoestadísticos como el Kriging utilizan el modelo ajustado del semivariograma para realizar la predicción de un valor no muestral sobre un lugar específico. Kriging asume que el fenómeno que se está estudiando toma valores en toda el área de estudio, es decir tiene continuidad espacial. Permite crear superficies de Predicción, Cuantiles, Probabilidades y Error estándar de predicción, a diferencia de los determinísticos que generan los de Predicción. Una de sus limitantes durante el presente estudio, se identificó en el análisis estructural, en donde la presencia de tendencia en los datos se limita el uso de algunos de estos métodos como el Kriging Ordinario, Simple e Indicador. Si en los datos muestrales no hay influencias direccionales (isotropía), los puntos utilizados para predecir se elegirán en todas direcciones desde el lugar no muestral, por lo tanto la forma que se tiene será un círculo. Pero si los datos tienen autocorrelación direccional (anisotropía) entonces se elegirá una elipse como forma de búsqueda de vecindad.

Autores como Díaz y González, 2009, consideran el uso de métodos geoestadísticos como las simulaciones, ya que ofrecen ventajas al no depender del número de puntos (estaciones) y de su distribución espacial. La simulación condicional puede ser perfeccionada agregándole toda la información cualitativa disponible del fenómeno real. Sin embargo son más demandantes computacionalmente. De igual manera Osorio, 2015, determina que las Simulaciones son capaces de estimar el valor en zonas no muestreadas, realizar análisis de sensibilidad respecto a diferentes escenarios, generar interpolaciones exactas cuando la simulación es condicional, una función de la variable regionalizada puede ser estimada sin sesgo por la misma función aplicada a los valores simulados, la dispersión de los valores simulados es la misma que la de los valores verdaderos, de esta manera no suaviza.

Lam (1983) citado por Hong et al (2005), concluye que hay muchos óptimos en lugar de un único y "mejor" opción, ya que el rendimiento de técnicas de interpolación depende fundamentalmente del tipo de datos, la precisión deseada y la eficiencia de la computación. Sin embargo, el conocimiento de un método que se ajuste a la cantidad y calidad de la información climática disponible se hace importante.

8. CONCLUSIONES

- La elección del conjunto de estaciones debe estar condicionada a factores como ubicación (diferenciadas para cada una de las zonas geográficas), y a estaciones que presenten registros confiables con mayor número de años o mensualmente más completos (de al menos 10 años).
- El número de estaciones resulta ser un factor condicionante para el uso de la herramienta Geoestatistical Analyst, impidiendo la transformación de los datos.
- Aspectos como el número de estaciones, años de muestreo, análisis de normalidad en los datos, permite la disminución del error cuadrático medio.
- La presencia de tendencia en los datos, genera condicionantes para el uso de métodos geoestadísticos como el Kriging.
- La mayoría de los métodos geoestadísticos sólo son óptimos si la variable de estudio (precipitación para este caso) sigue una distribución normal.
- El método determinístico IDW mostró, en general, los valores menores del error cuadrático medio. Sin embargo dichos valores fueron altos, valores entre los 627,4419 (año 1997) y 1566,408 (año 1998) respaldando el enunciado anterior (párrafo 3).
- El uso de la variable distancia asociado al método IDW, puede disminuir los valores de Error cuadrático medio.
- La capacidad de predicción, fue limitada para métodos como el K. Ordinario cuando no se contaba con estaciones vecinas.

9. RECOMENDACIONES

- Se hace importante aumentar el número de estaciones meteorológicas en el departamento, que en tiempo y regularidad de toma de datos permitan ser incluidas en este tipo de estudio.
- Vincular variables secundarias topográficas y geográficas puede tener relación directa con la variable de estudio y explicar mejor su comportamiento.
- Clasificar los datos de acuerdo a las épocas (seca y lluviosa) es utilizado para puntualizar los resultados de interpolación de variables climáticas como la lluvia.
- El uso de categorías o grupos de estaciones, de acuerdo a variables secundarias como distancia, altitud y épocas secas y de lluvia, pueden proveer resultados detallados y descriptivos de los datos interpolados, es posible que mejore la precisión de la predicción.
- Se recomienda utilizar métodos como Test Chi-cuadrado y Kolmogorov-Smirnov para el análisis exploratorio de los datos, ya que la herramienta Geostatistical analyst limita el análisis de los datos cuando se tienen menos de 20 estaciones, como en este caso.
- El uso de otros métodos geoestadísticos como las simulaciones, eventualmente podrían dar mejores resultados en la interpolación de la variable analizada en el presente estudio.
- Además de la validación cruzada es posible realizar la inclusión de un nuevo conjunto de datos test, para verificar los métodos de interpolación utilizados.
- El análisis exploratorio de datos como herramienta de observación puede ser complementado con el uso de semivariogramas experimentales que en conjunto facilitarían la posterior interpolación de variables en la región Andina Ecuatorial.

BIBLIOGRAFÍA

Chen, D. OU, T. Gong, L. XU, C. LI, W. HO, C. 2010. Weihong, Q. Spatial Interpolation of Daily Precipitation in China: 1951- 2005. *Advances in atmospheric sciences*, Vol. 27, No. 6, 1221-1232.

Chica, R.H, A; Peña, Q. A, J; Giraldo, J. J. F; Obando, B. D; Riaño, H. N, M. 2014. SueMulador: Herramienta para la Simulación de Datos Faltantes en Series Climáticas Diarias de Zonas Ecuatoriales *Revista Facultad Nacional de Agronomía - Medellín*, vol. 67, núm. 2, pp. 7365-7373. Universidad Nacional de Colombia, Medellín, Colombia.

Díaz, V. M.A; González, C. R. 2009. *Geoestadística Aplicada. Tema: Métodos de Simulación*. Universidad Nacional Autónoma de México.

Feng, W. C; Chen, W. L, 2012. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. *Paddy Water Environ* DOI 10.1007/s10333-012-0319-1.

Giraldo, H. R. 2000. *Introducción a la geoestadística: Teoría y aplicación*. Universidad Nacional de Colombia, sede Bogotá.

Hao, W. Chang, X. 2013. Comparison of Spatial Interpolation Methods for Precipitation in Ningxia, China. *International Journal of Science and Research (IJSR)*, India Online ISSN: 2319-7064. Volume 2 Issue 8.

Hong, Y. Nix, H, A. Hutchinson, M. Booth, T. H. 2005. Spatial Interpolation of monthly mean climate data for China. *International Journal of Climatology Int. J. Climatol.* 25: 1369–1379. Published online in Wiley InterScience. DOI: 10.1002/joc.1187

Boada, R. A. 2015 Documento Métodos Determinísticos. Módulo Estadística Espacial, Especialización en SIG. Universidad de Manizales, Manizales, Caldas.

Boada, R. A. 2015. Documento Métodos no Determinísticos, Kriging. Módulo Estadística Espacial, Especialización en SIG. Universidad de Manizales, Manizales, Caldas.

Boada, R. A. 2015 Documento Análisis Exploratorio de Datos. Módulo Estadística Espacial, Especialización en SIG. Universidad de Manizales, Manizales, Caldas.

Jaramillo R., A. 2005. Lluvias máximas en 24 horas para la región andina de Colombia. *Cenicafé* 56(4): 250-268.

Osorio, A. J. 2015. Comparación entre los métodos de estimación de recursos mineros empleando métodos tradicionales y geoestadísticos en la planeación minera de placeres. Monografía para optar por el título magister en ingeniería – recursos minerales. Universidad Nacional de Colombia, Facultad de Minas Medellín.

Pizarro, R., Ramírez, C., Flores, J. P. 2003. Análisis comparativo de cinco métodos para la estimación de precipitaciones areales anuales en períodos extremos. Comparison analyses of five mean areal rainfall estimation methods in widely varying years. BOSQUE 24(3): 31-38. NOTA TECNICA

Portalés, C. Boronat-Zarceño, N. Pardo-Pascual, J.,E. Balaguer-Beser, A. 2008. Un nuevo método para el cálculo de precipitaciones medias mediante técnicas de interpolación Geoestadística considerando las características Geográficas y topográficas del territorio. Congreso Internacional de Ingeniería Geomántica y topográfica. Valencia.

Ranses, J. Vázquez-Montenegro Oscar J. Solano-Ojeda. 2012. Modelación espacial de la lluvia y la evapotranspiración teniendo en cuenta parámetros geográficos. Instituto de Meteorología Loma de Casablanca, Regla. La Habana. Cuba. Diciembre 30.

Rodríguez, S. J. M. 2014. Aplicación de métodos de interpolación para el cálculo de precipitación por modelamiento geoestadístico y análisis espacial para el departamento de Cundinamarca. Universidad Militar Nueva Granada, Geomática.

Eslava R., J. A. 1993. Climatología y diversidad climática de Colombia. Revista de la Academia Colombiana de Ciencias Exactas y Naturales 18(71):507-538.

Taesombat, W. Sriwongsitanon, N. 2009. Areal rainfall estimation using spatial Interpolation techniques. Department of Water Resources Engineering, Faculty of Engineering, Kasetsart University, Bangkok 10900, Thailand. ScienceAsia 35 (2009): 268–275

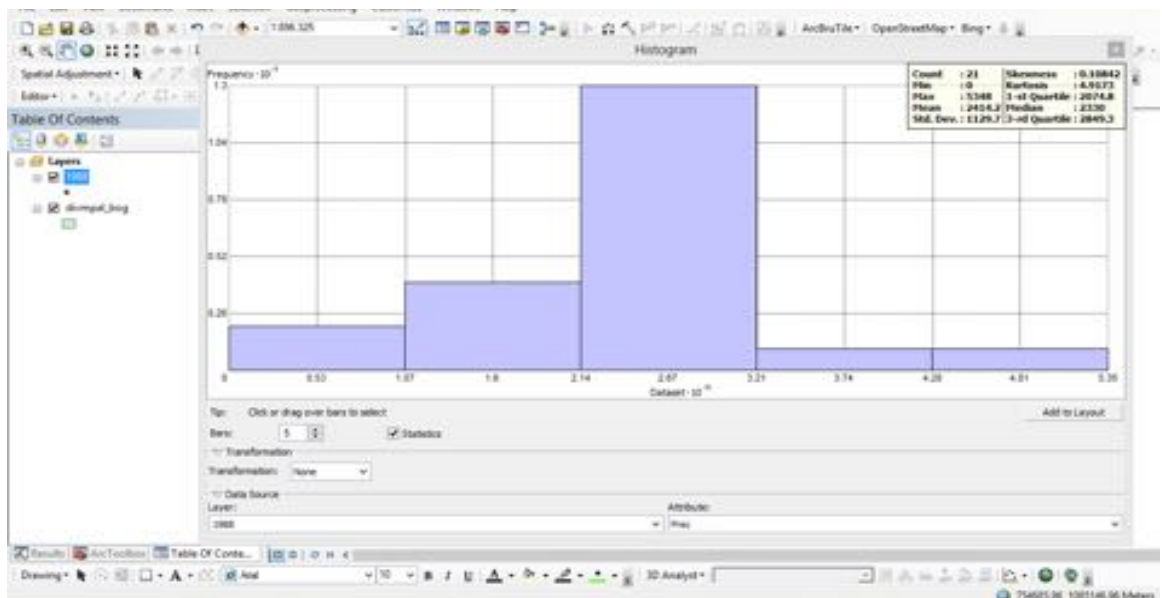
Valencia, J. A. Castillo, O. G. Betancur, J. F. Peña, J.A. 2015. Determinación del área de influencia de estaciones pluviométricas en la región cafetera del departamento de Caldas, Colombia. Global Advanced Research Journal of Agricultural Science (ISSN: 2315-5094) Vol. 4(2) pp. 118-126, Available online <http://garj.org/garjas/index.htm>. Copyright © 2015 Global Advanced Research Journal.

Viera, D. M. 2002. Geoestadística aplicada. Instituto de Geofísica, UNAM Instituto de Geofísica y Astronomía, CITMA, Cuba.

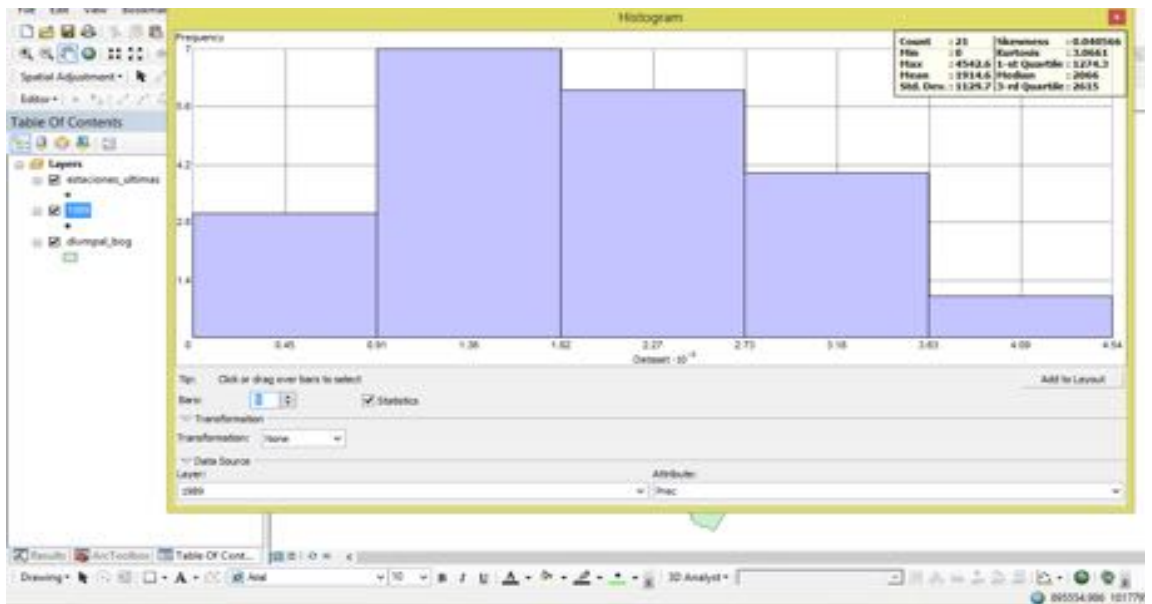
ANEXOS

Anexo A, Análisis exploratorio de datos: Histogramas anuales

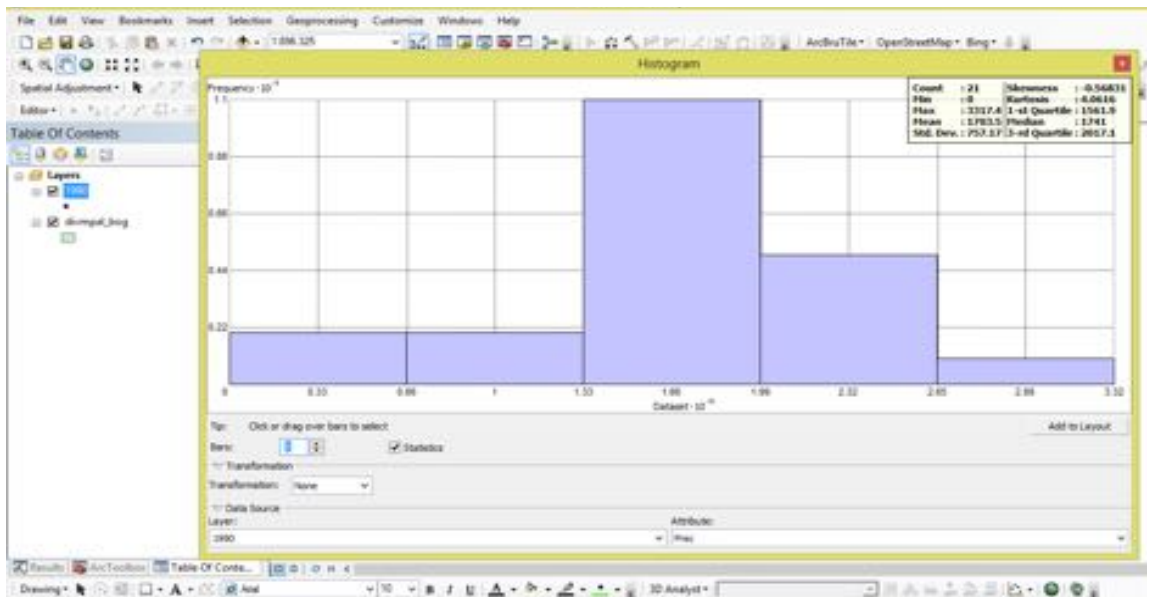
Año 1988:



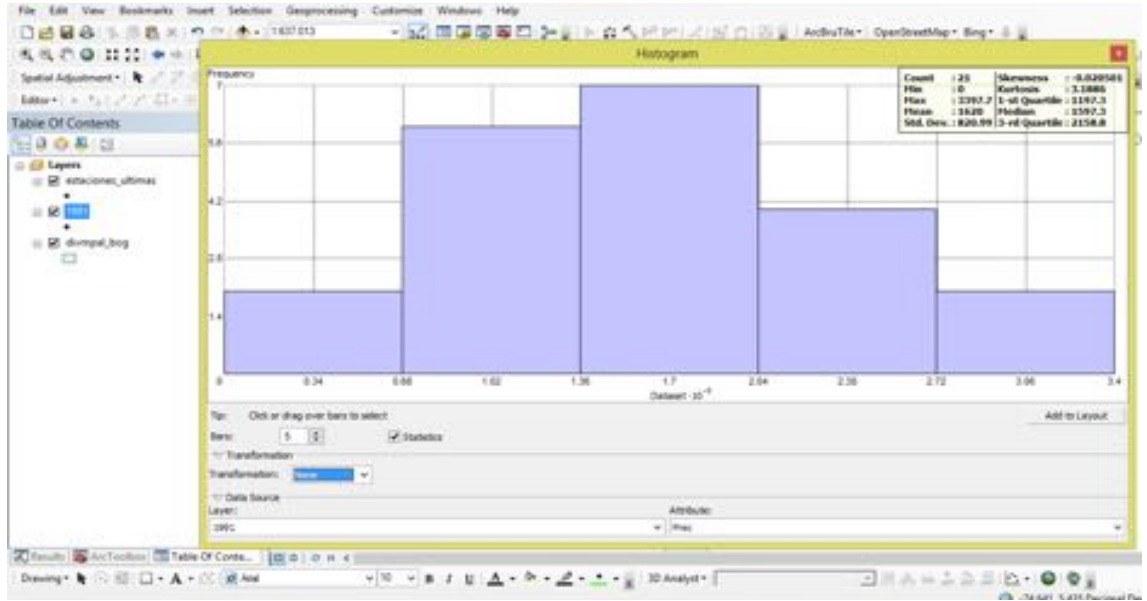
Año 1989:



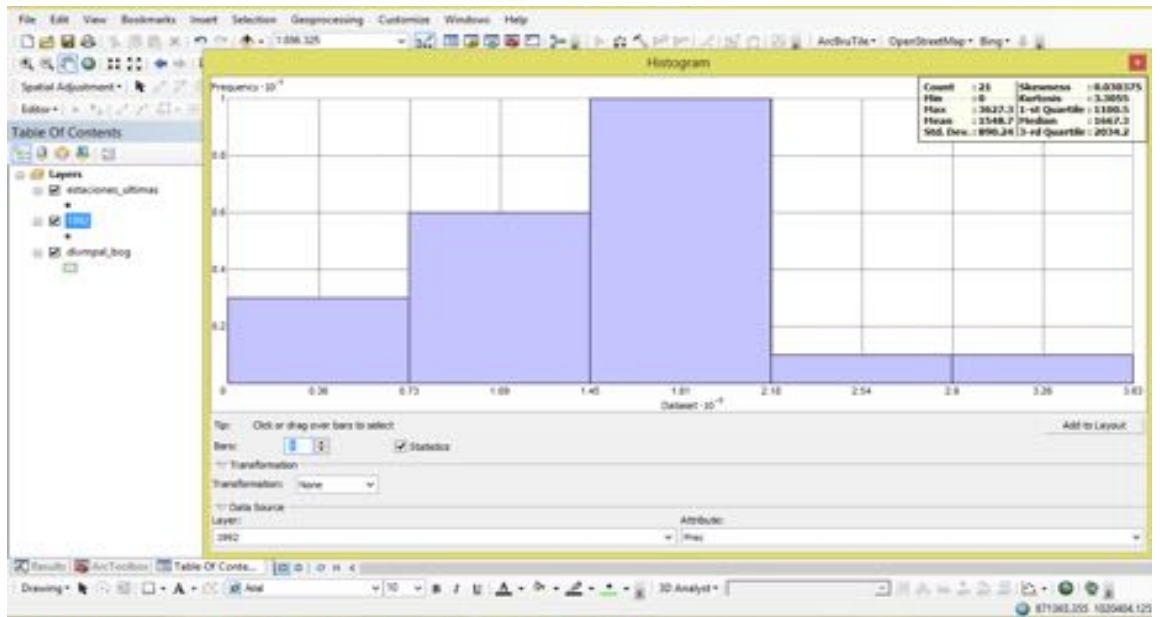
Año 1990:



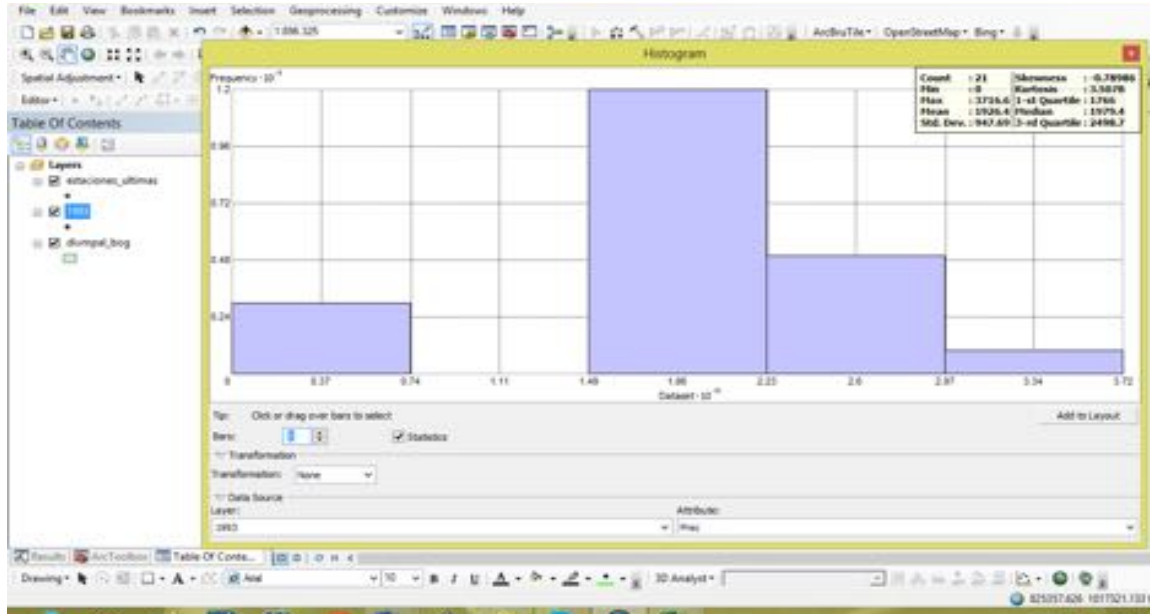
Año 1991:



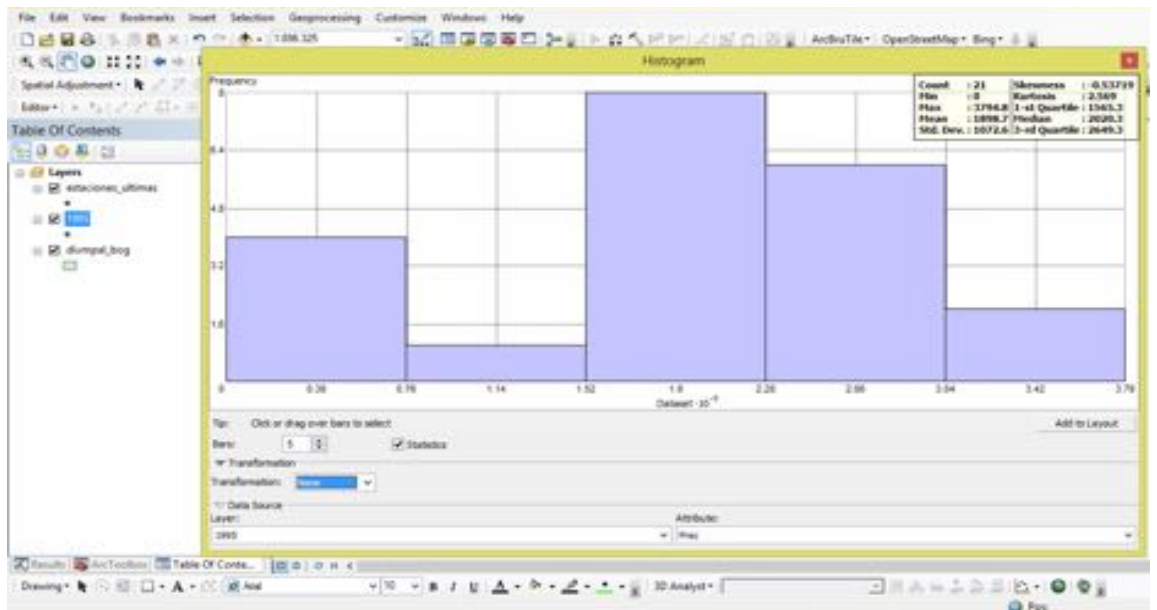
Año 1992:



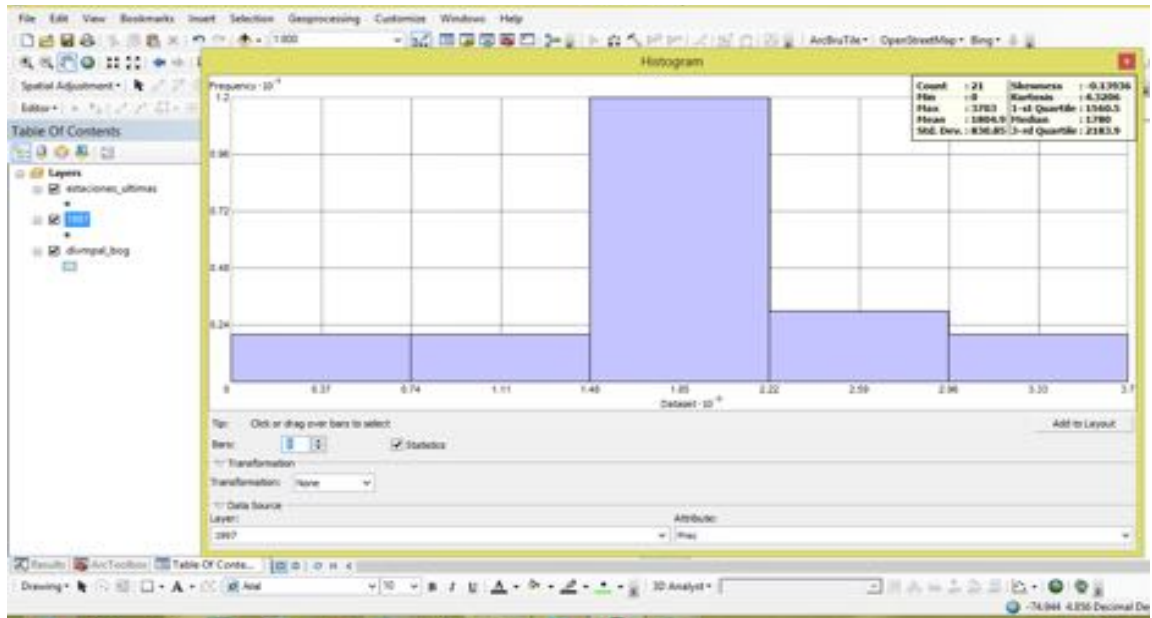
Año 1993:



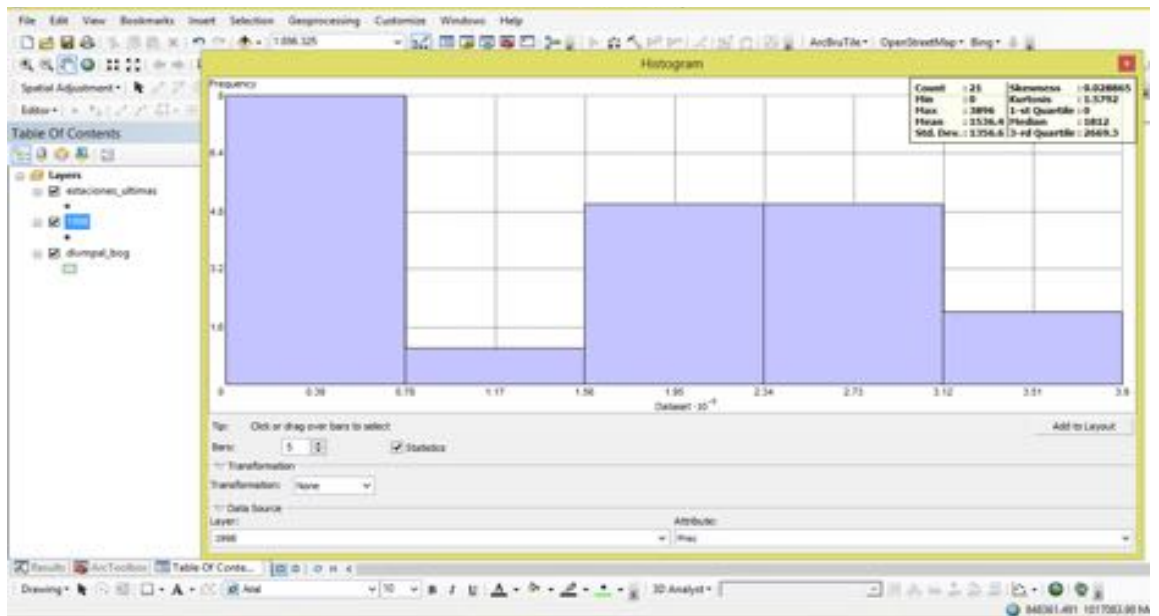
Año 1995:



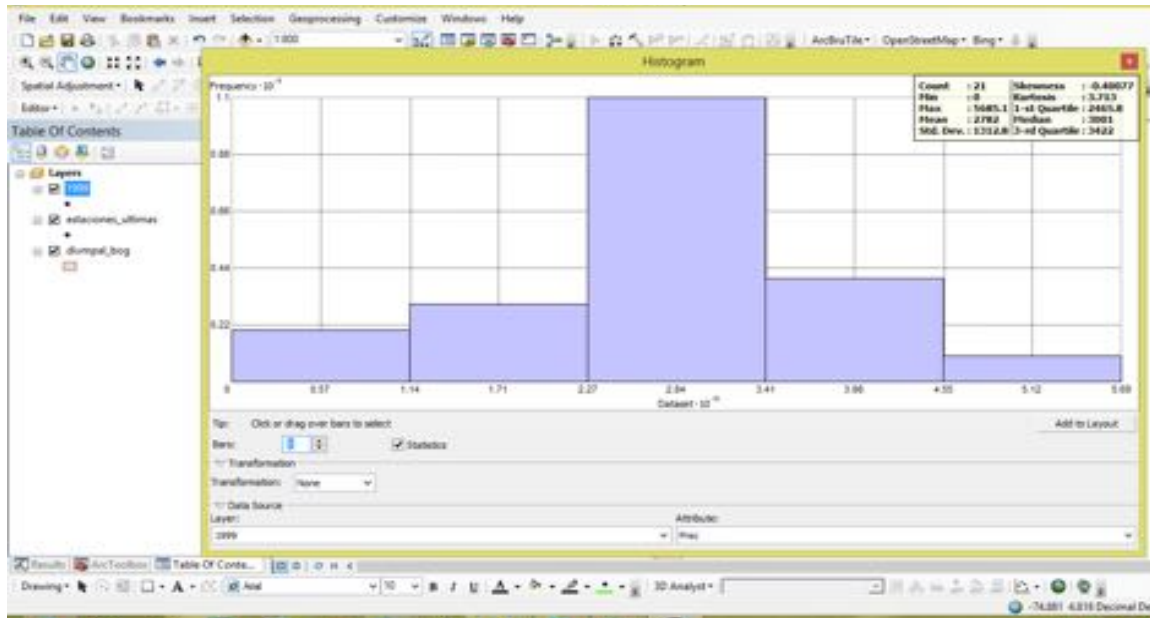
Año 1997:



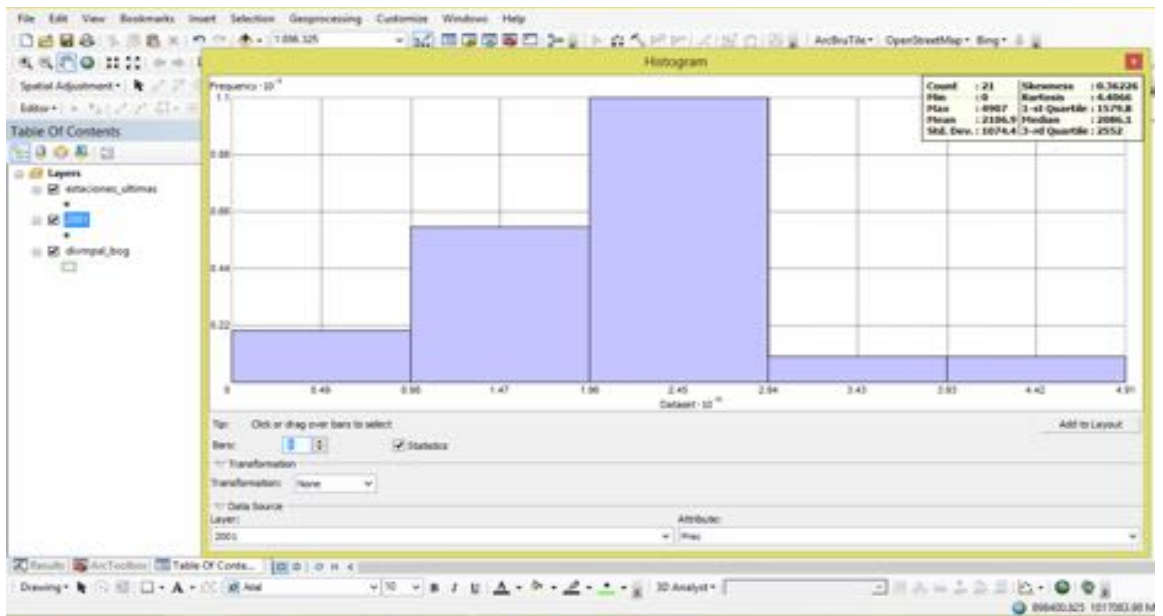
Año 1998:



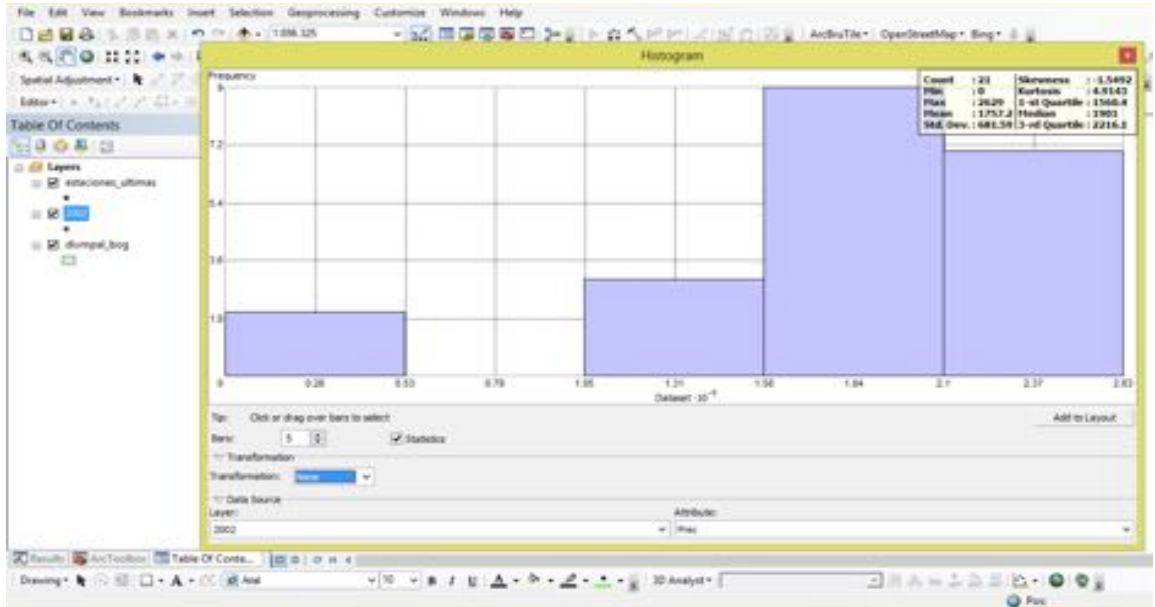
Año 1999:



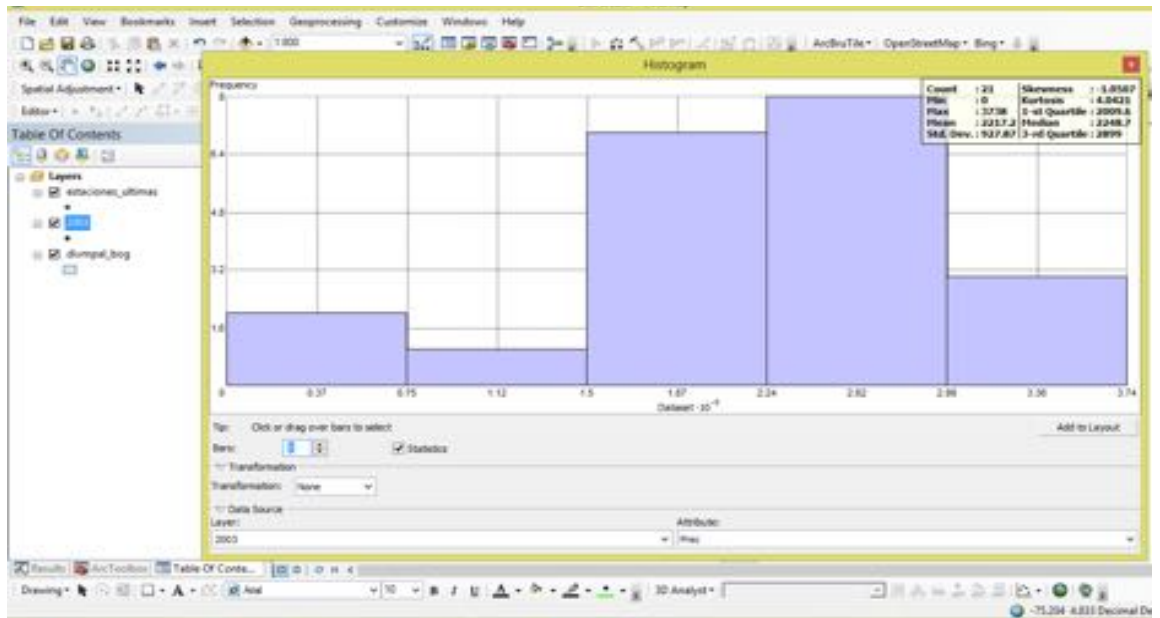
Año 2001:



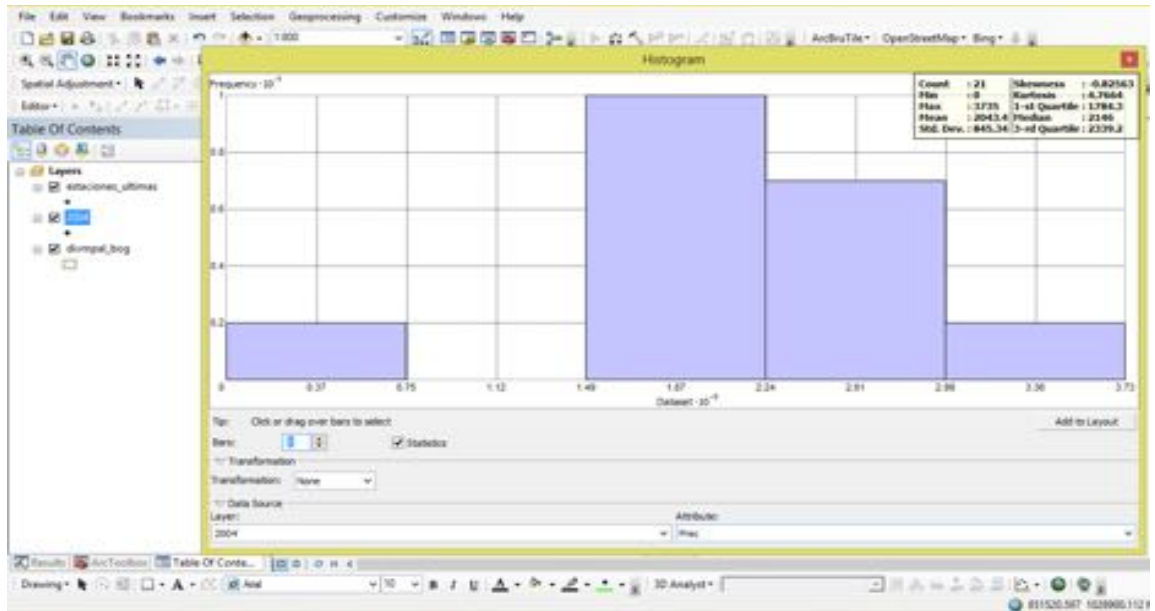
Año 2002:



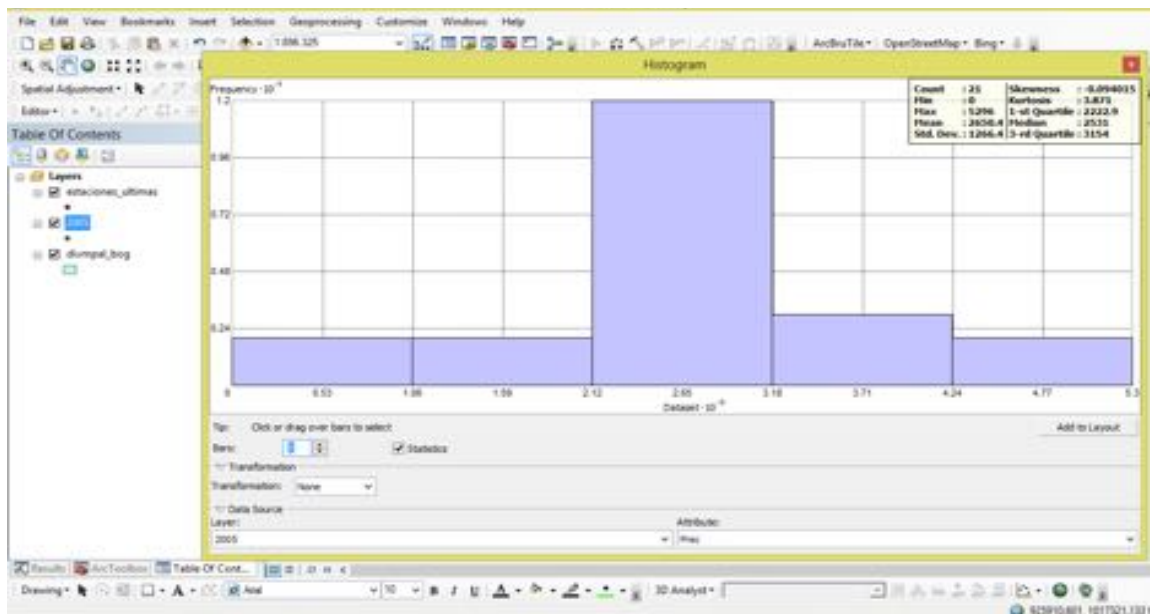
Año 2003:



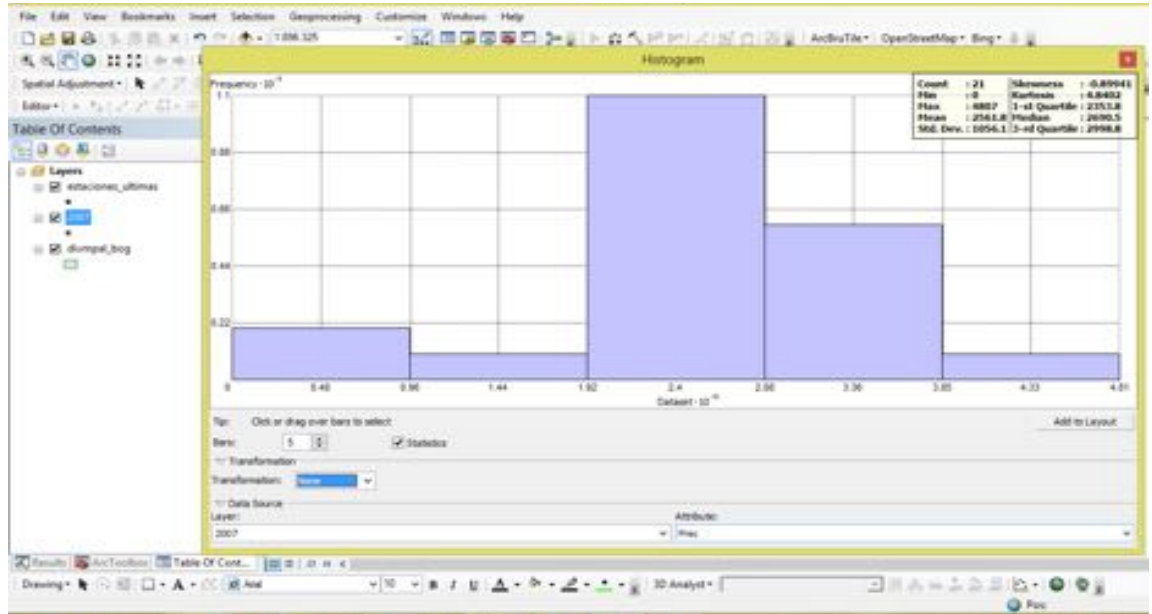
Año 2004:



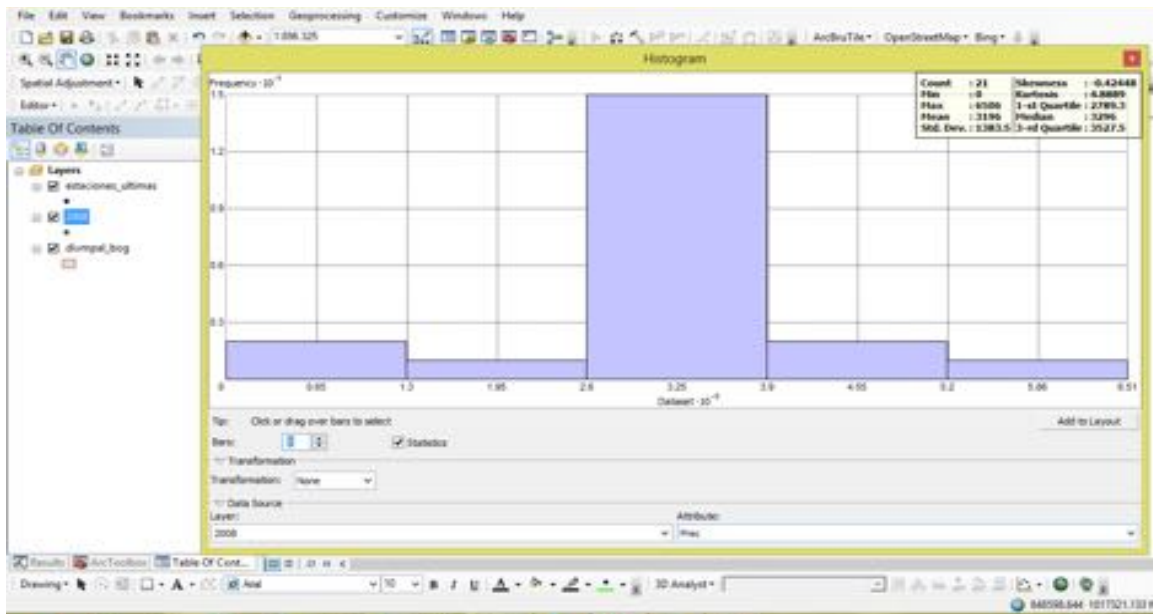
Año 2005:



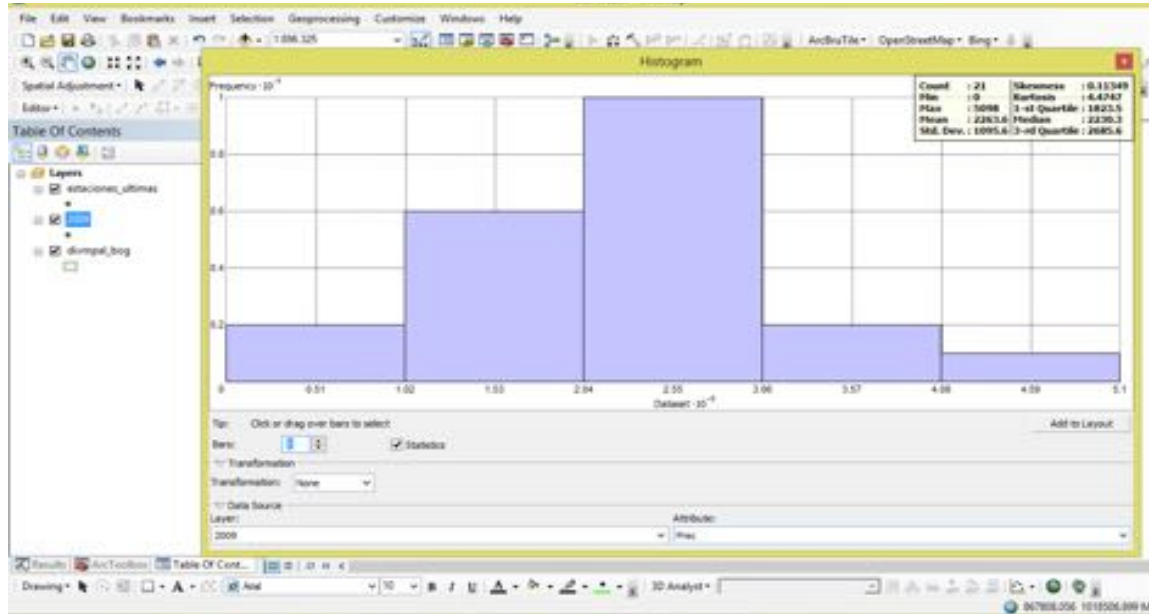
Año 2007:



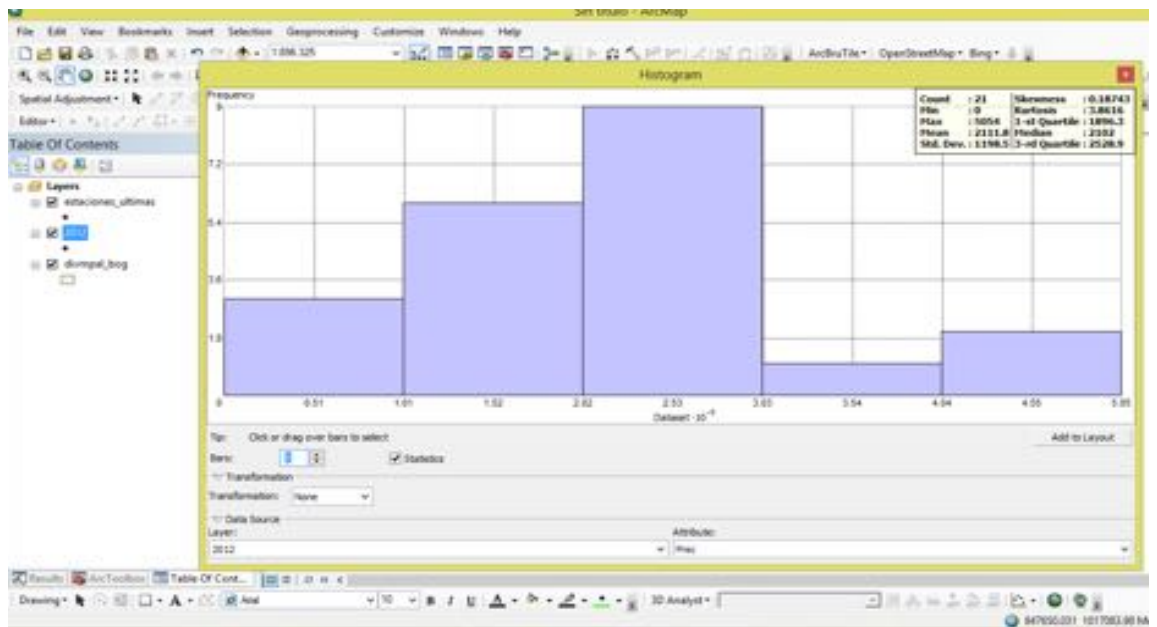
Año 2008:



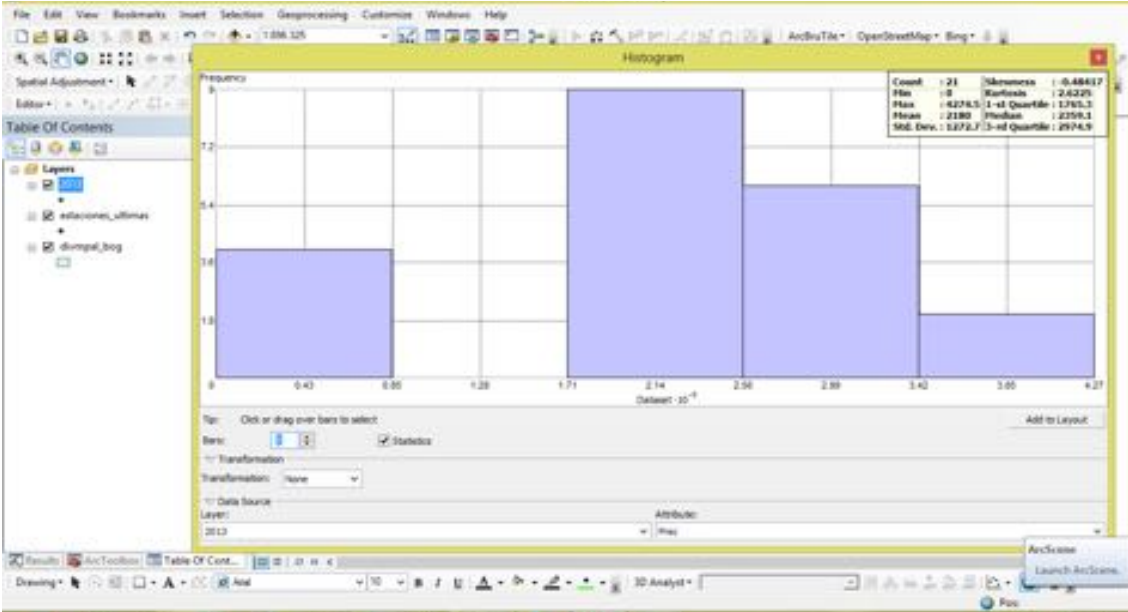
Año 2009:



Año 2012:



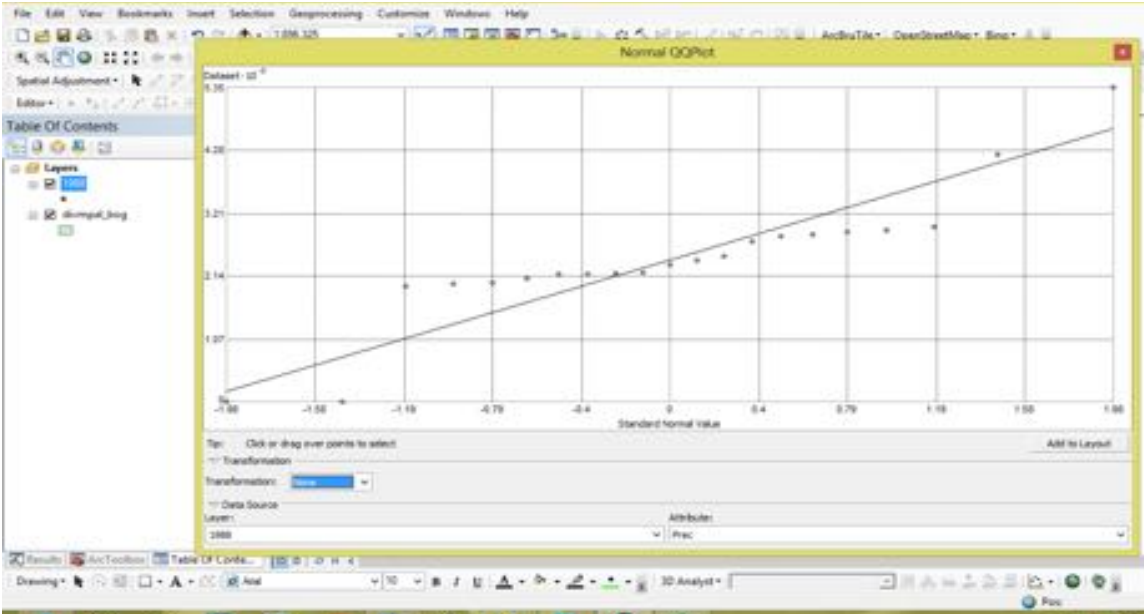
Año 2013:



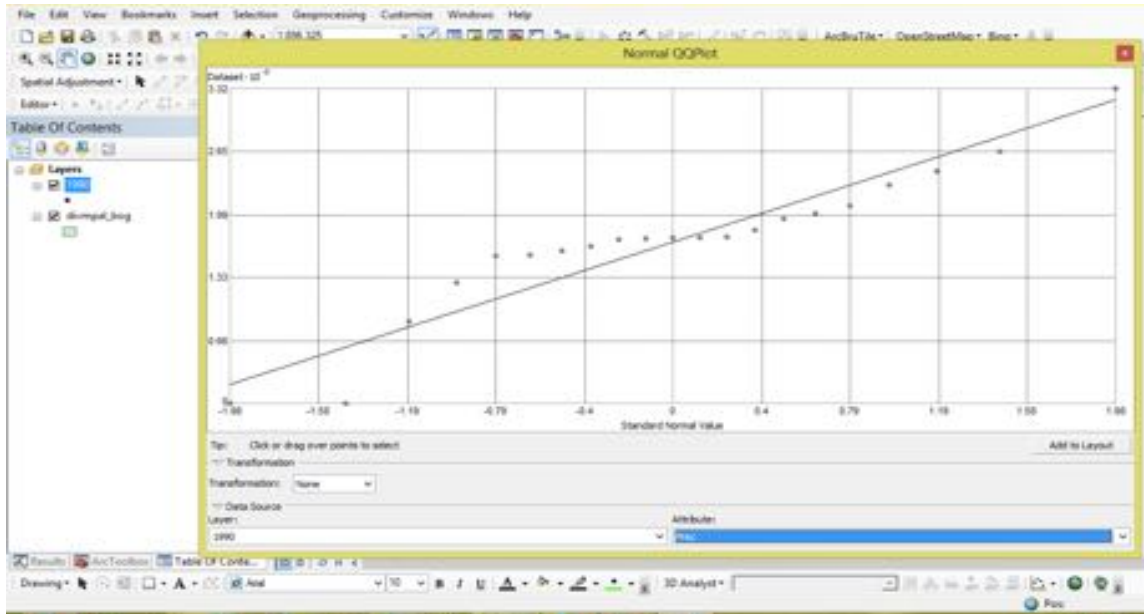
Anexo B

Gráficos anuales Normal QQ Plot

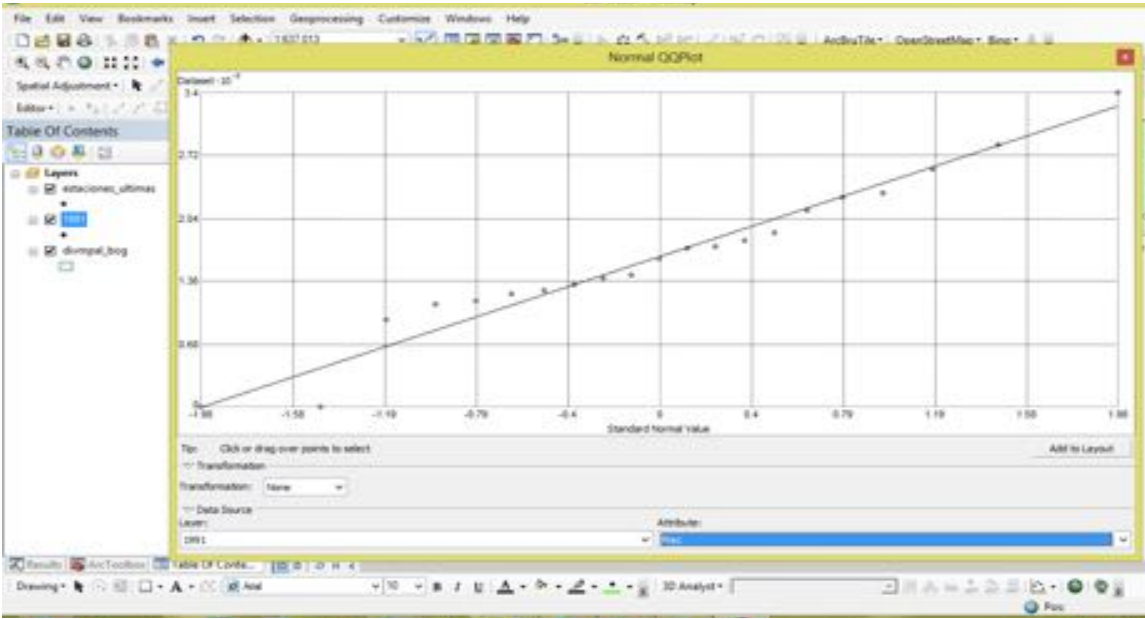
Año 1988:



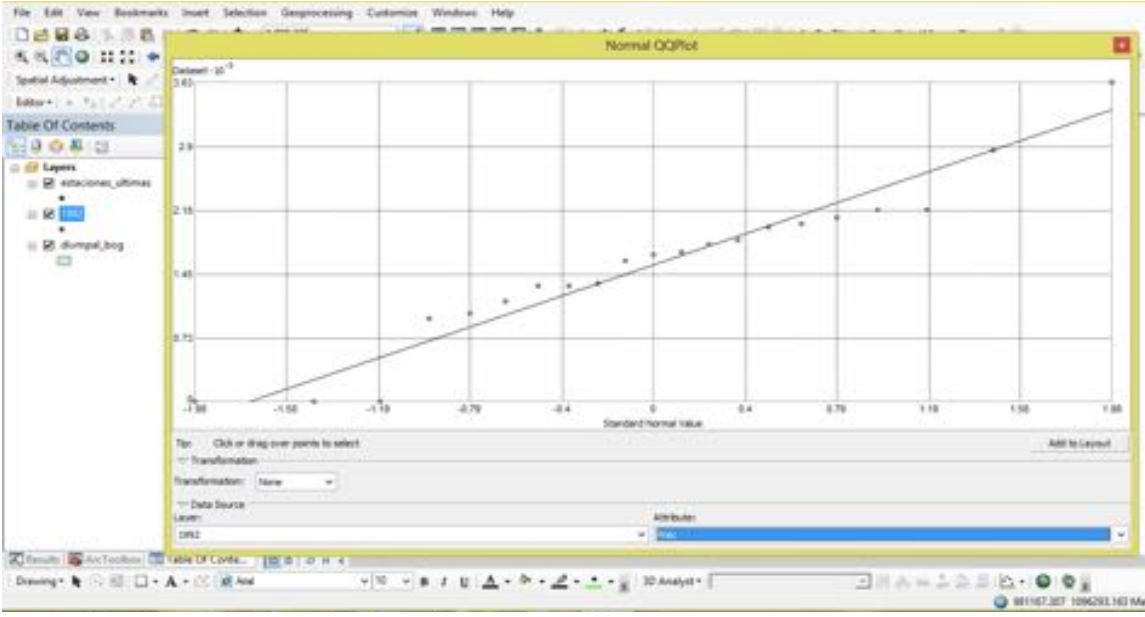
Año 1990:



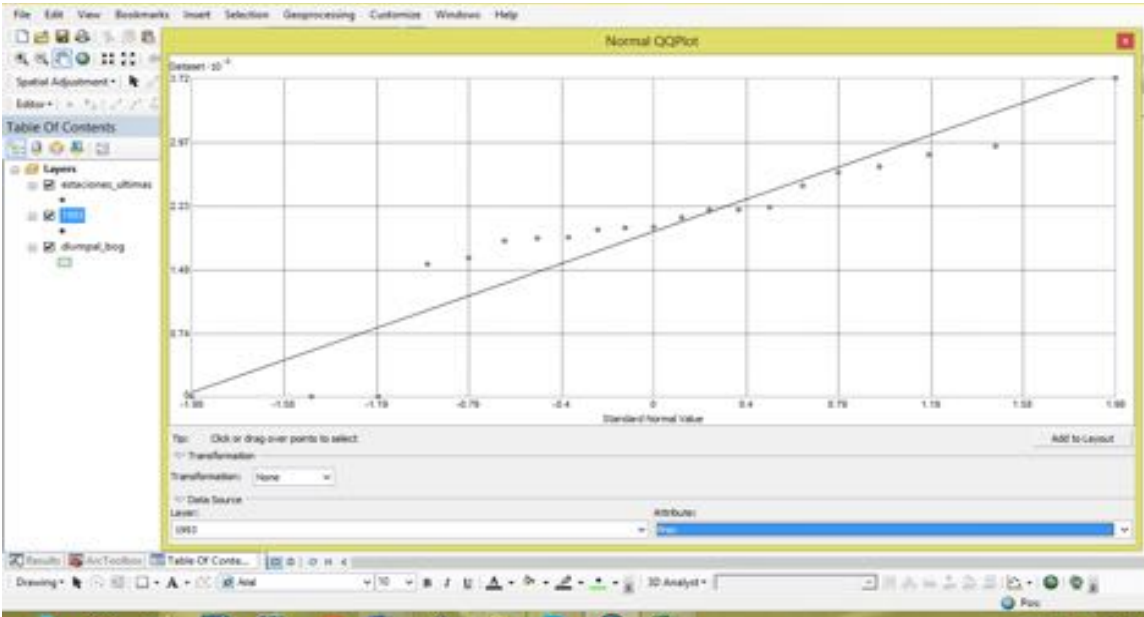
Año 1991:



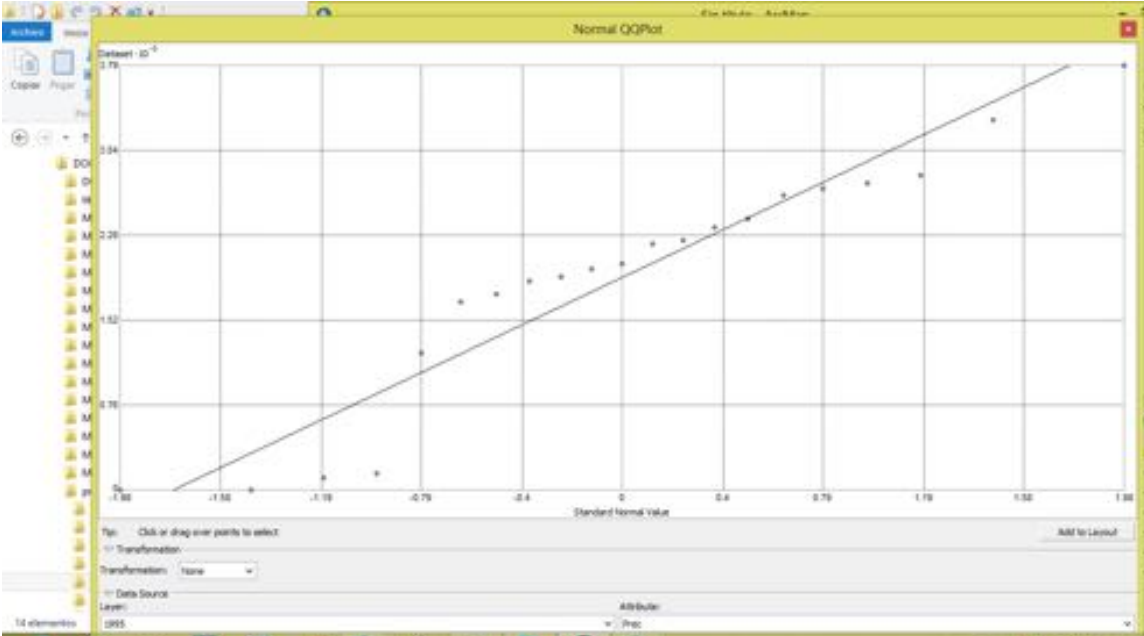
Año 1992:



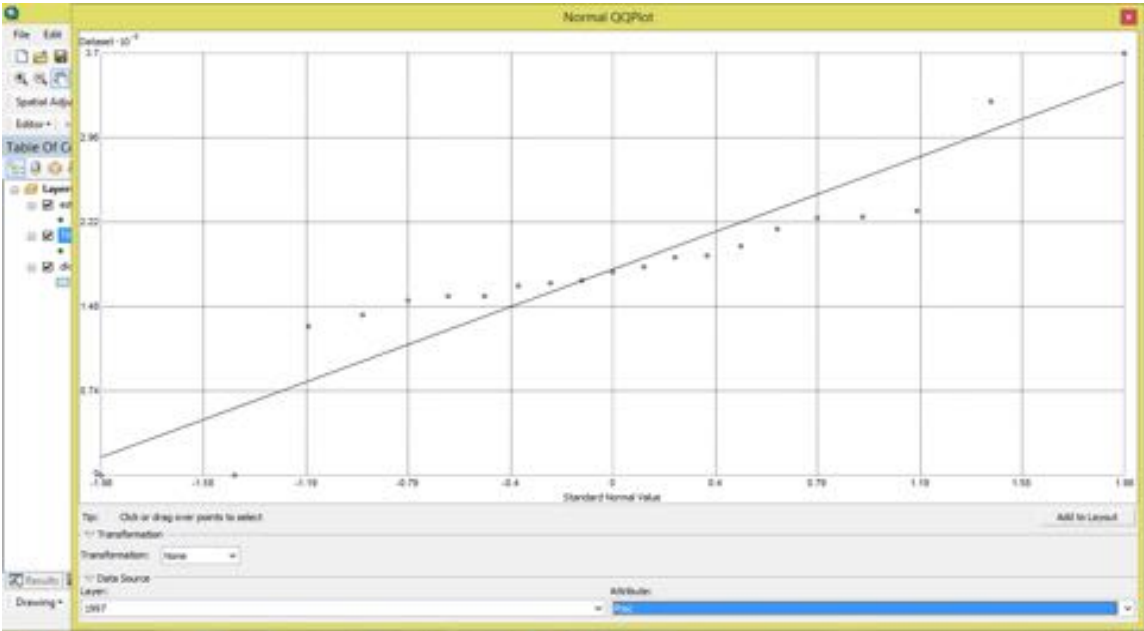
Año 1993:



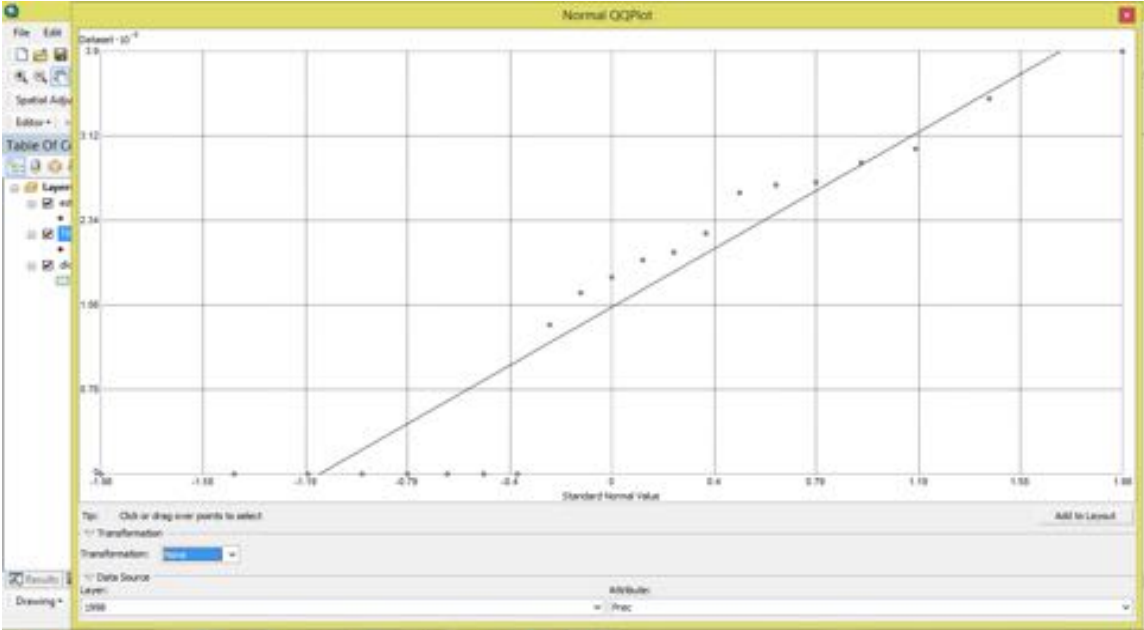
Año 1995:



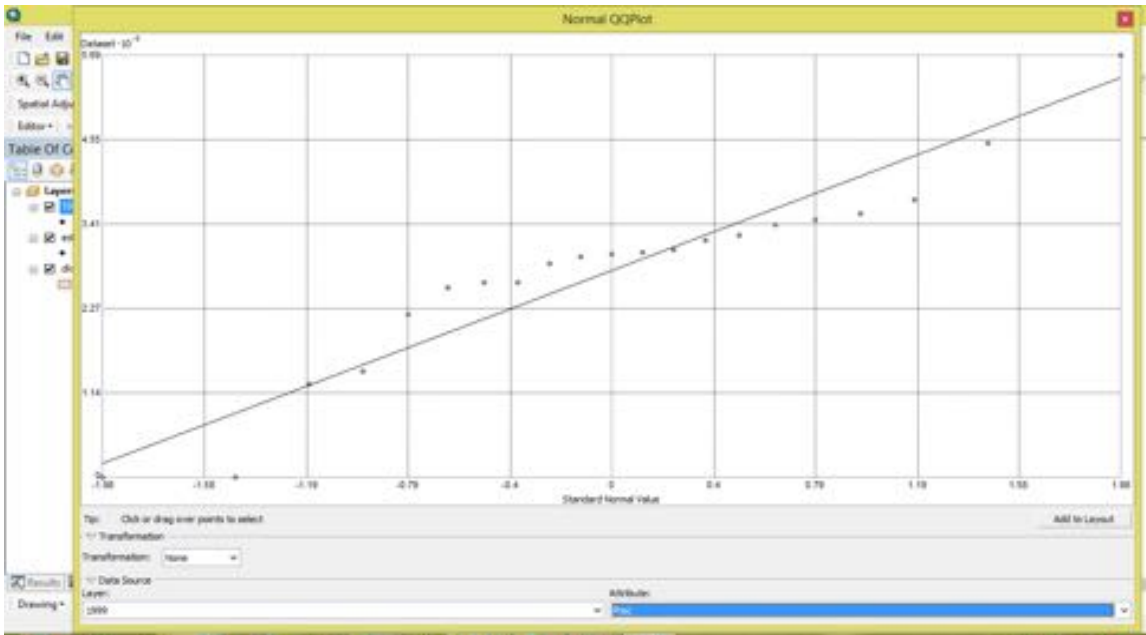
Año 1997:



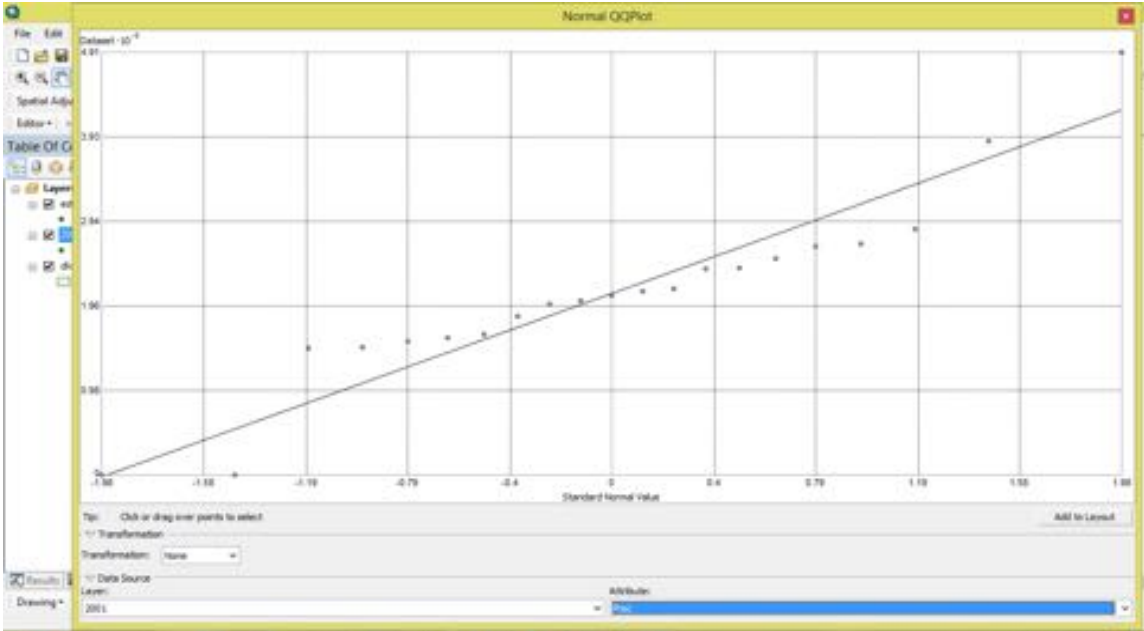
Año 1998:



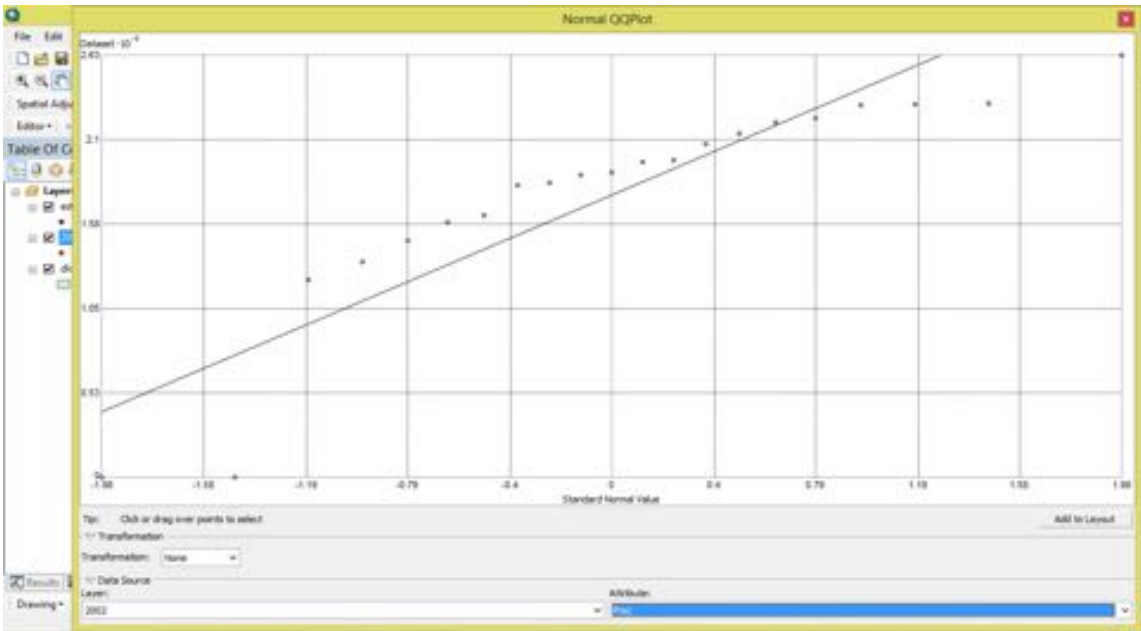
Año 1999:



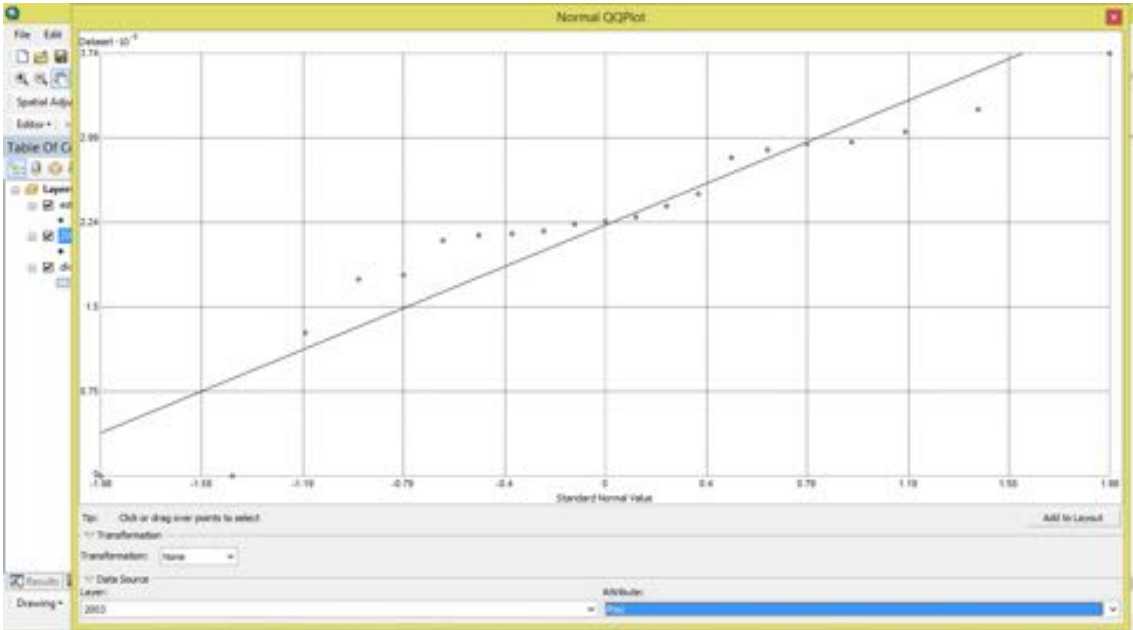
Año 2001:



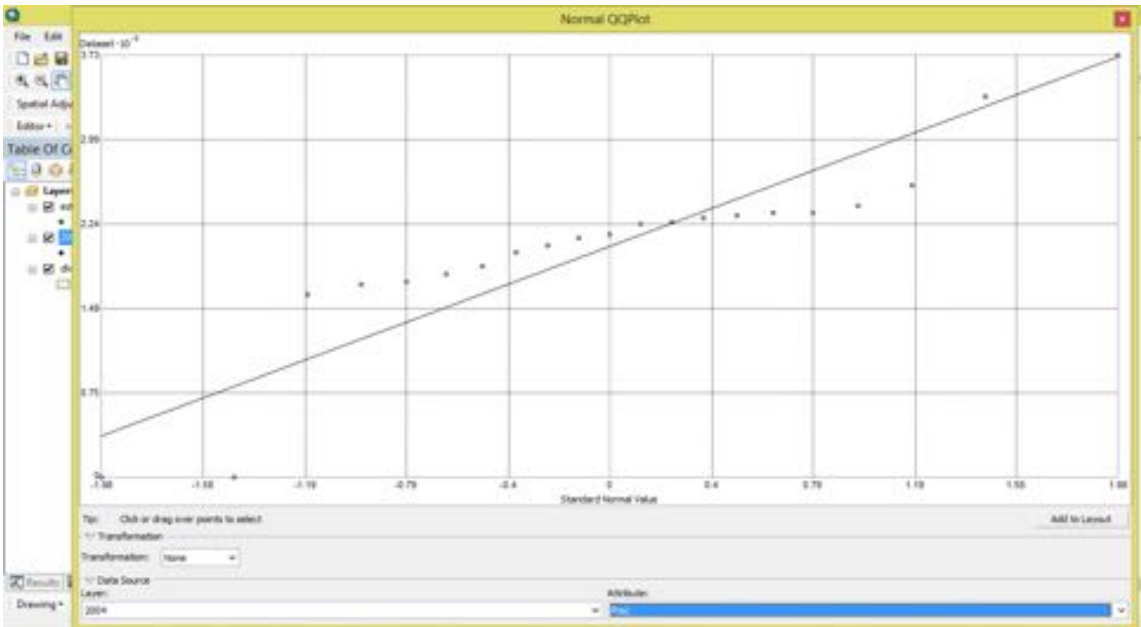
Año 2002:



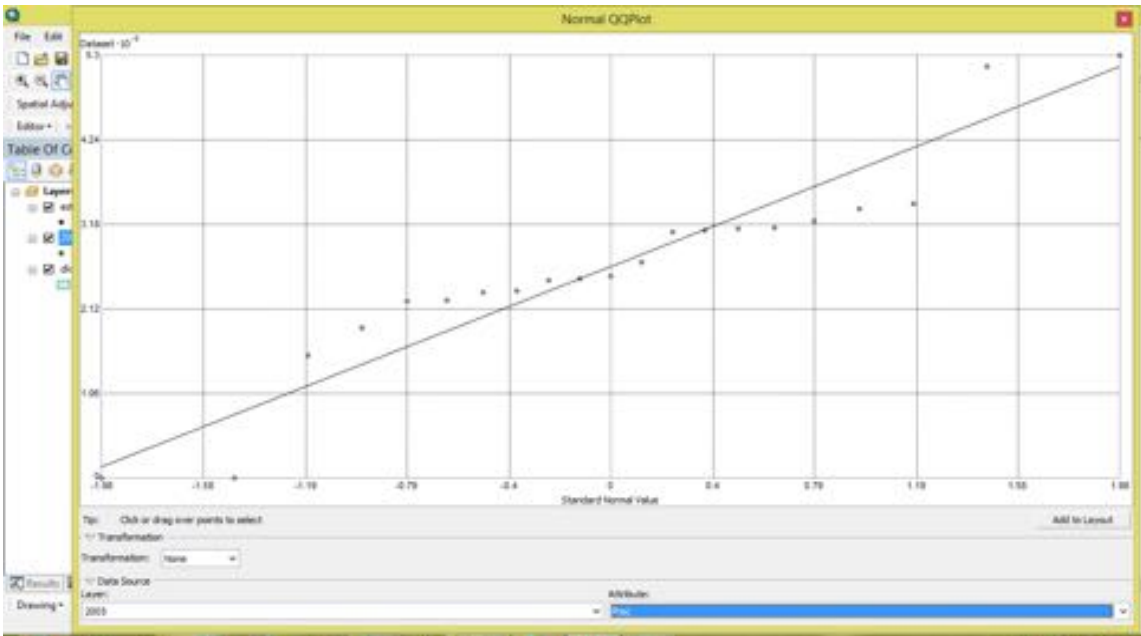
Año 2003:



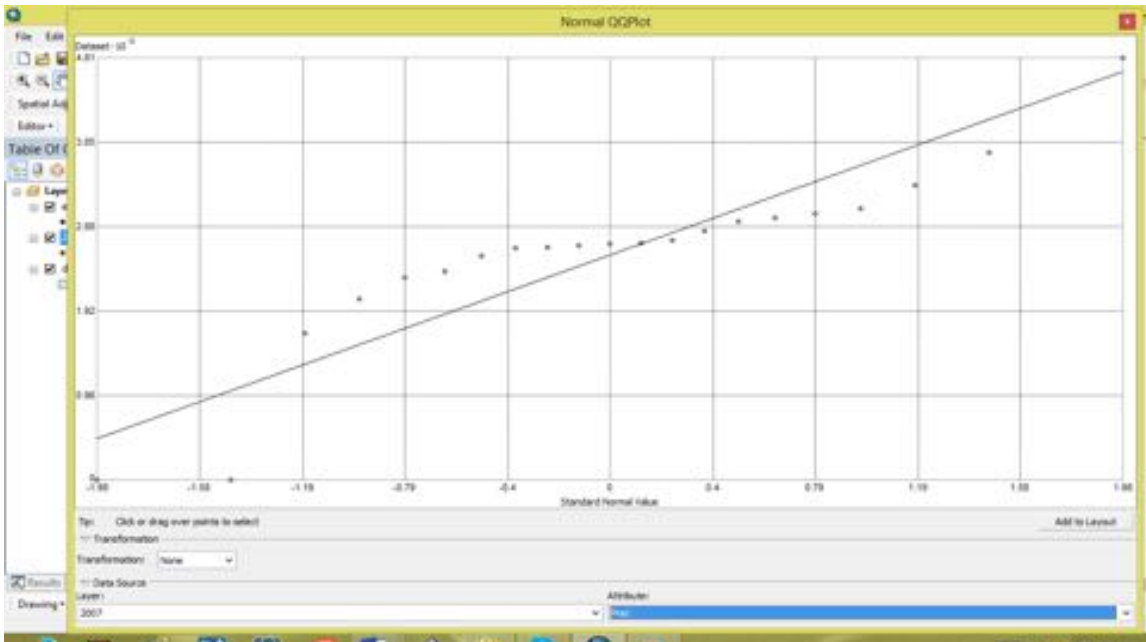
Año 2004:



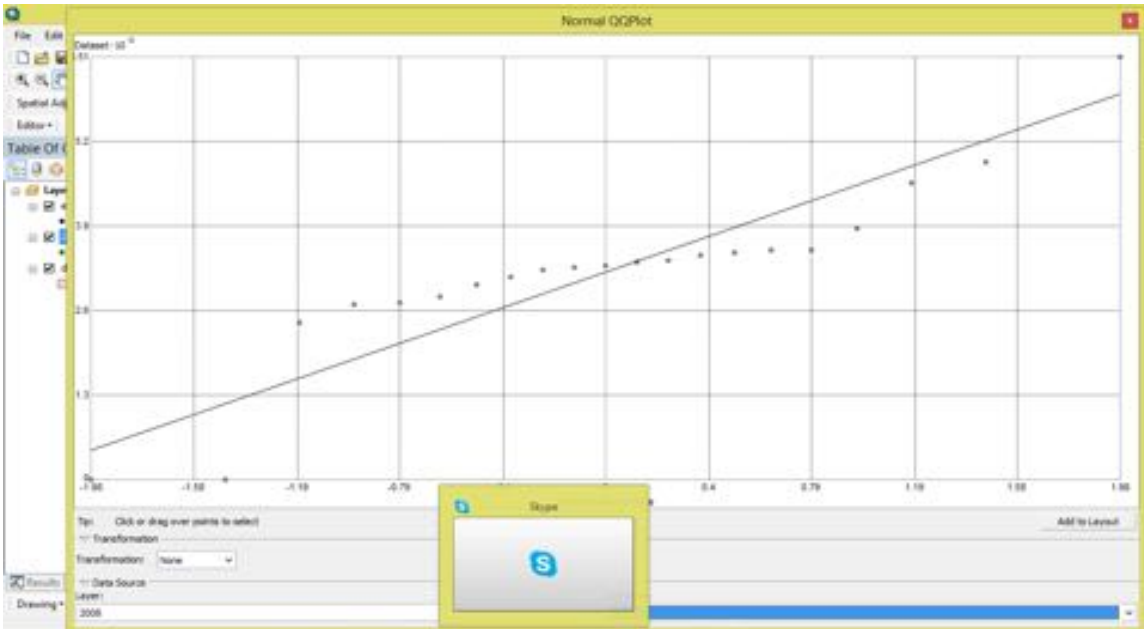
Año 2005:



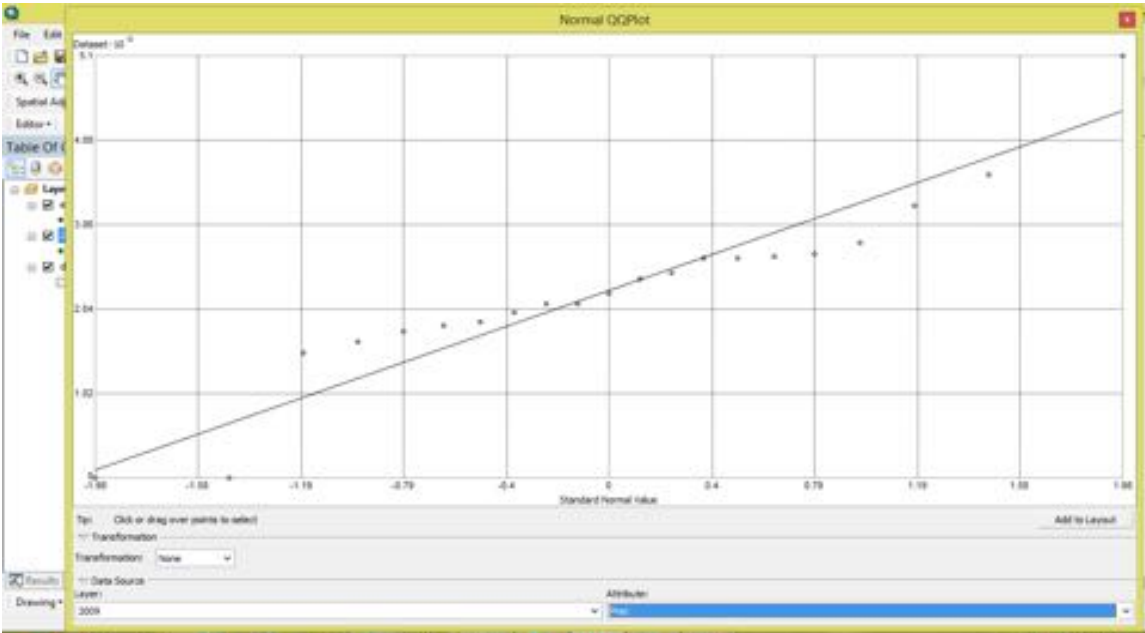
Año 2007:



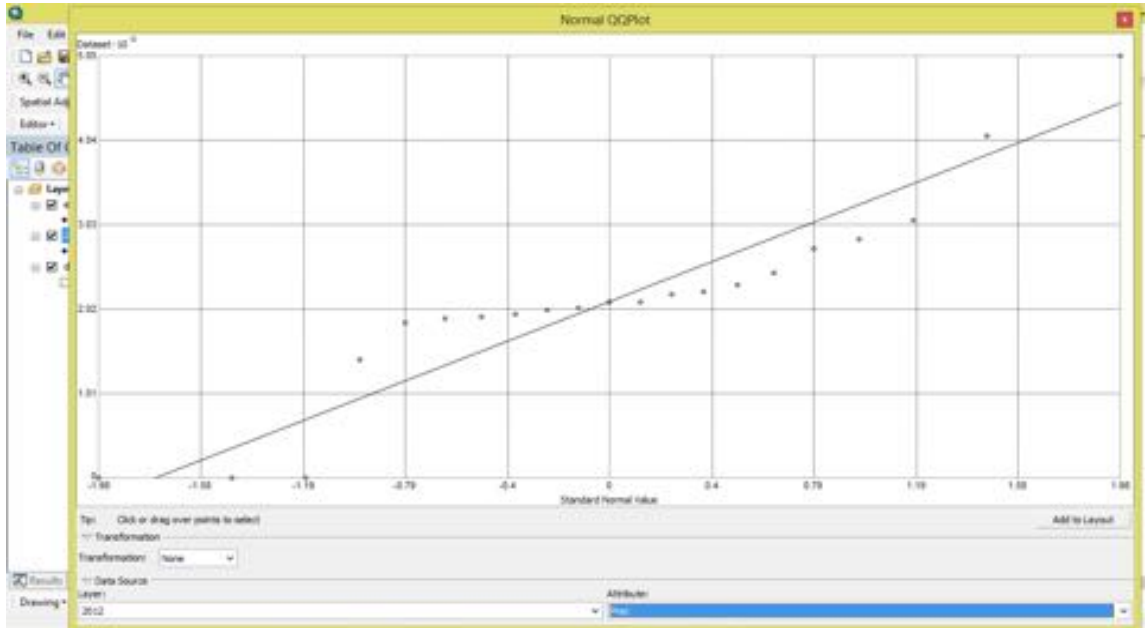
Año 2008:



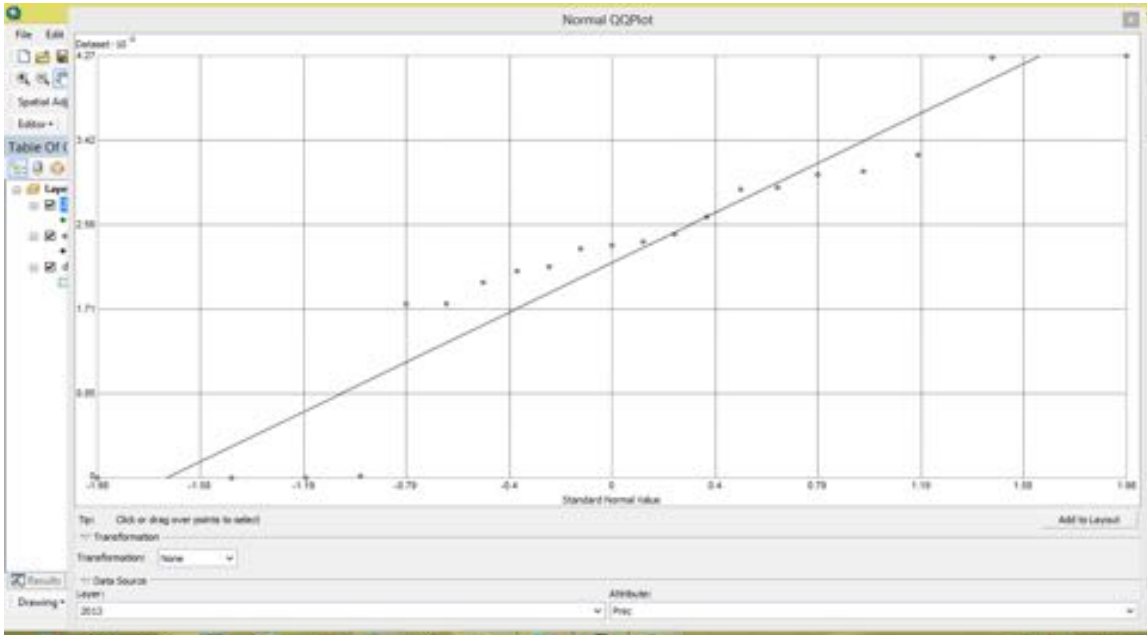
Año 2009:



Año 2012:



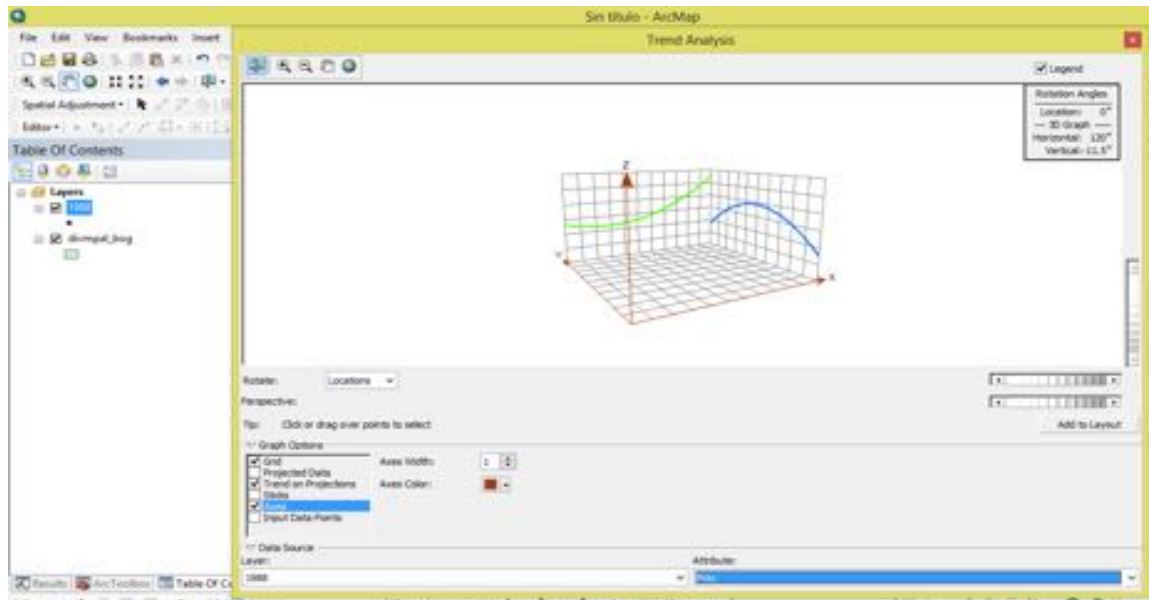
Año 2013:



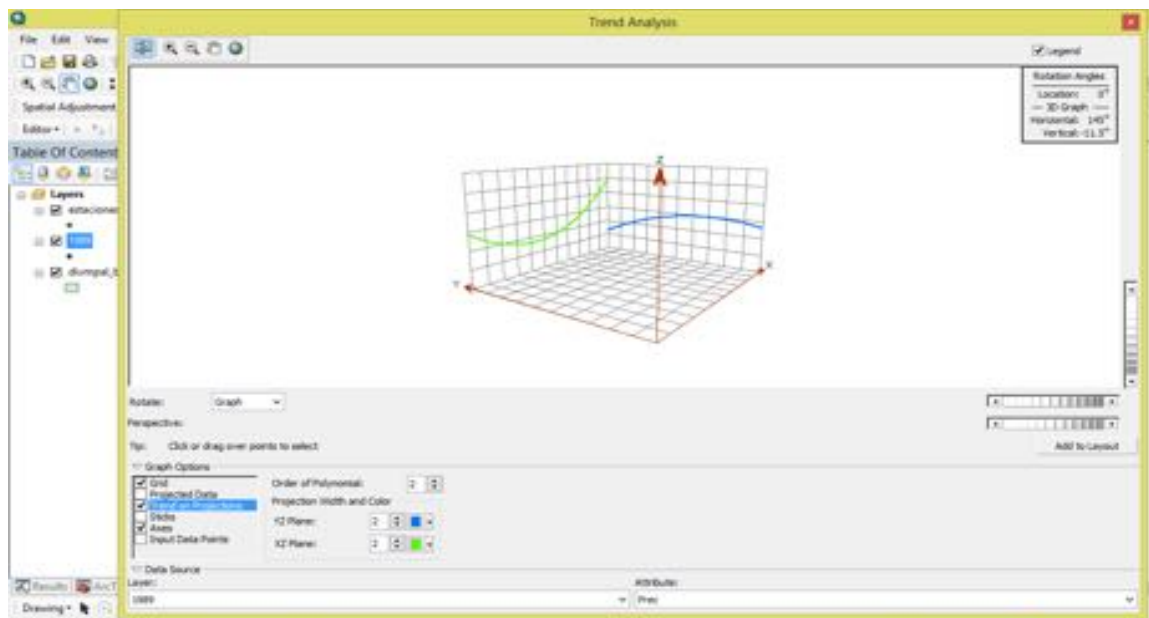
Anexo C

Análisis Exploratorio de Datos: Gráfico Trend

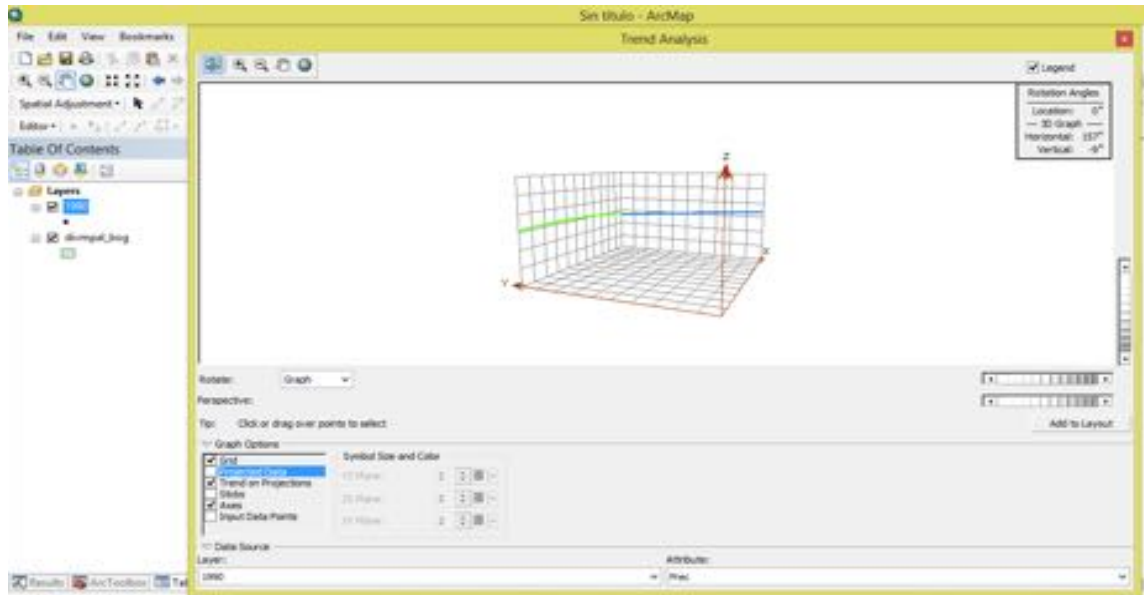
1988: Cuadrático, de segundo orden



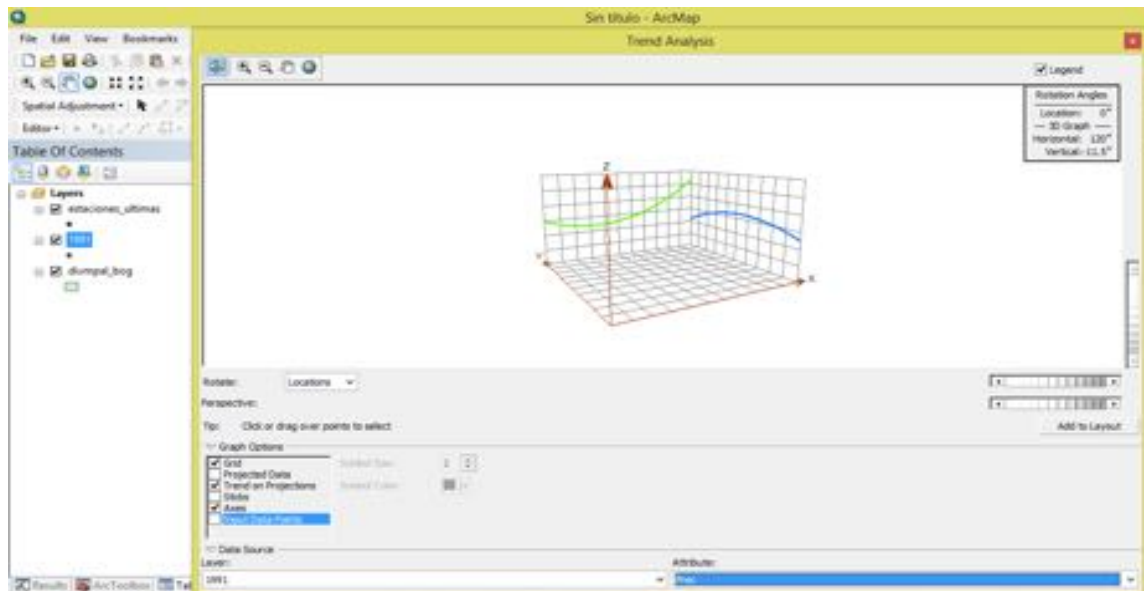
Año 1989: Cuadrático, de segundo orden



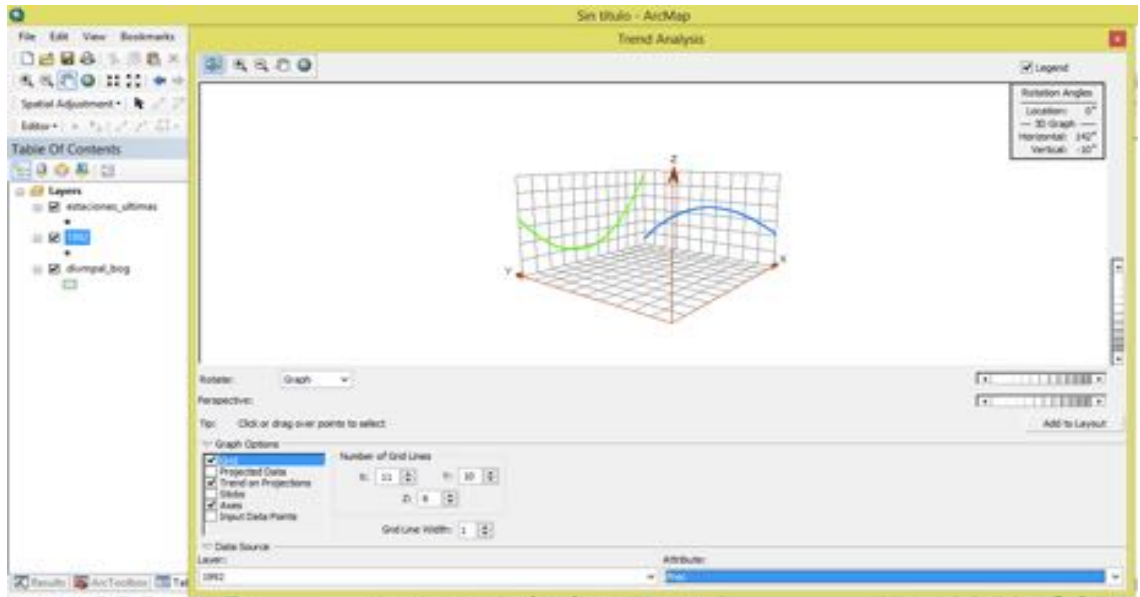
Año 1990: Lineal, de primer orden



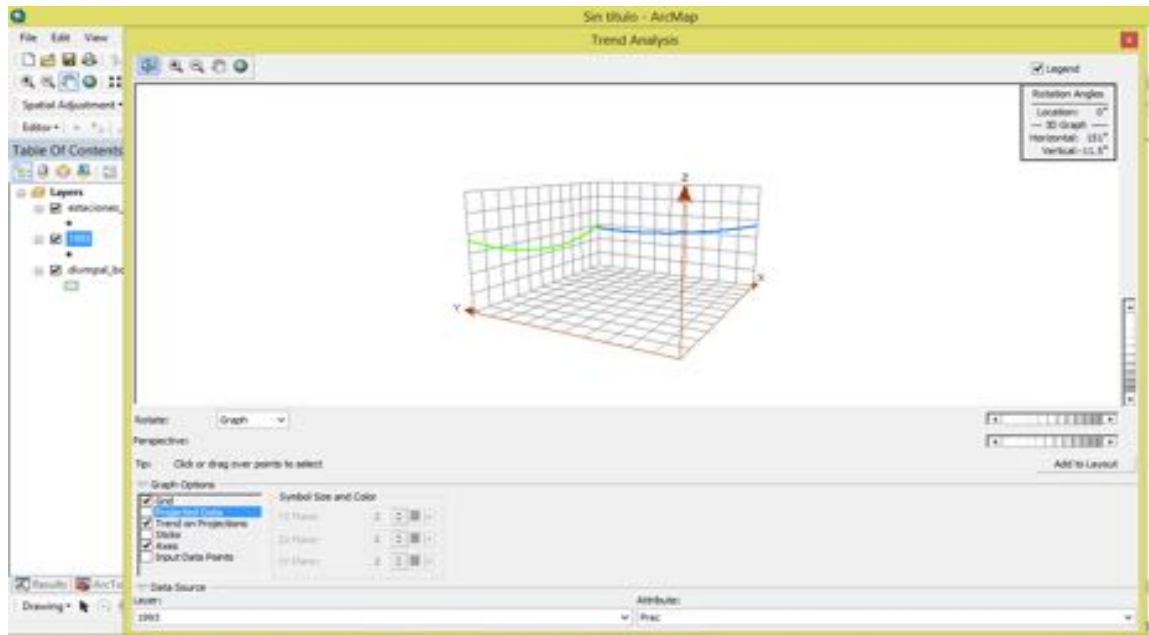
Año 1991: Cuadrático, de segundo orden



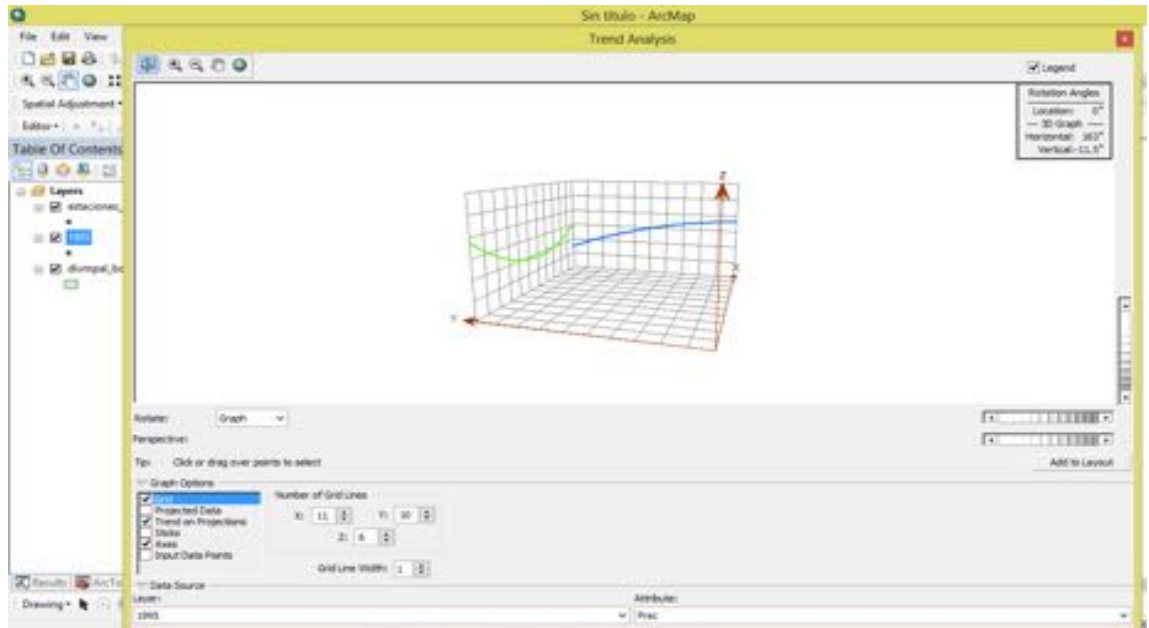
Año 1992: Cuadrático, de segundo orden



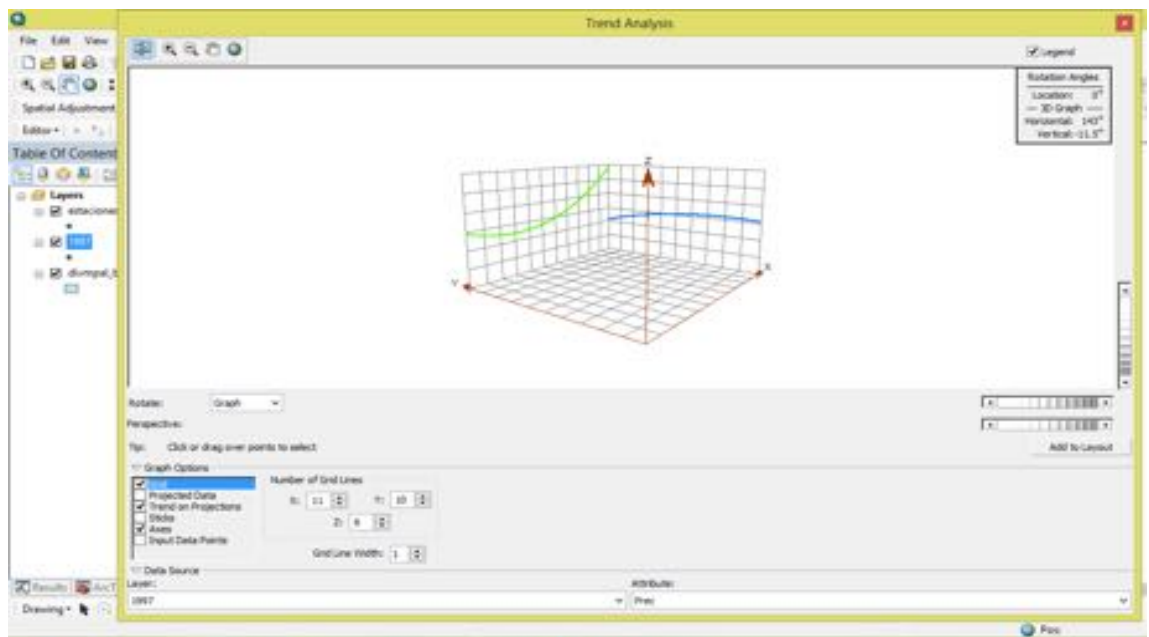
Año 1993: Cuadrático, de segundo orden



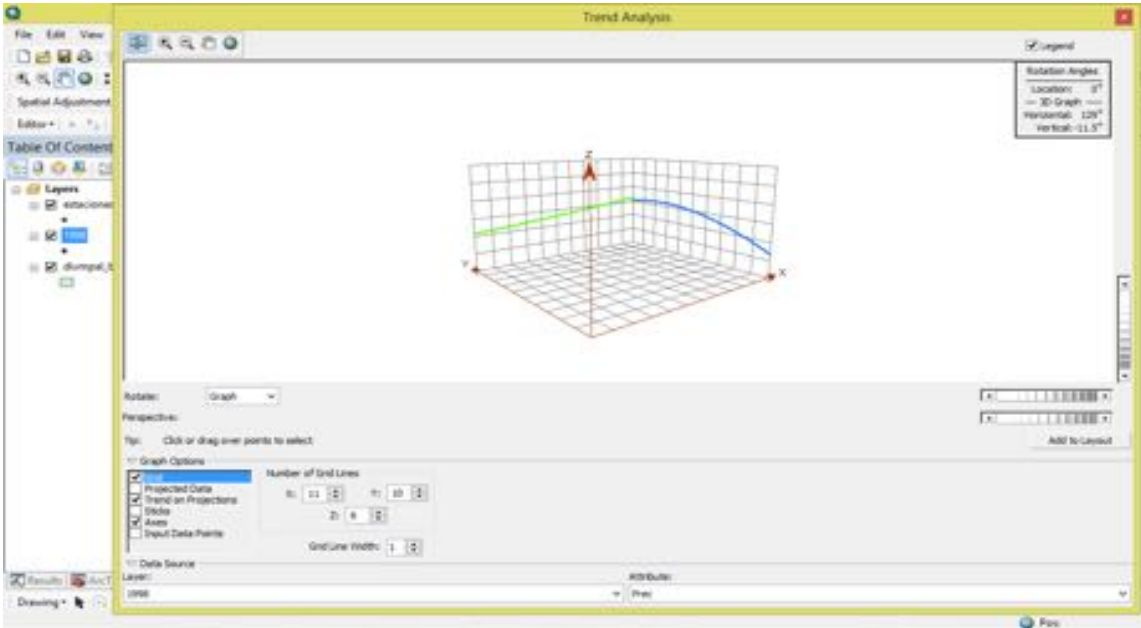
Año 1995: Cuadrático, de segundo orden



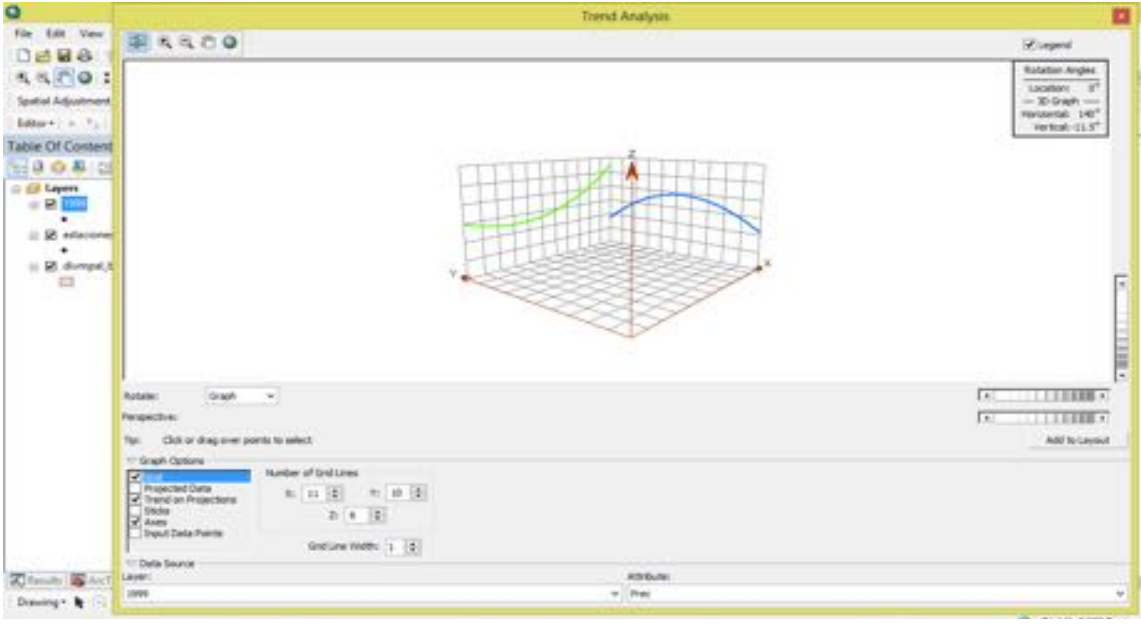
Año 1997: Cuadrático, de segundo orden



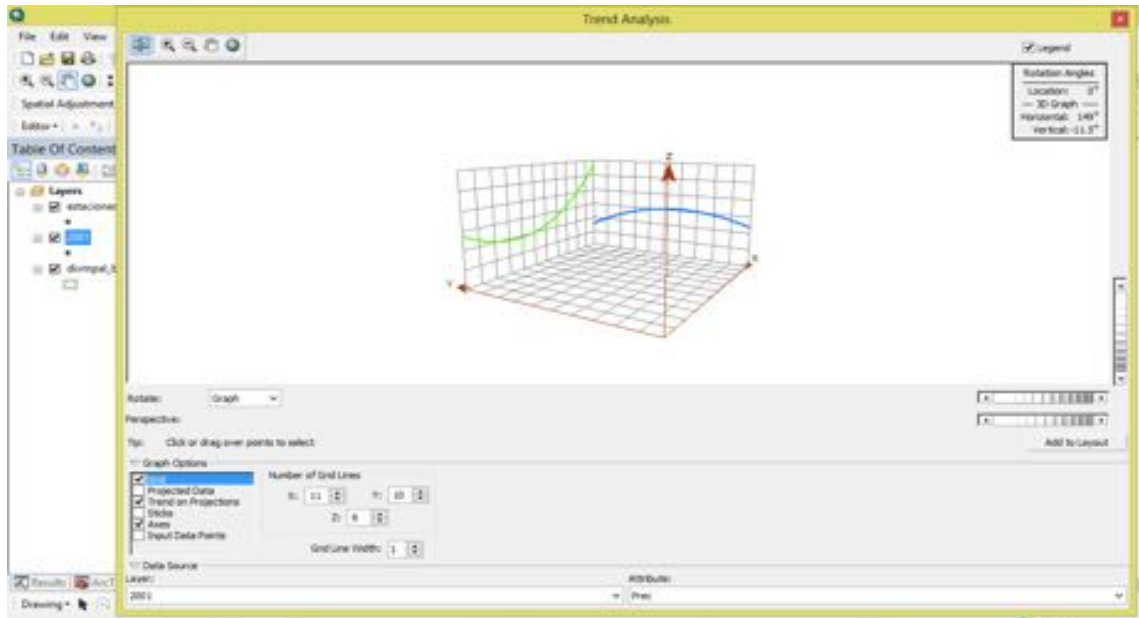
Año 1998: Cuadrático, de segundo orden



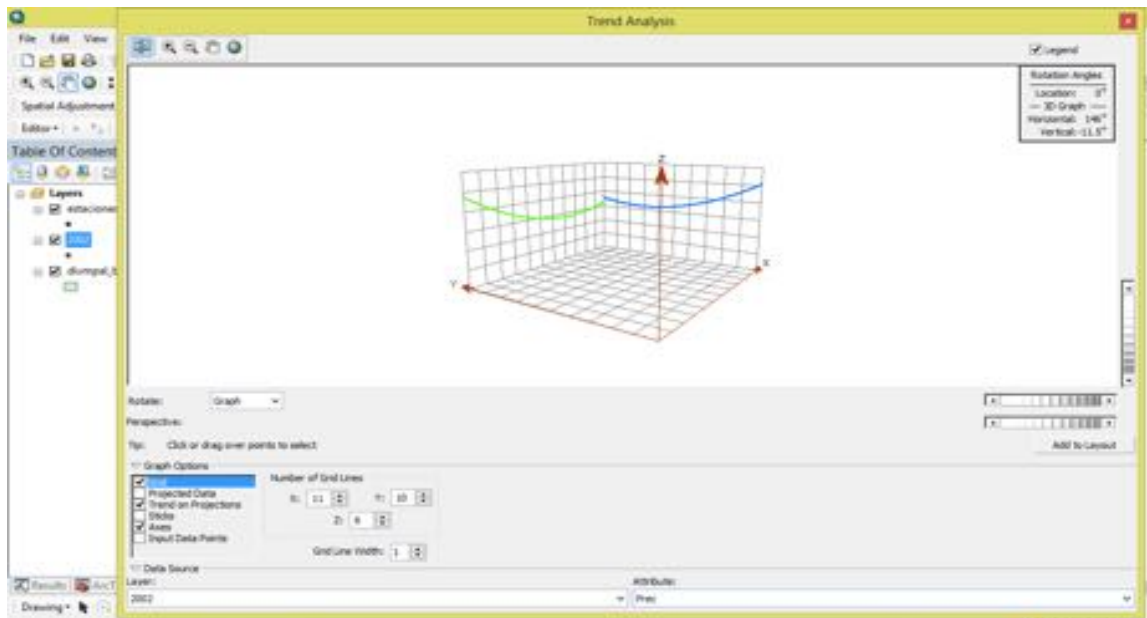
Año 1999: Cuadrático, de segundo orden



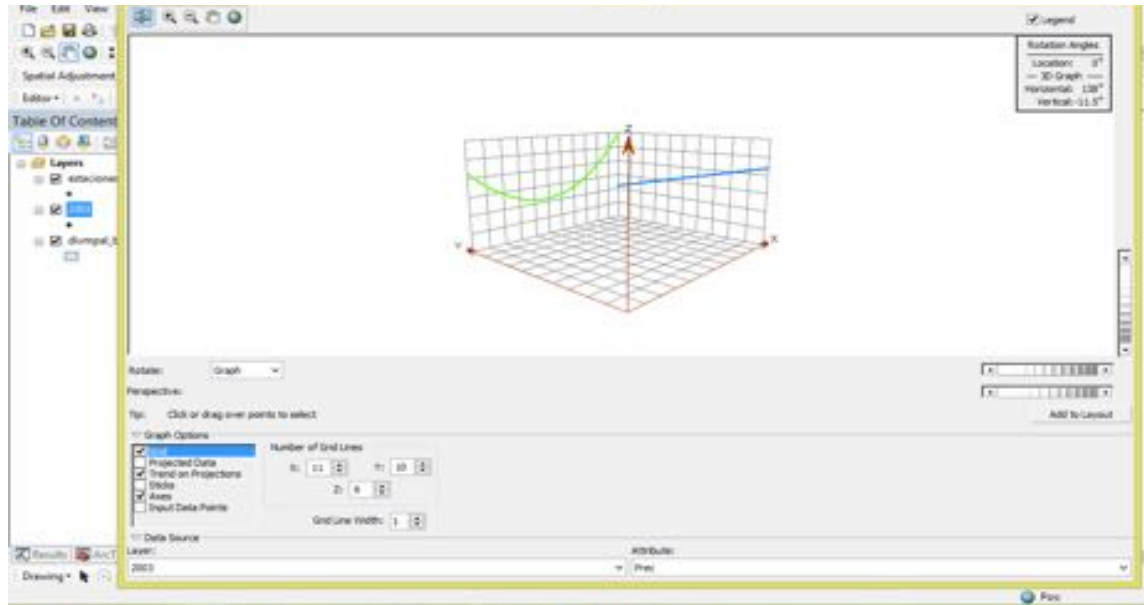
Año 2001: Cuadrático, de segundo orden



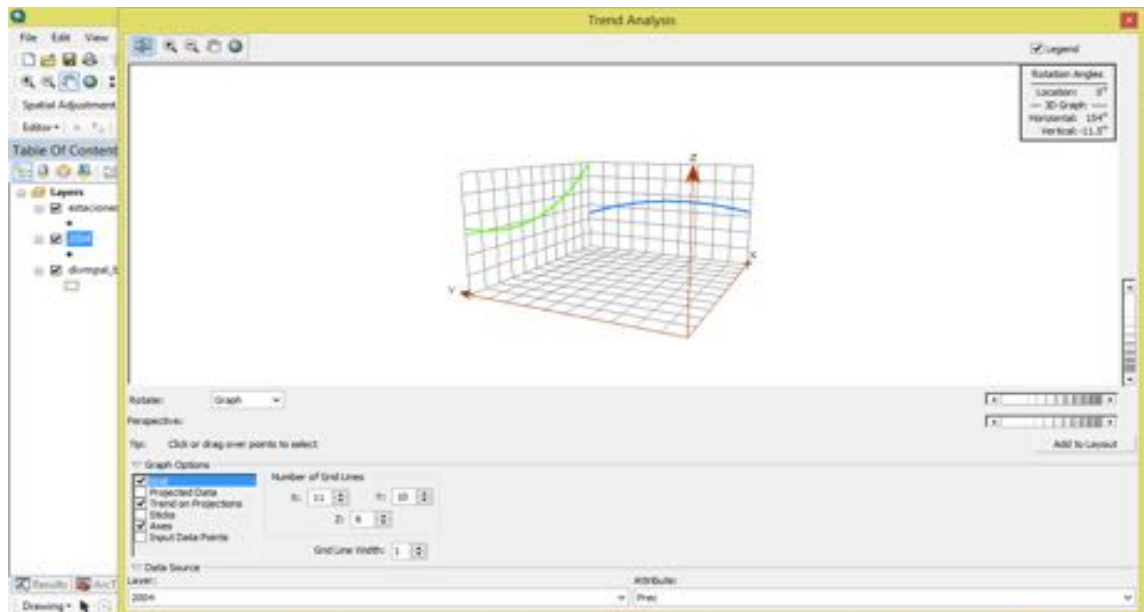
Año 2002: Cuadrático, de segundo orden



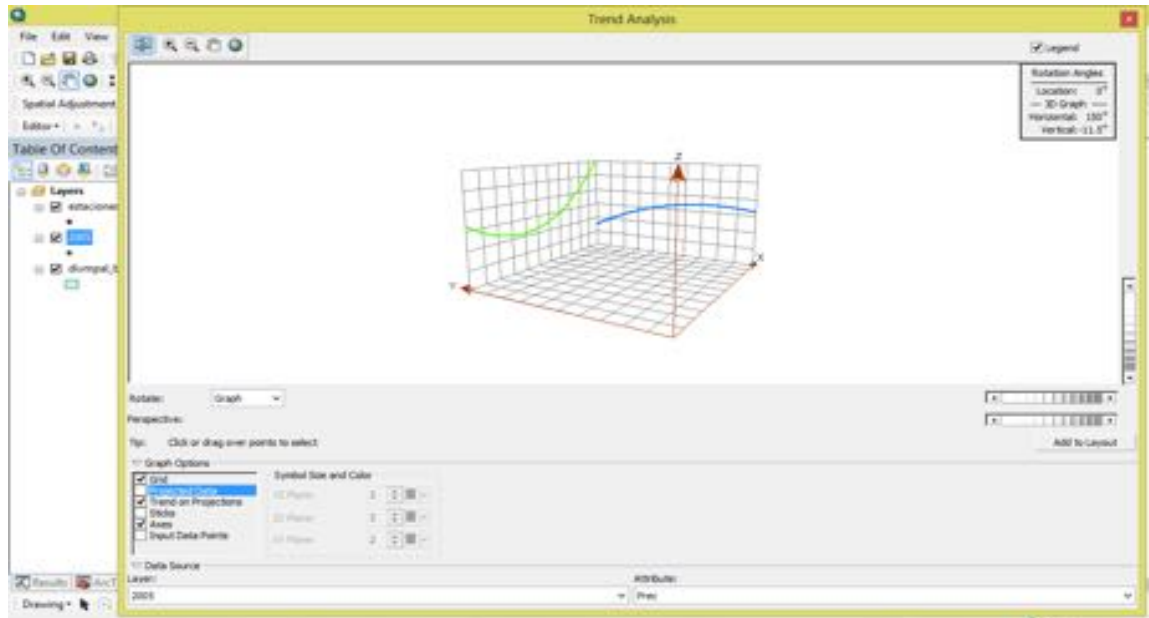
Año 2003: Lineal, de primer orden



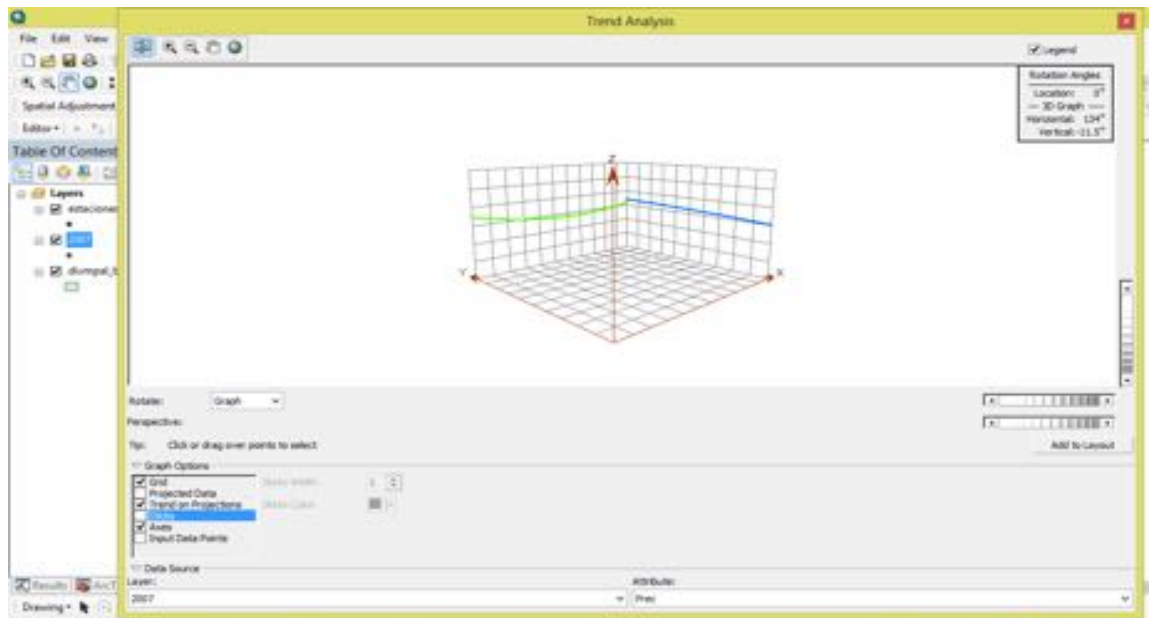
Año 2004: Cuadrático, de segundo orden



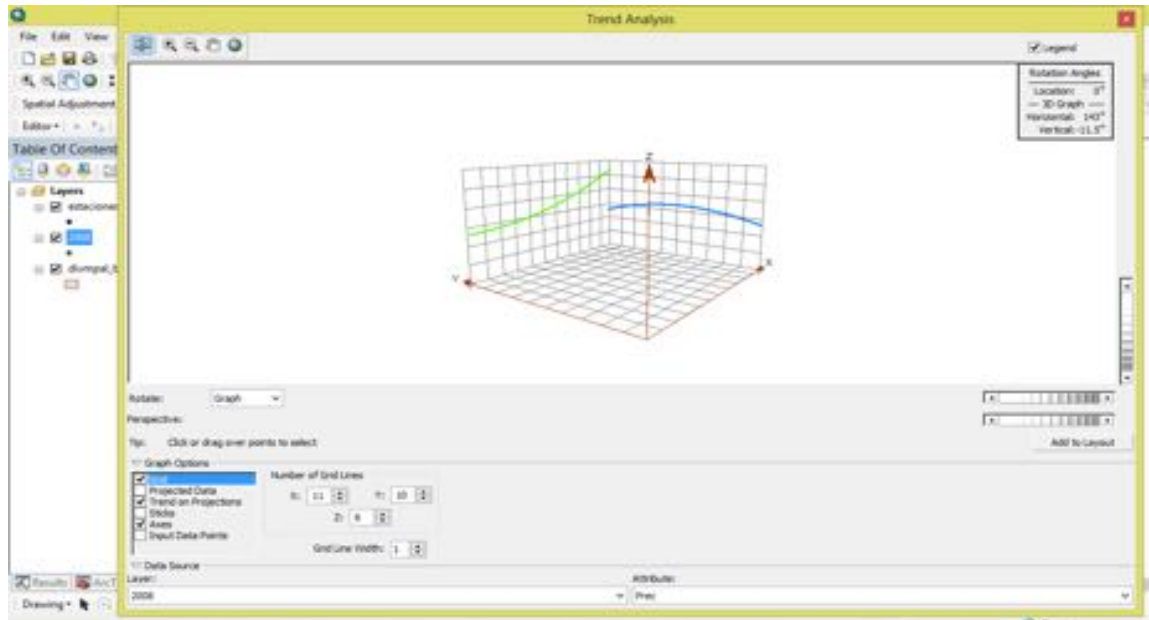
Año 2005: Cuadrático, de segundo orden



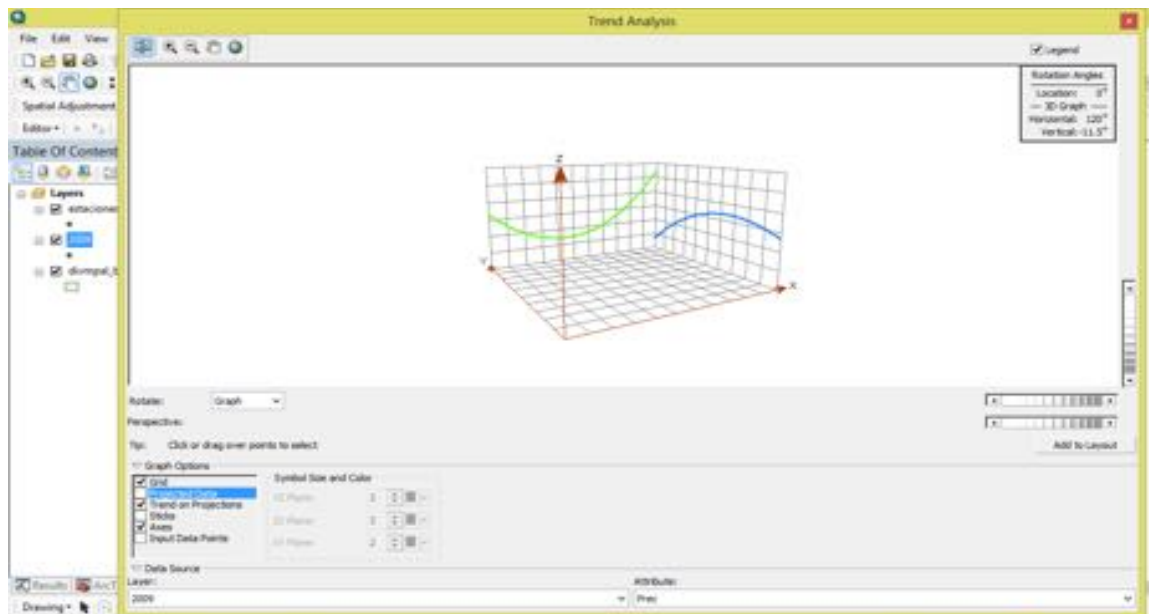
Año 2007: Lineal, de primer orden



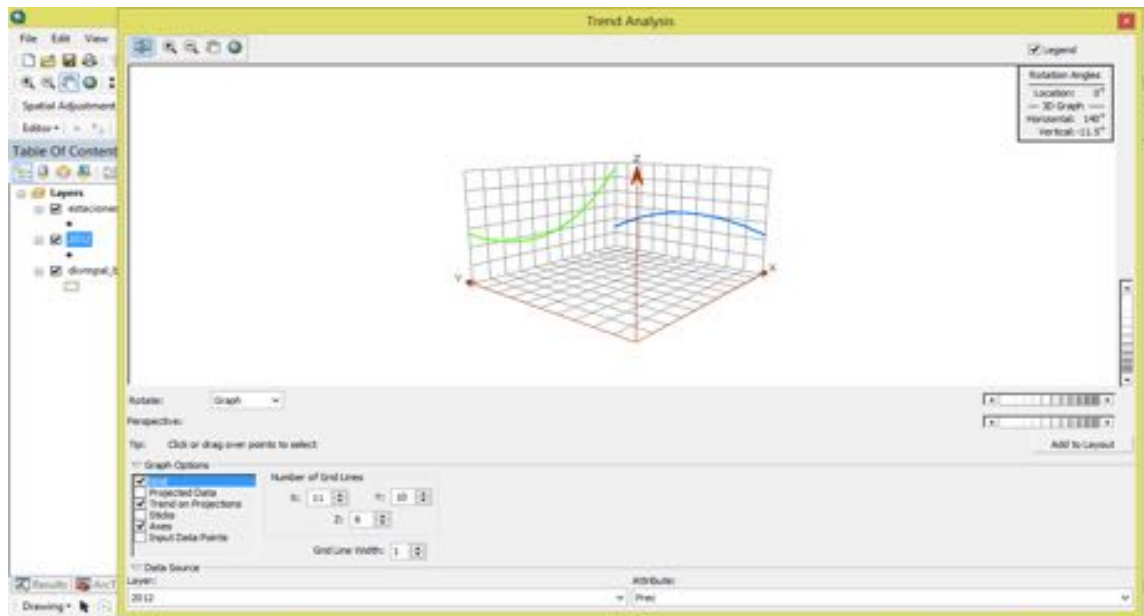
Año 2008: Cuadrático, de segundo orden



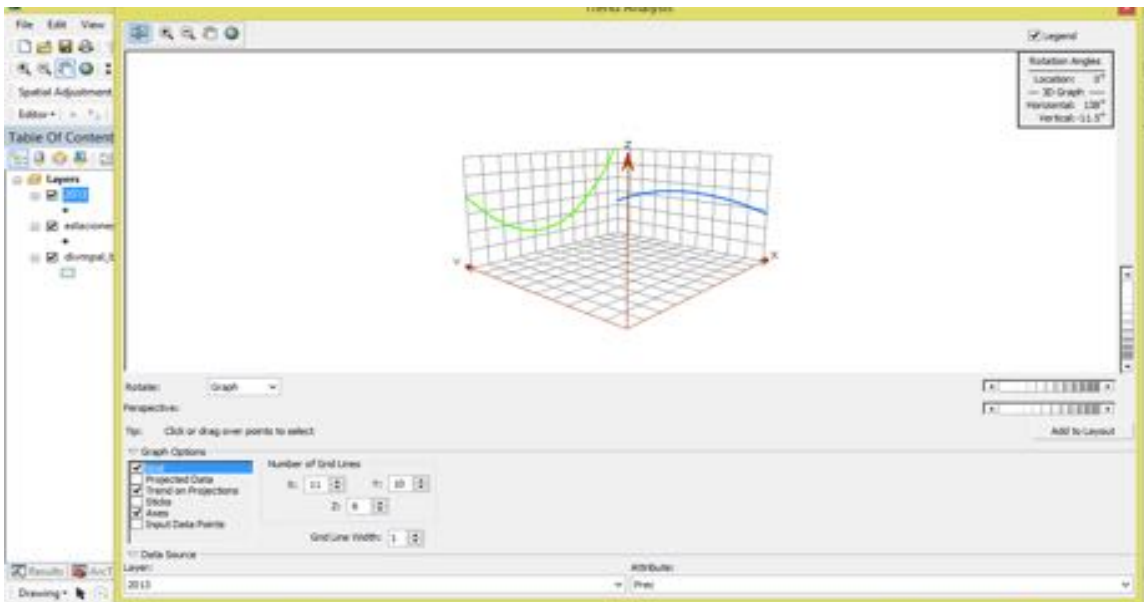
Año 2009: Cuadrático, de segundo orden



Año 2012: Cuadrático, de segundo orden



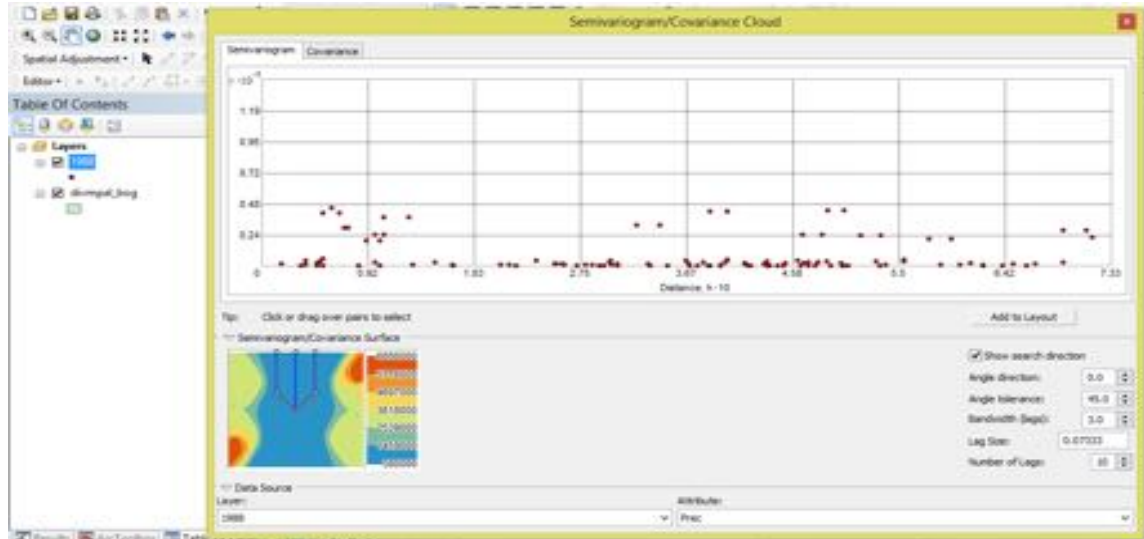
Año 2013: Cuadrático, de segundo orden



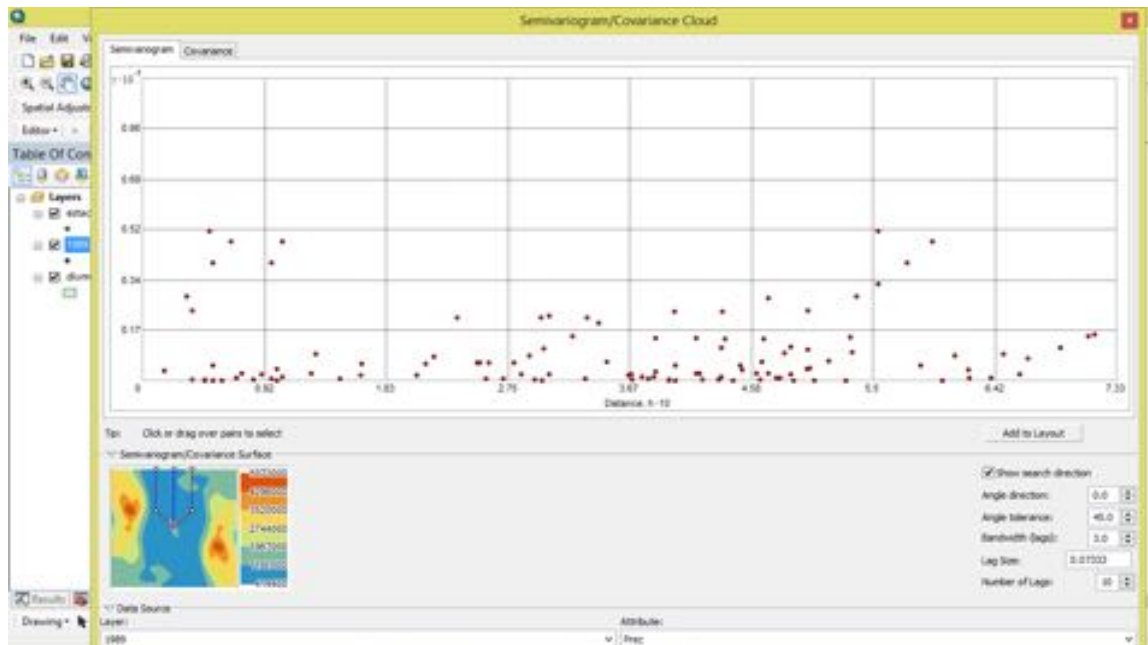
Anexo D

Análisis Exploratorio de Datos: Semivariogramas

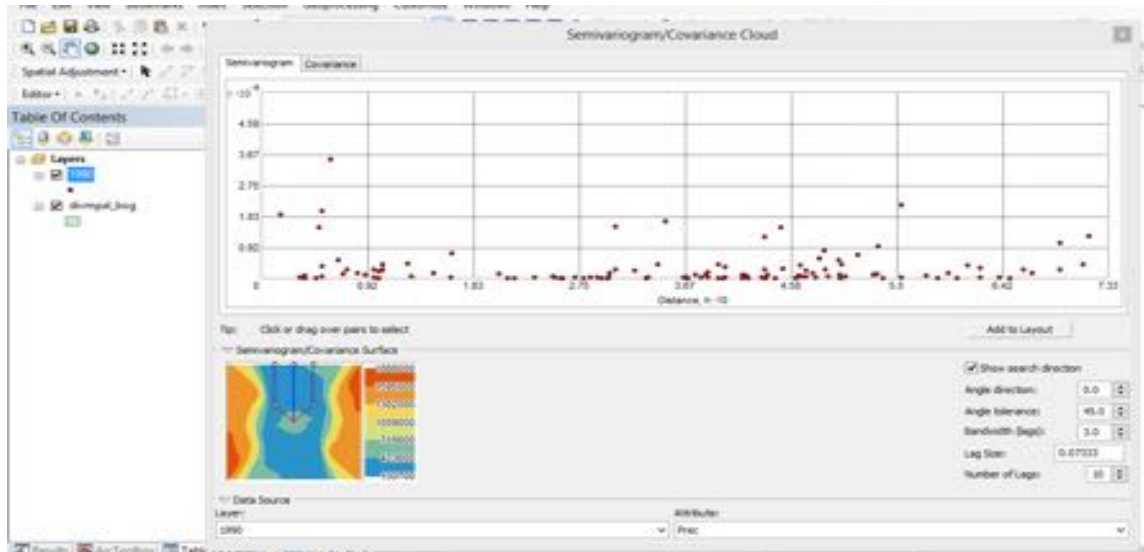
Año 1988:



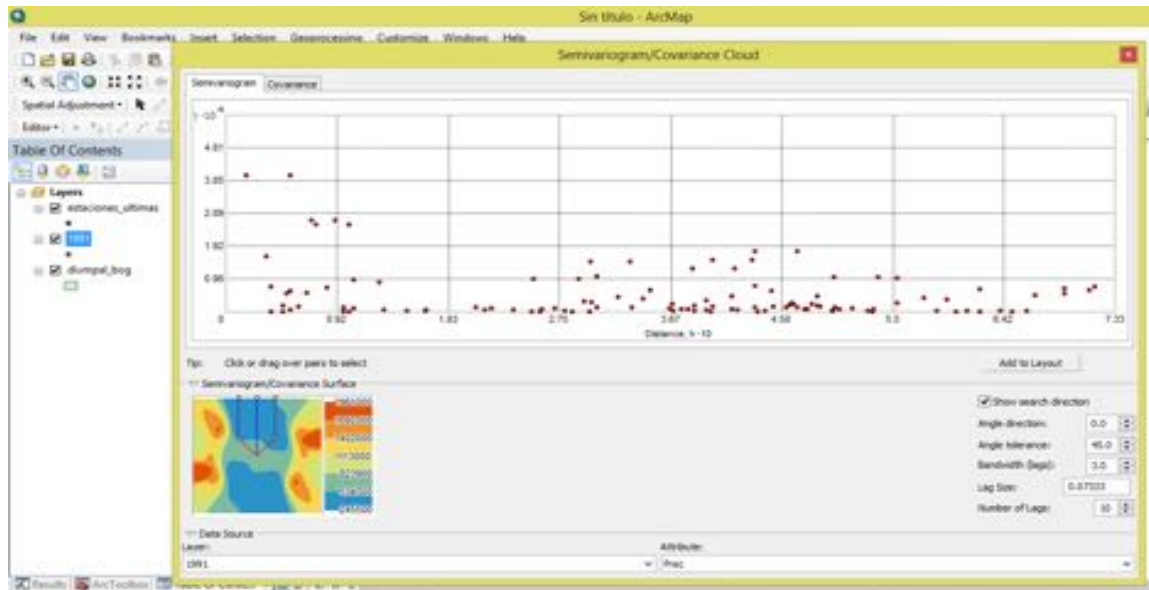
Año 1989:



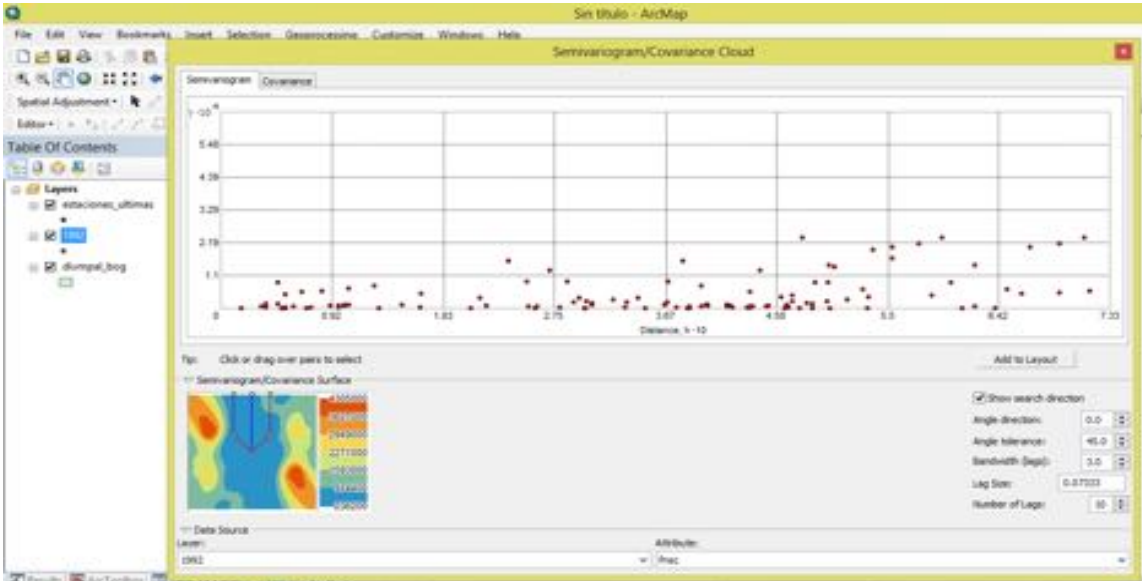
Año 1990:



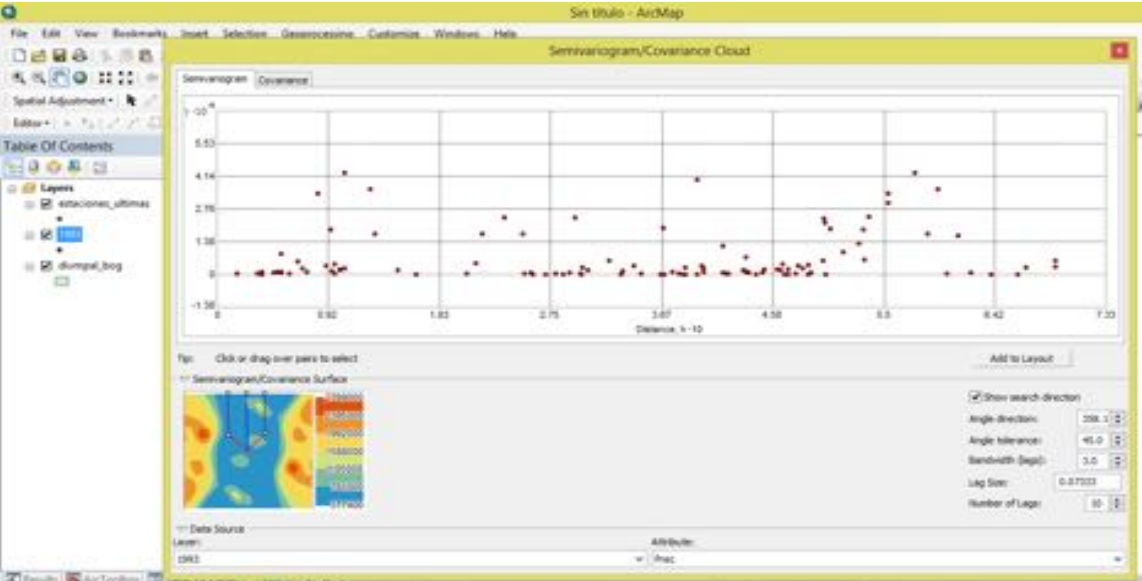
Año 1991:



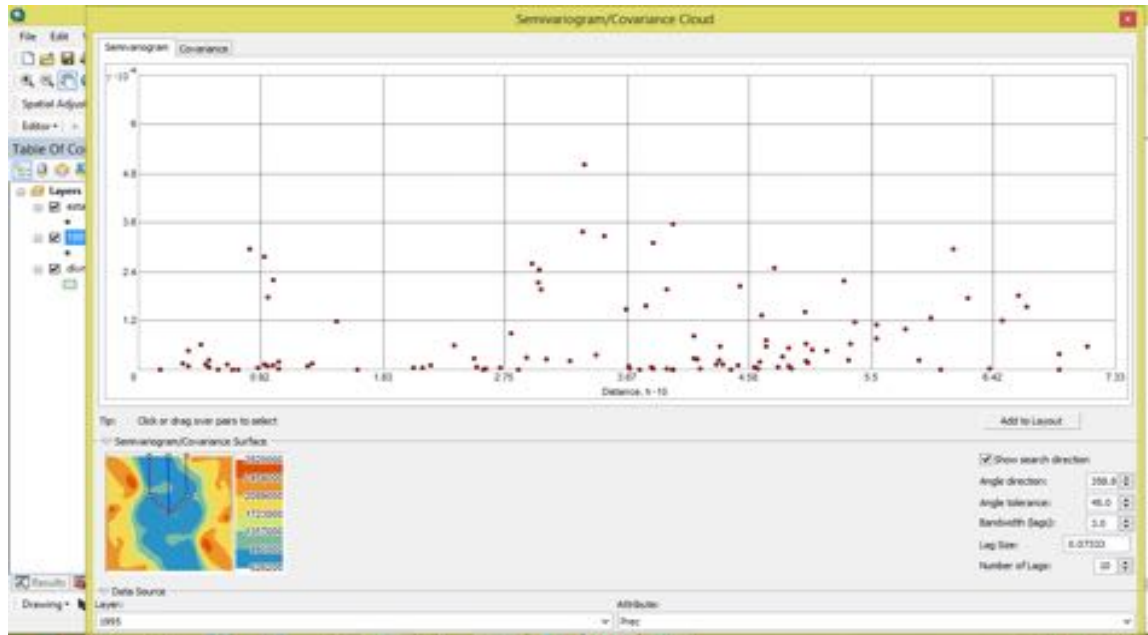
Año 1992:



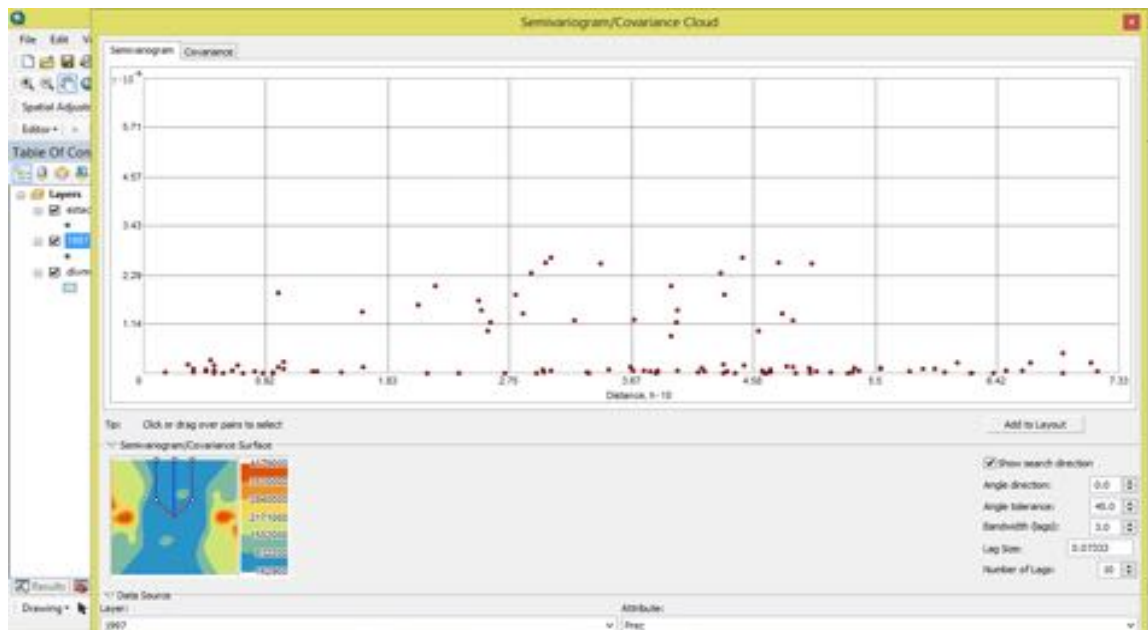
Año 1993:



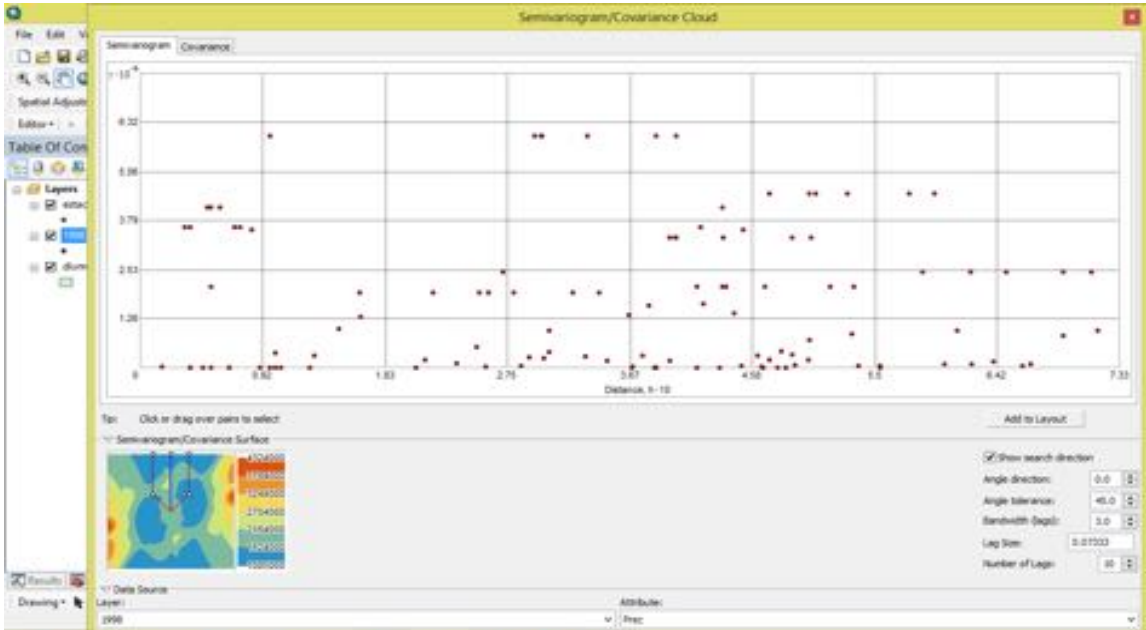
Año 1995:



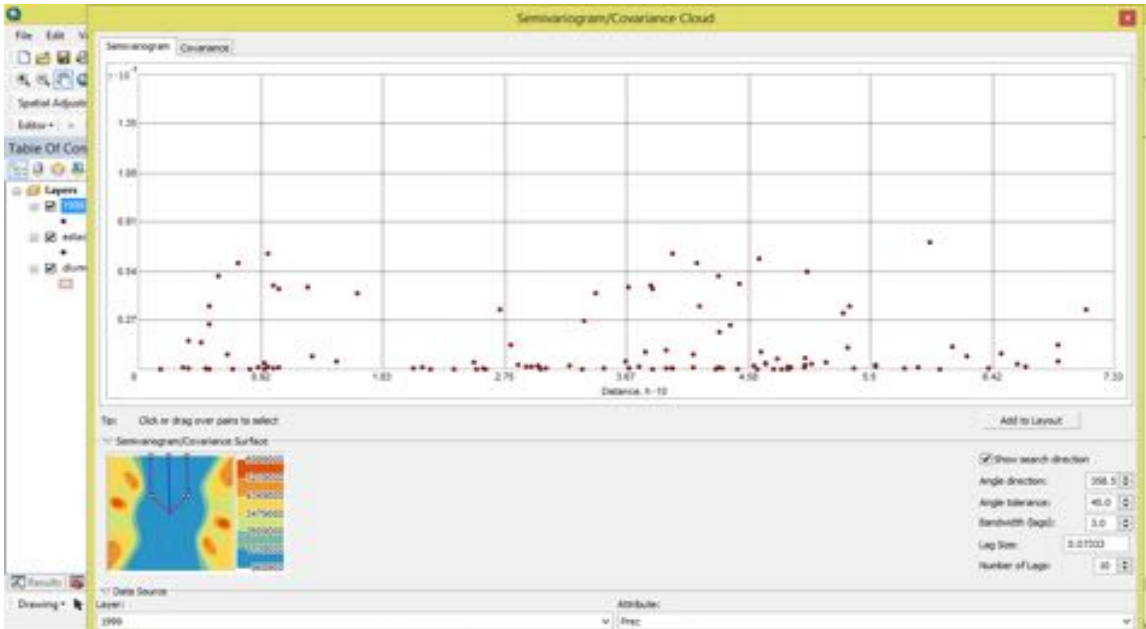
Año 1997:



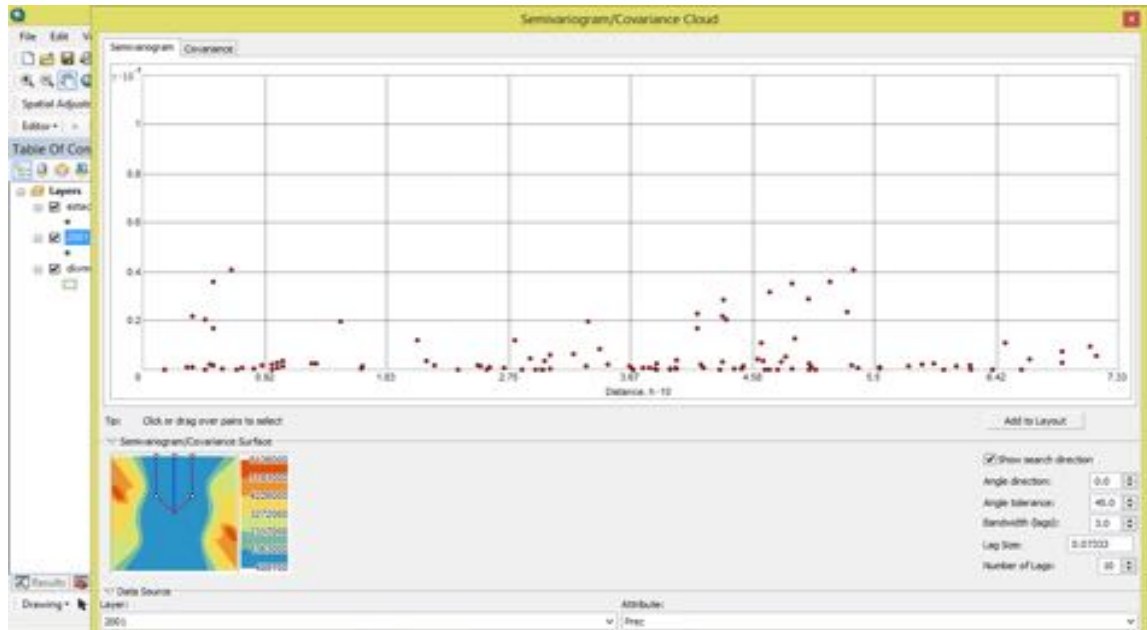
Año 1998:



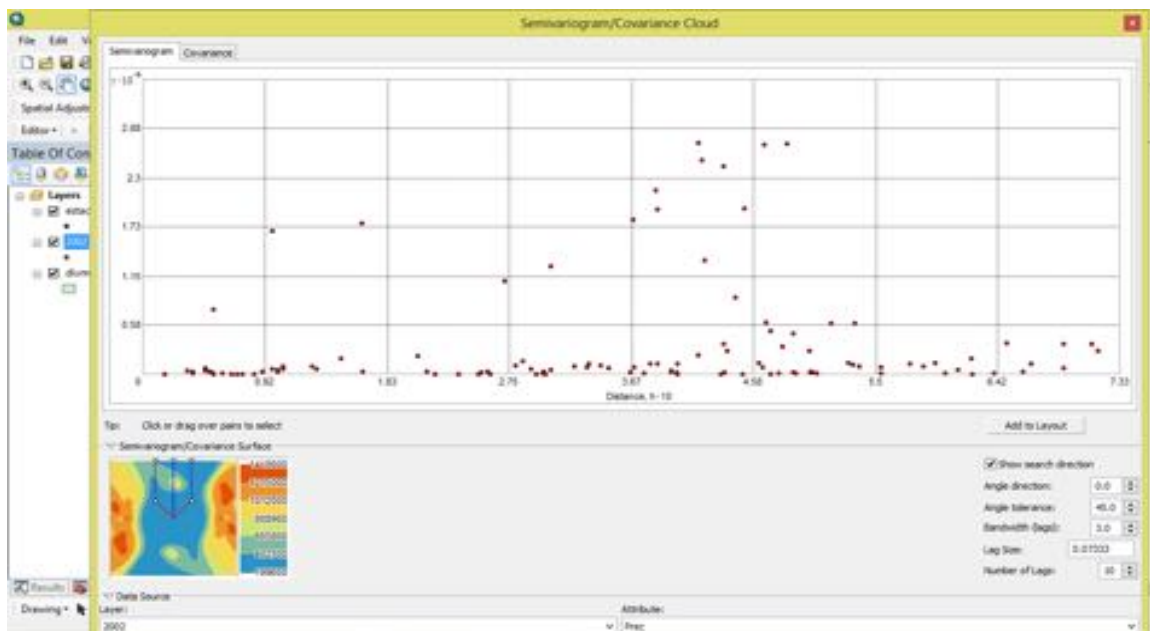
Año 1999:



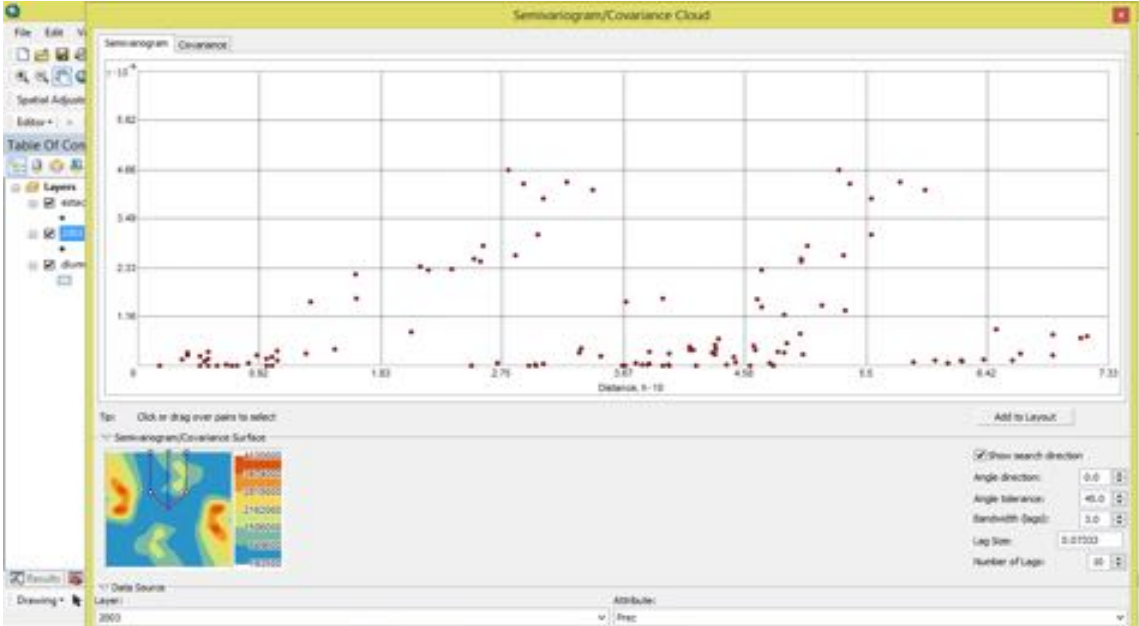
Año 2001:



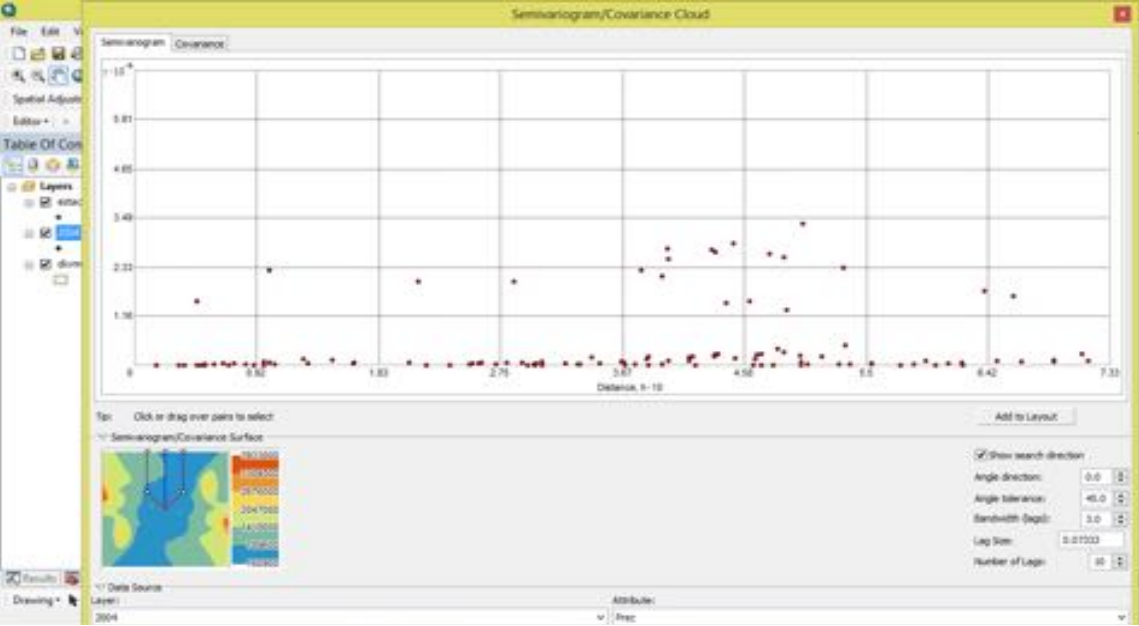
Año 2002:



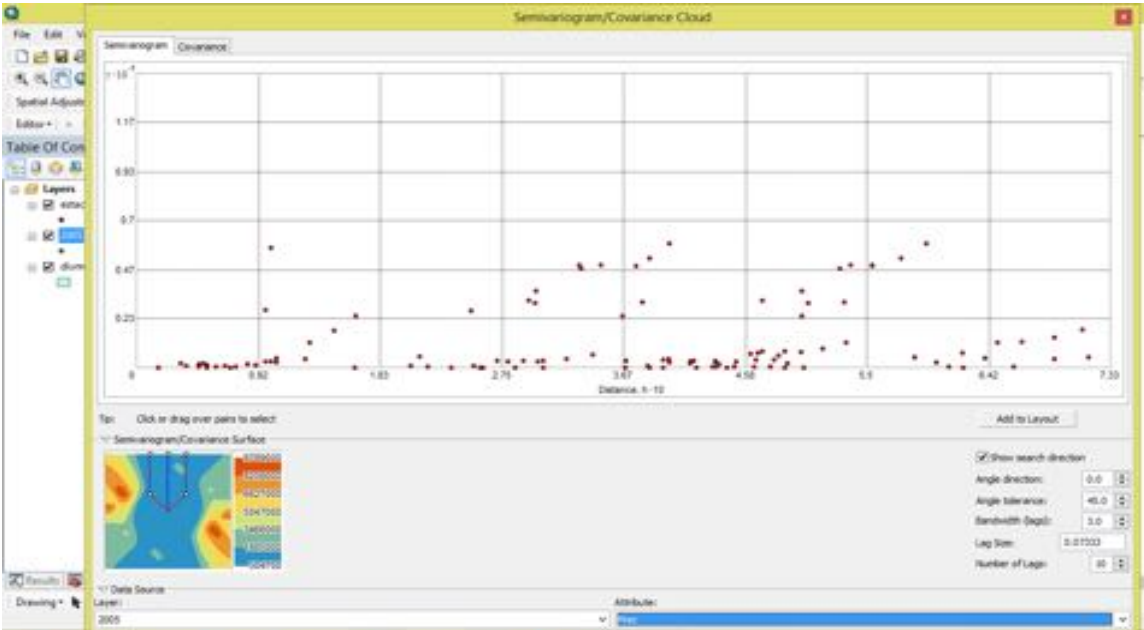
Año 2003:



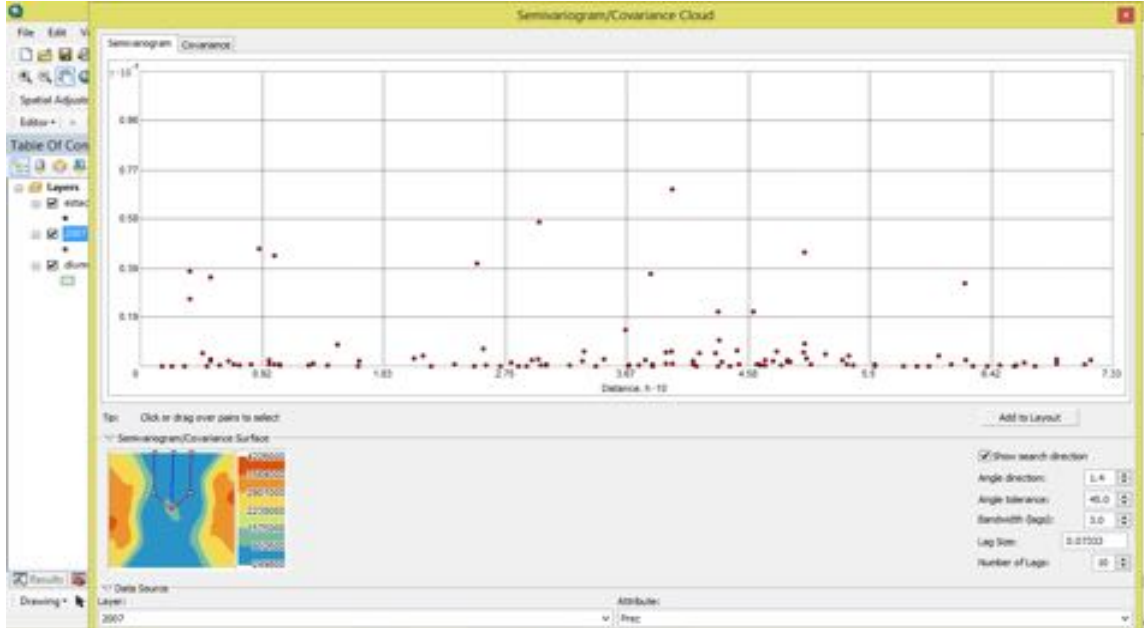
Año 2004:



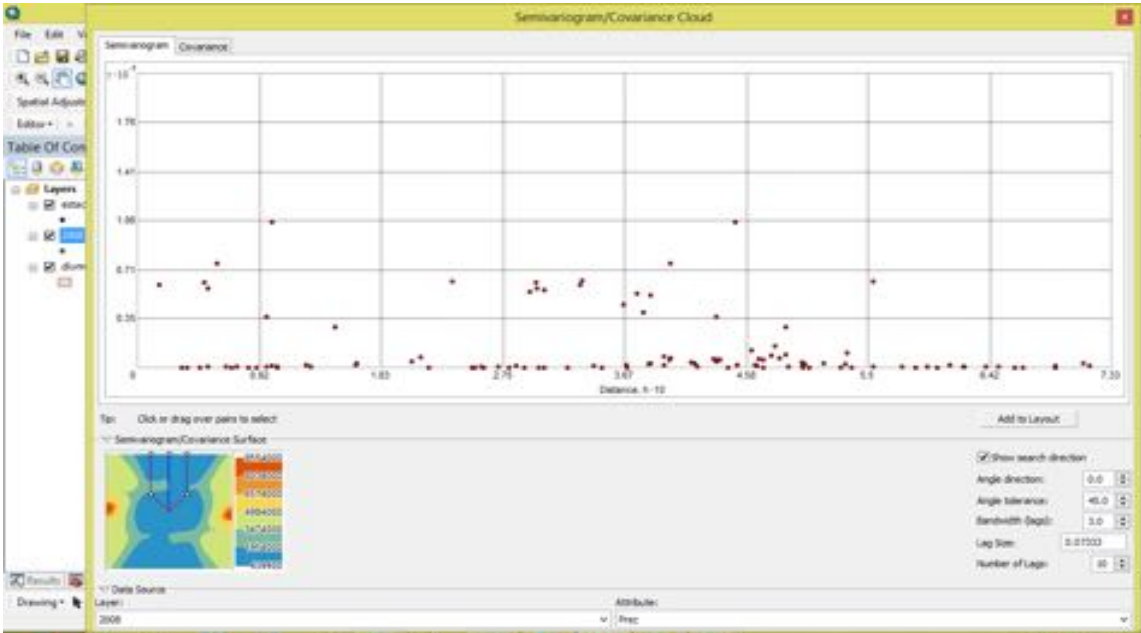
Año 2005:



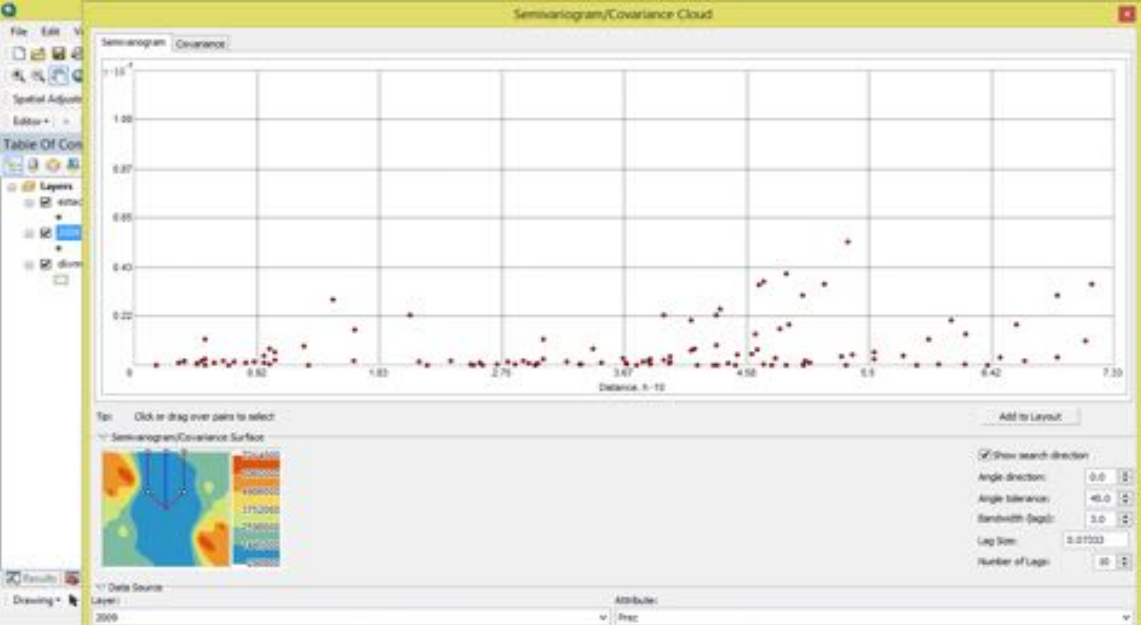
Año 2007:



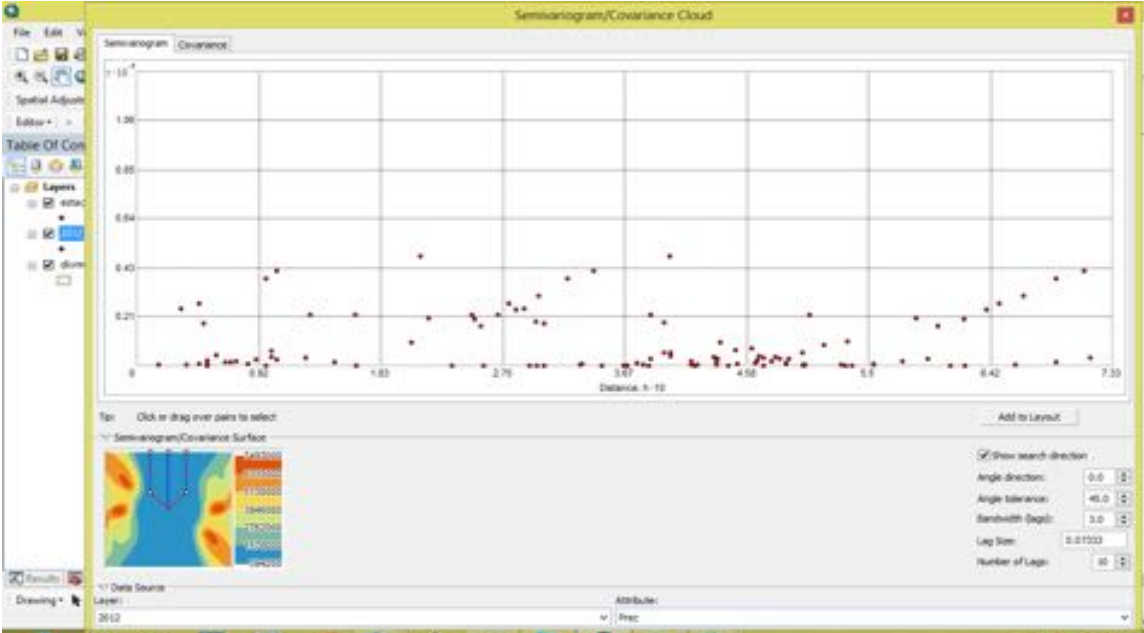
Año 2008:



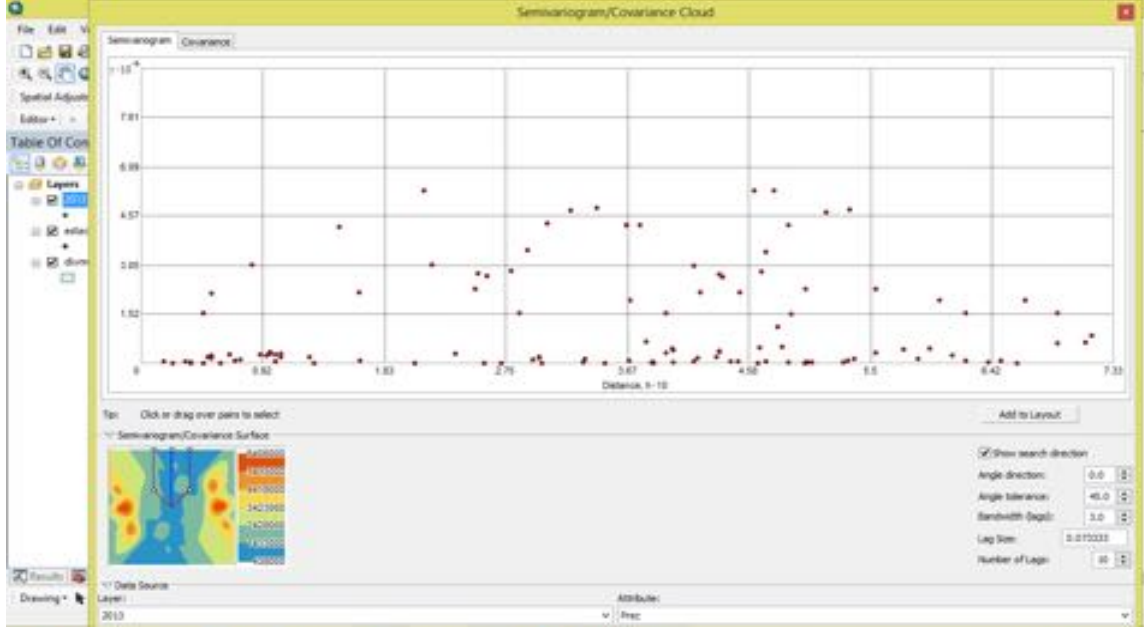
Año 2009:



Año 2012:

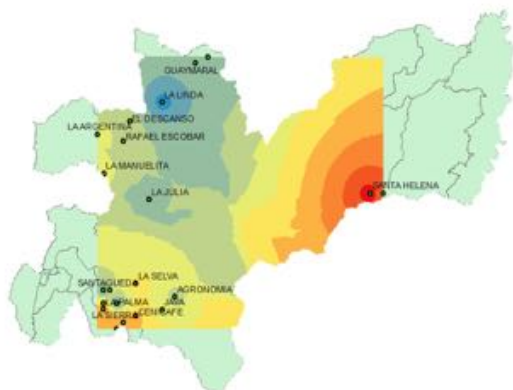


Año 2013:



Anexo E

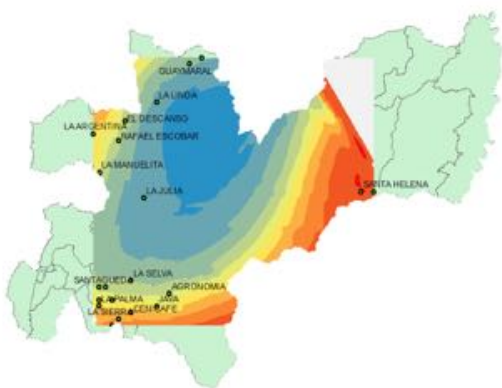
Superficies de interpolación para datos de precipitación anual



Método IDW año 1989



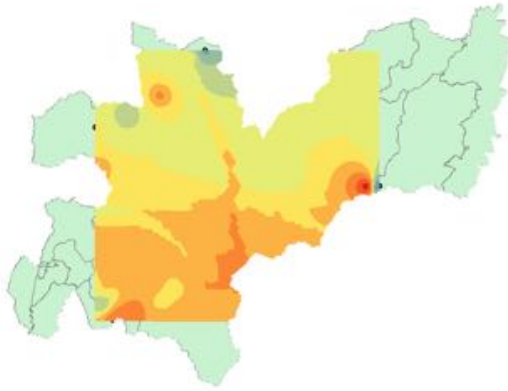
Método GPI año 1989



Método Kriging Ordinario año 1989



Método Kriging Universal año 1989



Método IDW año 1990



Método GPI año 1990



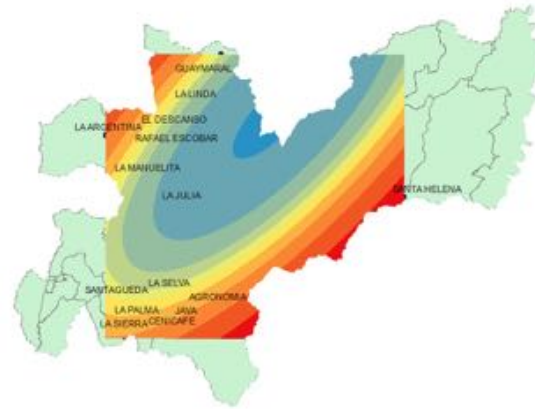
Método Kriging Ordinario año 1990



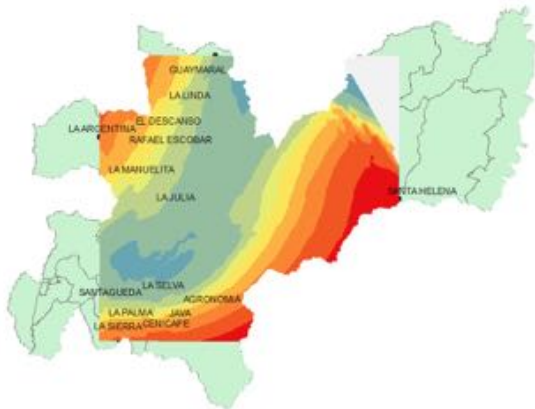
Método Kriging Universal año 1990



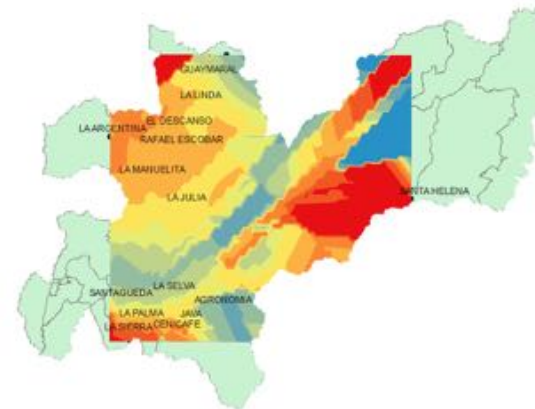
Método IDW año 1991



Método GPI año 1991



Método Kriging Ordinario año 1991



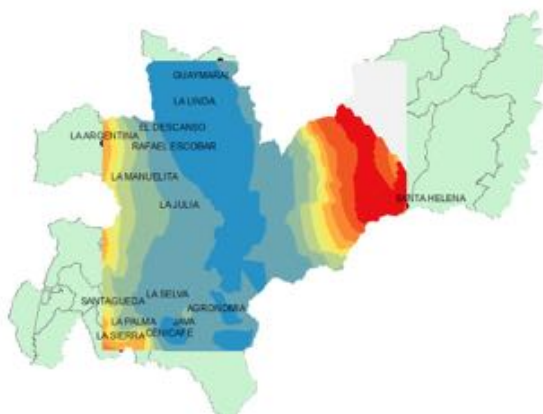
Método Kriging Universal año 1991



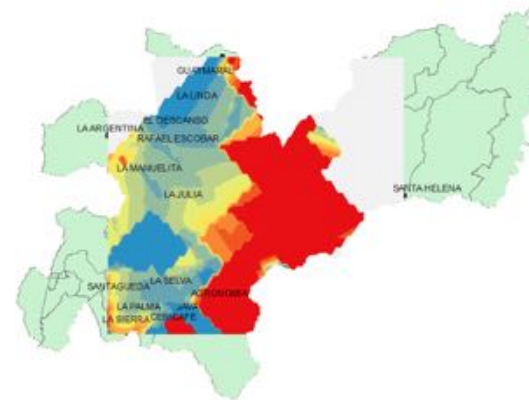
Método IDW año 1992



Método GPI año 1992



Método Kriging Ordinario año 1992



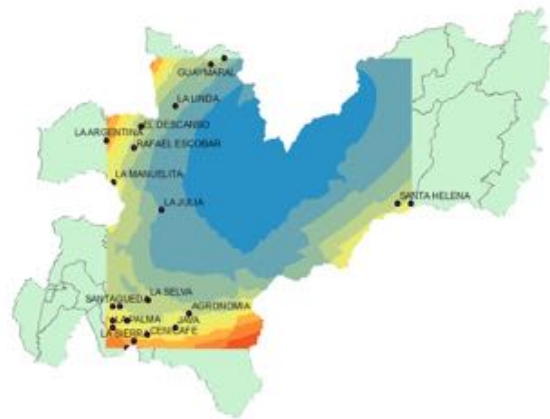
Método Kriging Universal año 1992



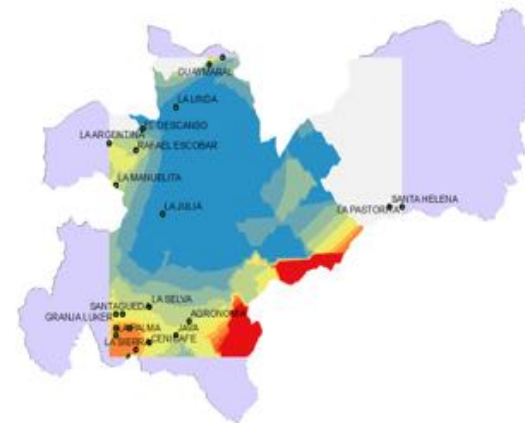
Método IDW año 1993



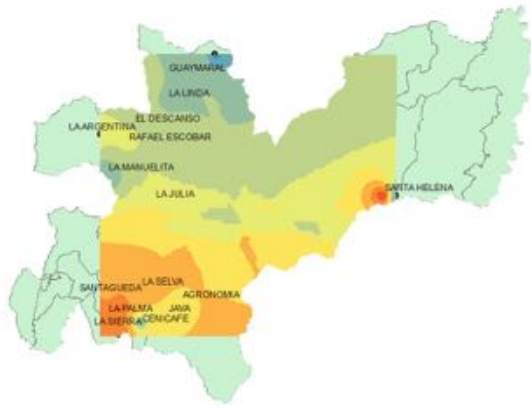
Método GPI año 1993



Método Kriging Ordinario año 1993



Método Kriging Universal año 1993



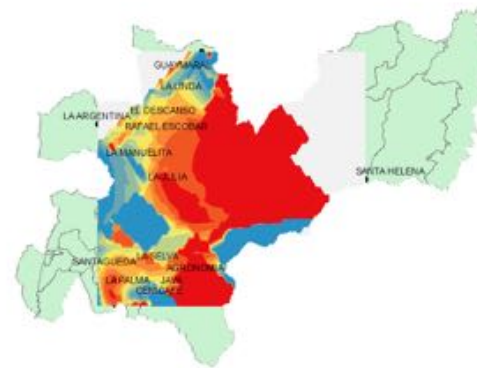
Método IDW año 1995



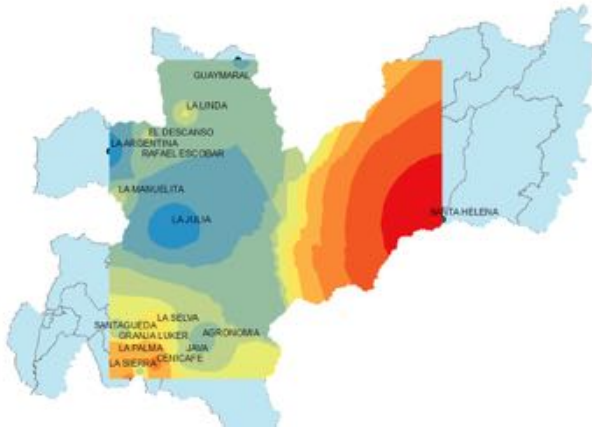
Método GPI año 1995



Método Kriging Ordinario año 1995



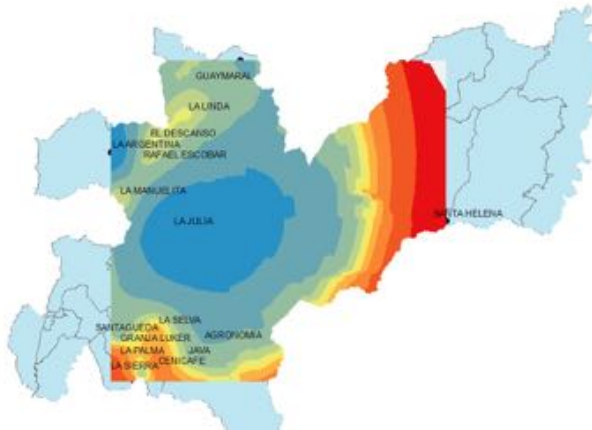
Método Kriging Universal año 1995



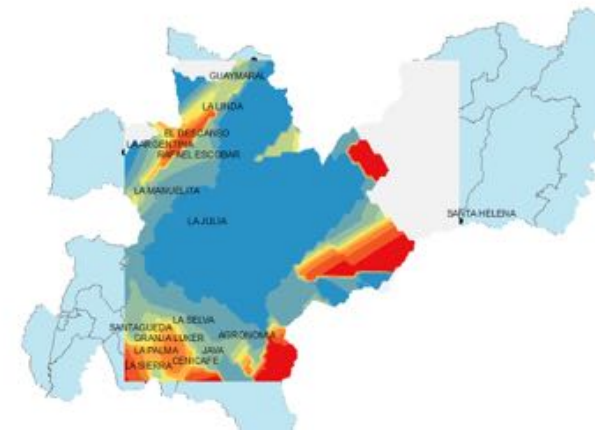
Método IDW año 1997



Método GPI año 1997



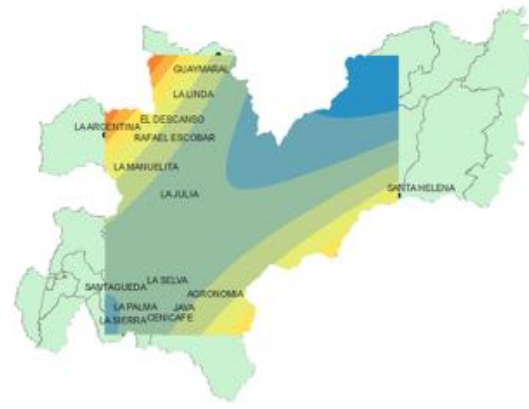
Método Kriging Ordinario año 1997



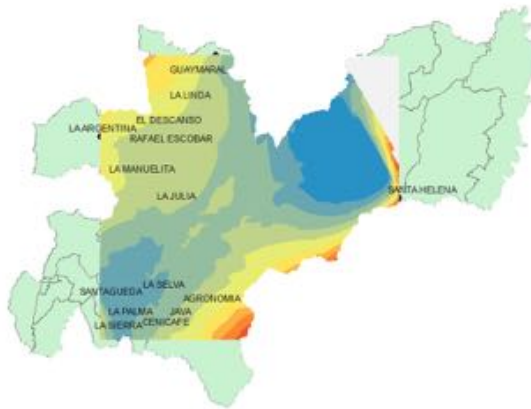
Método Kriging Universal año 1997



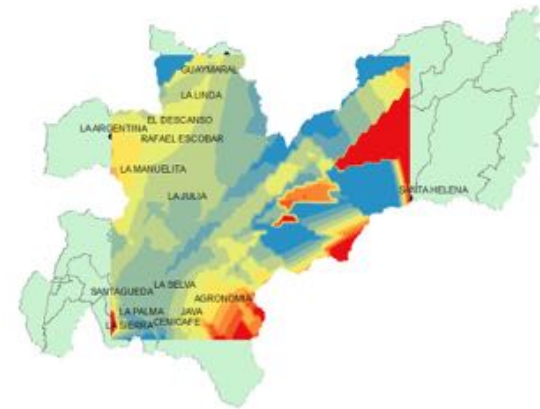
Método IDW año 1998



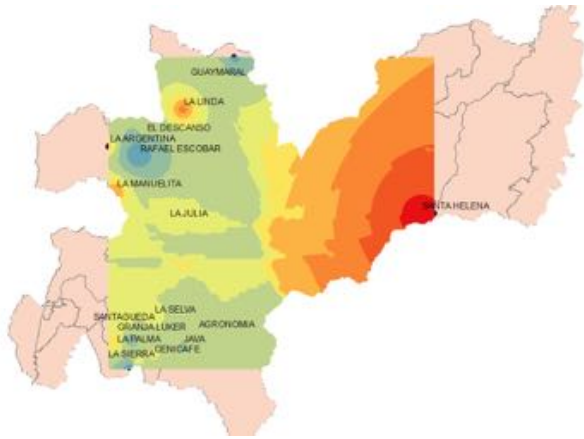
Método GPI año 1998



Método Kriging Ordinario año 1998



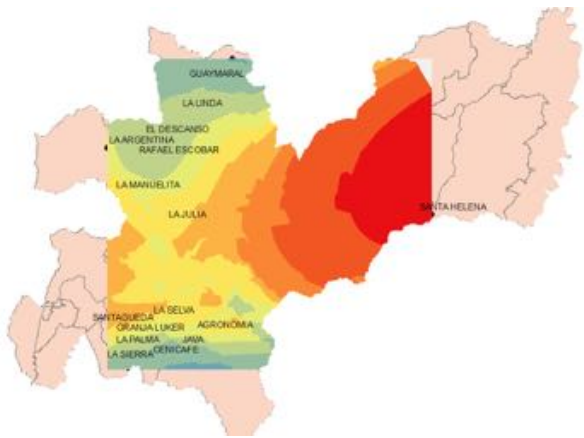
Método Kriging Universal año 1998



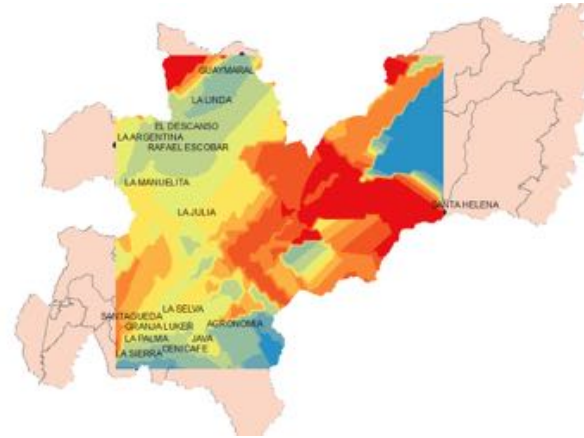
Método IDW año 1999



Método GPI año 1999



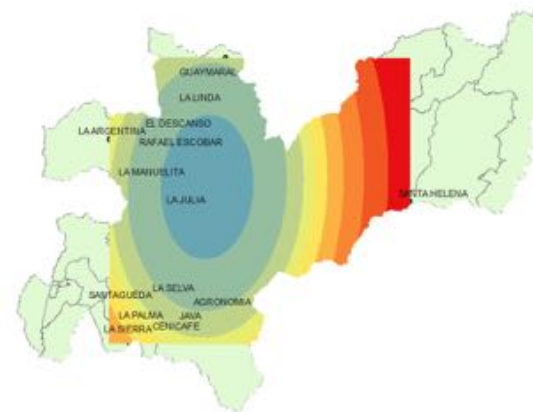
Método Kriging Ordinario año 1999



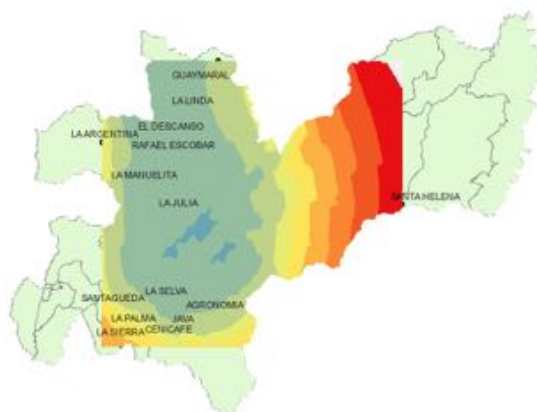
Método Kriging Universal año 1999



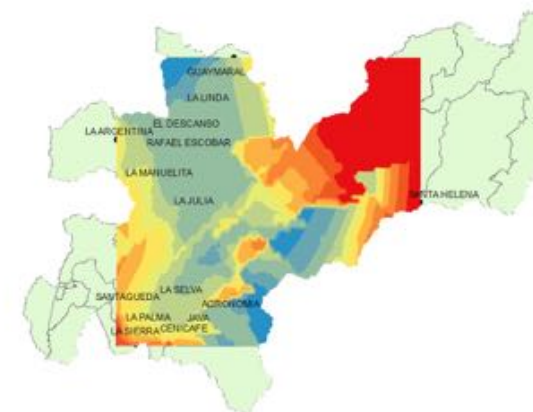
Método IDW año 2001



Método GPI año 2001



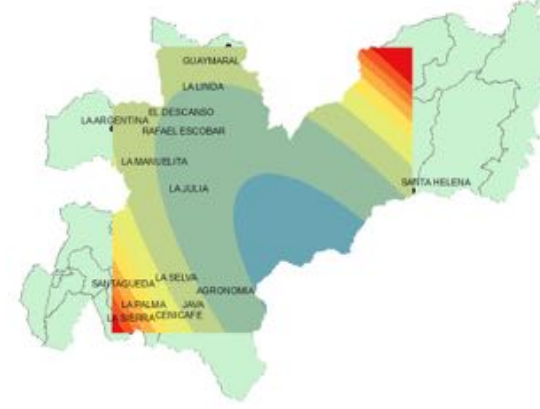
Método Kriging Ordinario año 2001



Método Kriging Universal año 2001



Método IDW año 2002



Método GPI año 2002



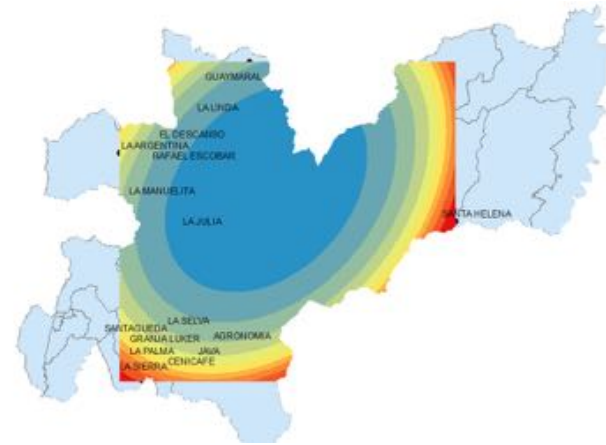
Método Kriging Ordinario año 2002



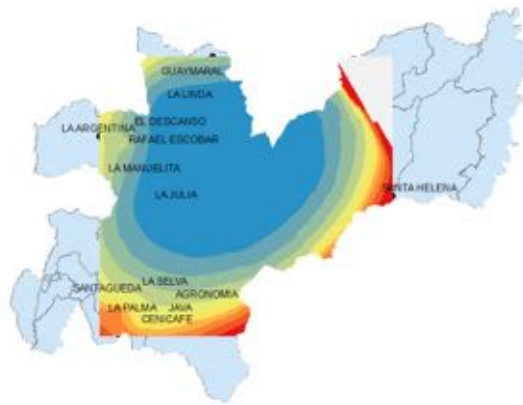
Método Kriging Universal año 2002



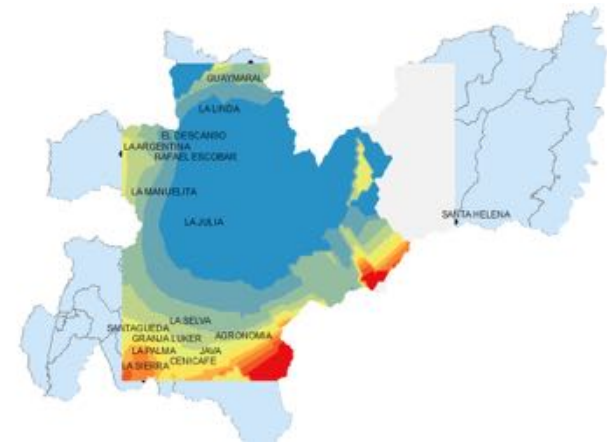
Método IDW año 2003



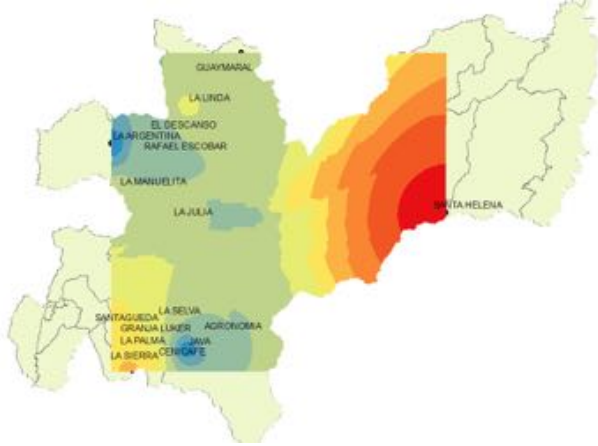
Método GPI año 2003



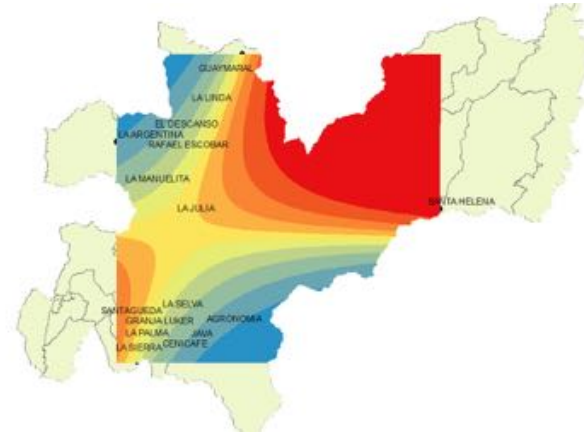
Método Kriging Ordinario año 2003



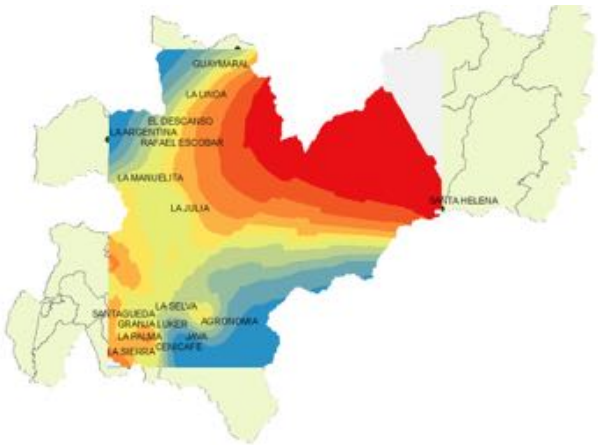
Método Kriging Universal año 2003



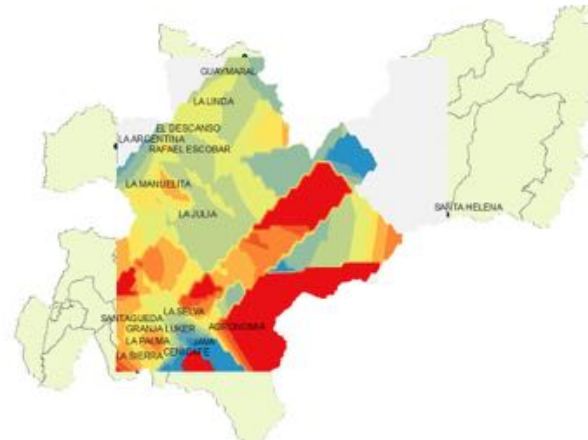
Método IDW año 2004



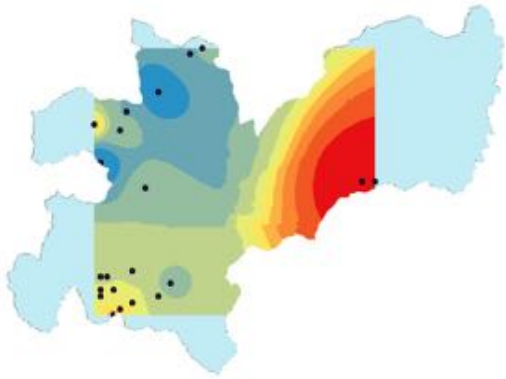
Método GPI año 2004



Método Kriging Ordinario año 2004



Método Kriging Universal año 2004



Método IDW año 2005



Método GPI año 2005



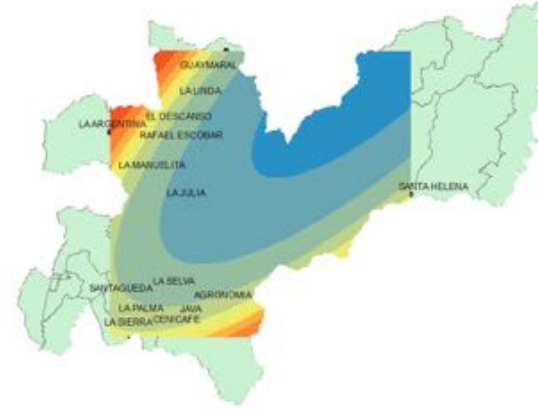
Método Kriging Ordinario año 2005



Método Kriging Universal año 2005



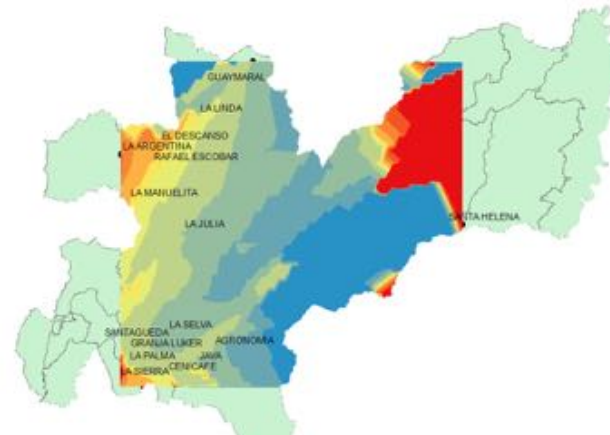
Método IDW año 2007



Método GPI año 2007



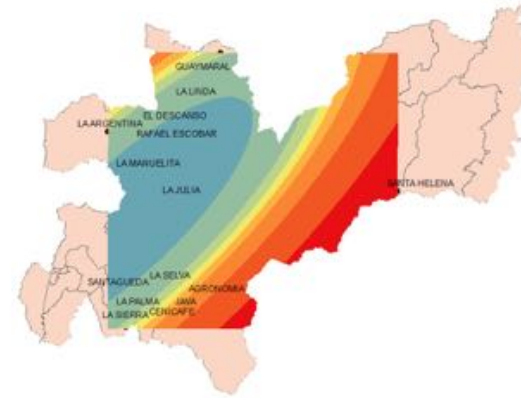
Método Kriging Ordinario año 2007



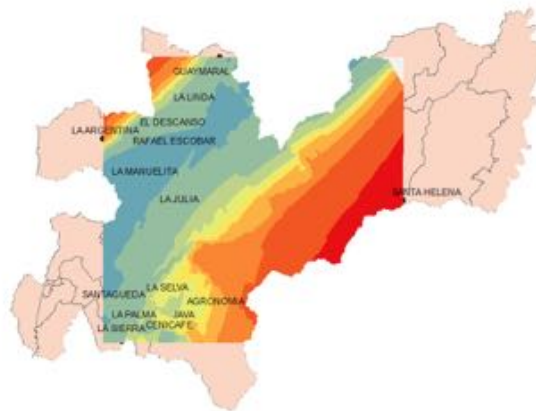
Método Kriging Universal año 2007



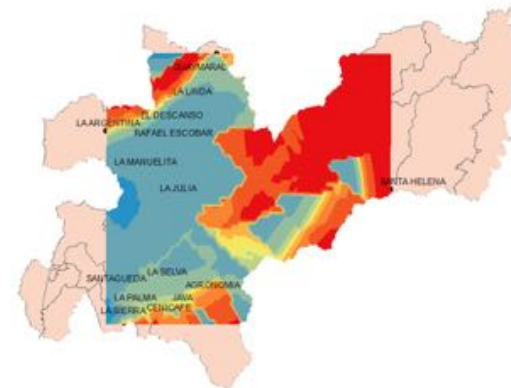
Método IDW año 2008



Método GPI año 2008



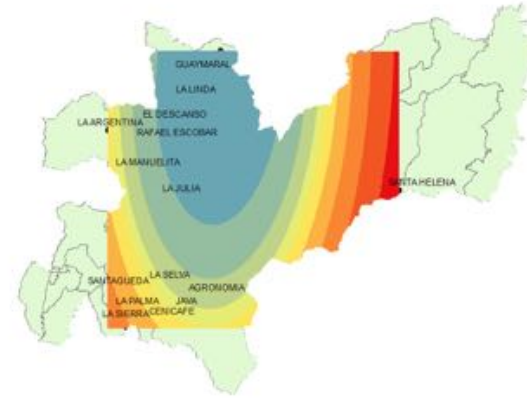
Método Kriging Ordinario año 2008



Método Kriging Universal año 2008



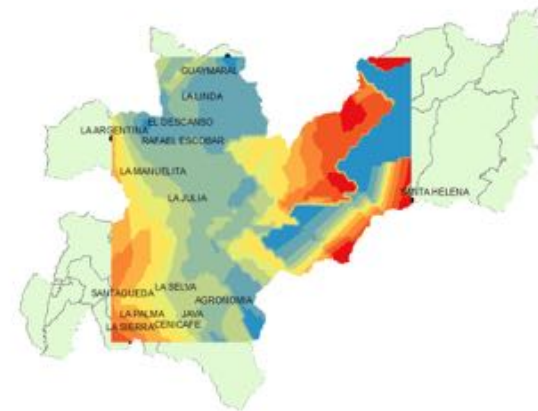
Método IDW año 2009



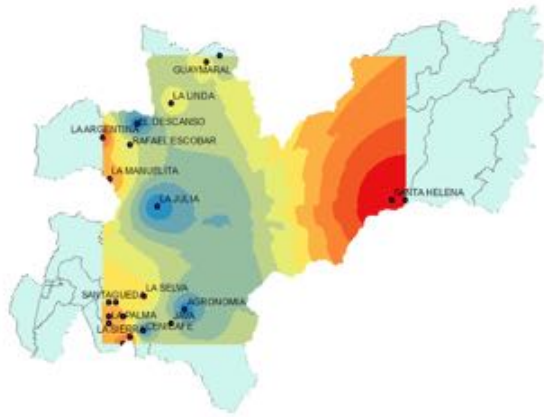
Método GPI año 2008



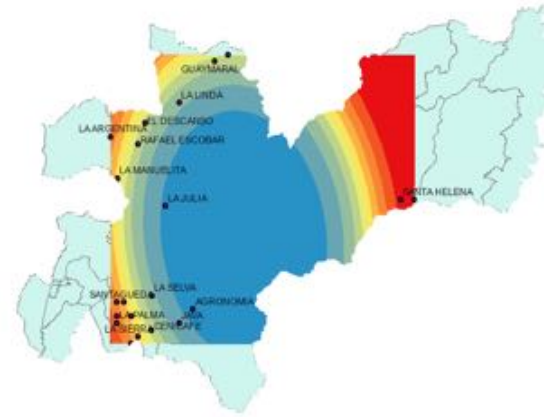
Método Kriging Ordinario año 2009



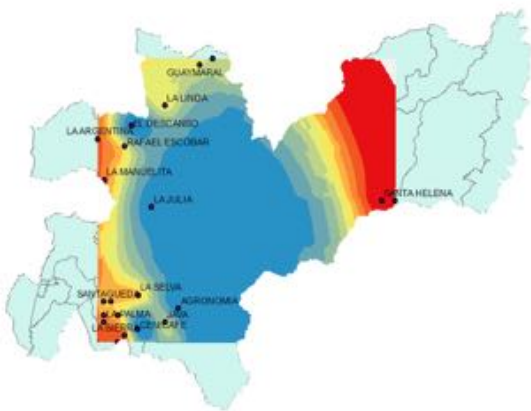
Método Kriging Universal año 2009



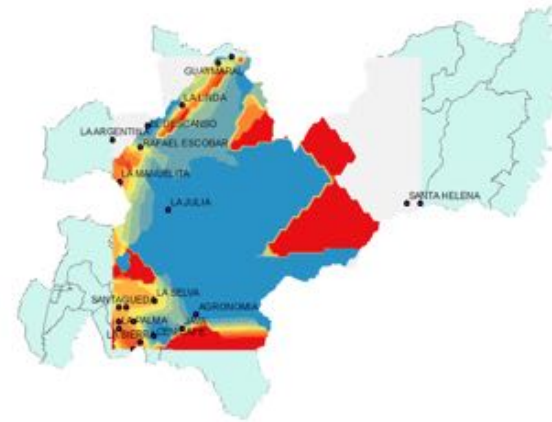
Método IDW año 2013



Método GPI año 2013



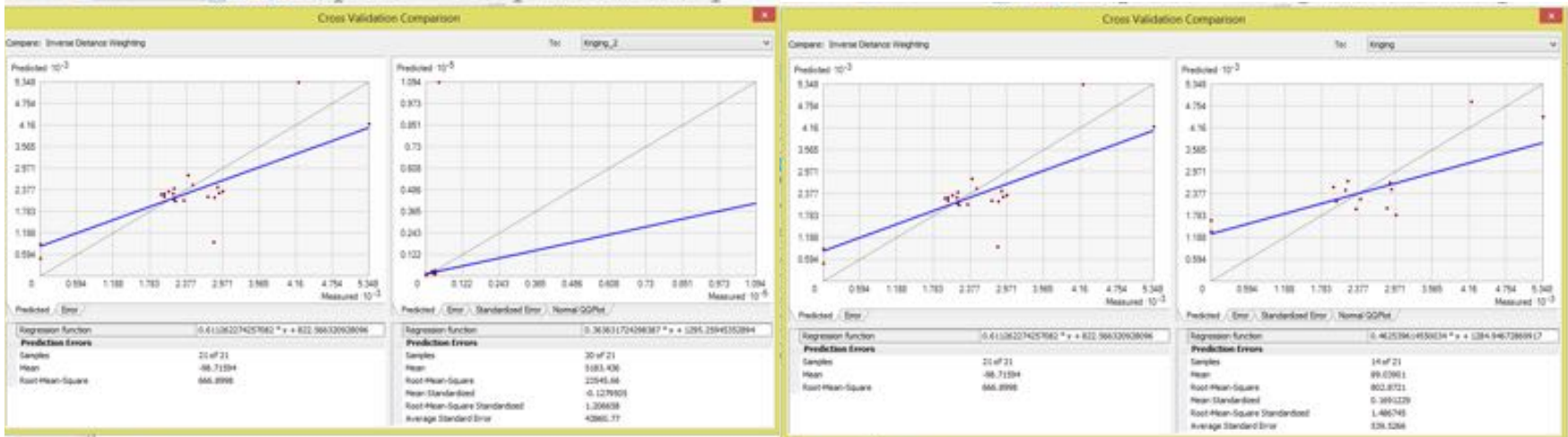
Método Kriging Ordinario año 2013



Método Kriging Universal año 2013

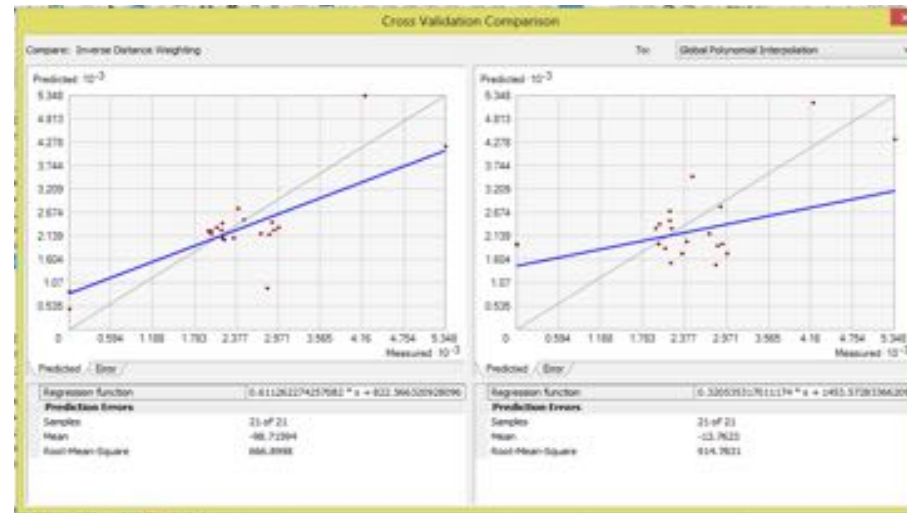
ANEXO F

Validaciones cruzadas, métodos de interpolación

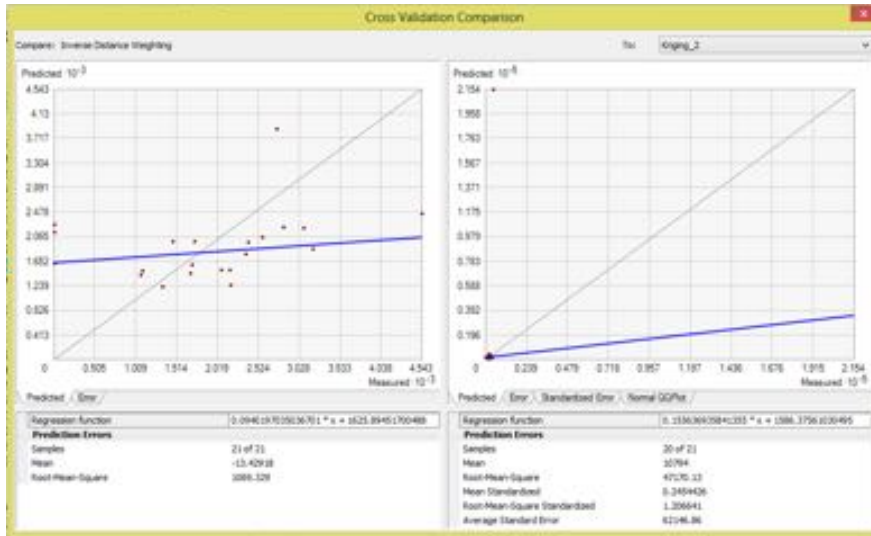


IDW VS. Kriging Universal, año 1988

IDW Vs. Kriging Ordinario, año 1988



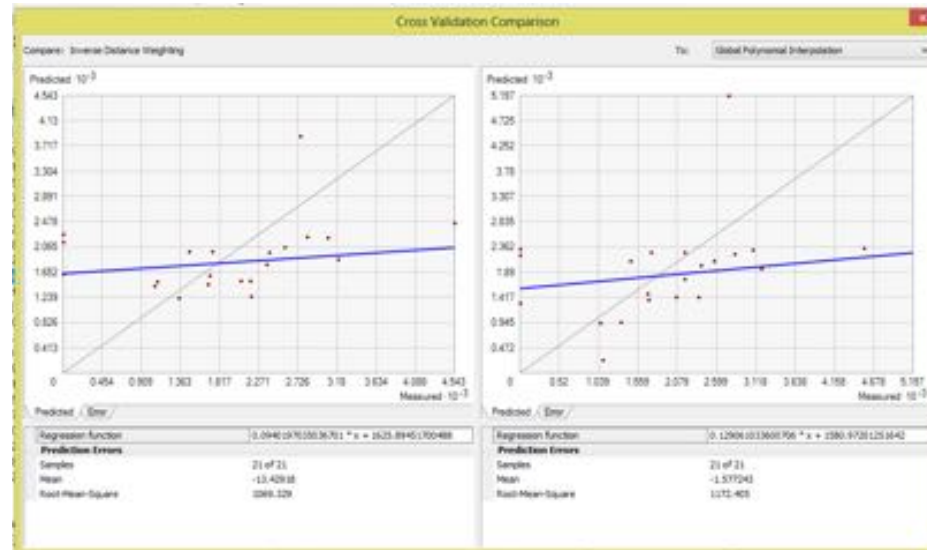
IDW vs. GPI, año 1988



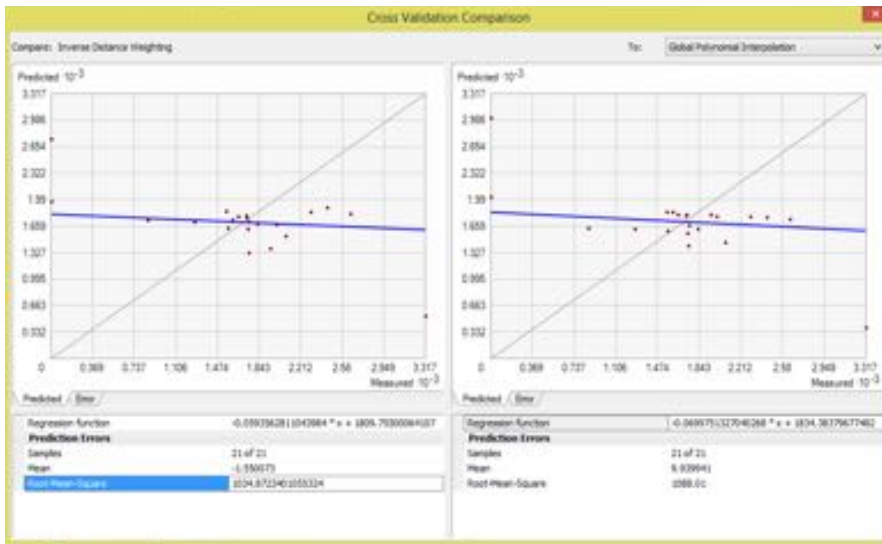
IDW VS. Kriging Universal, año 1989



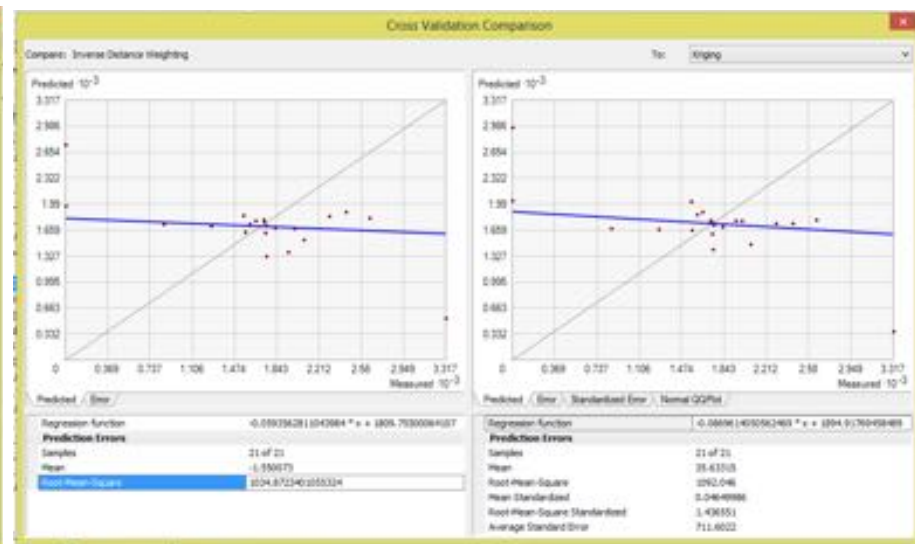
IDW Vs. Kriging Ordinario, año 1989



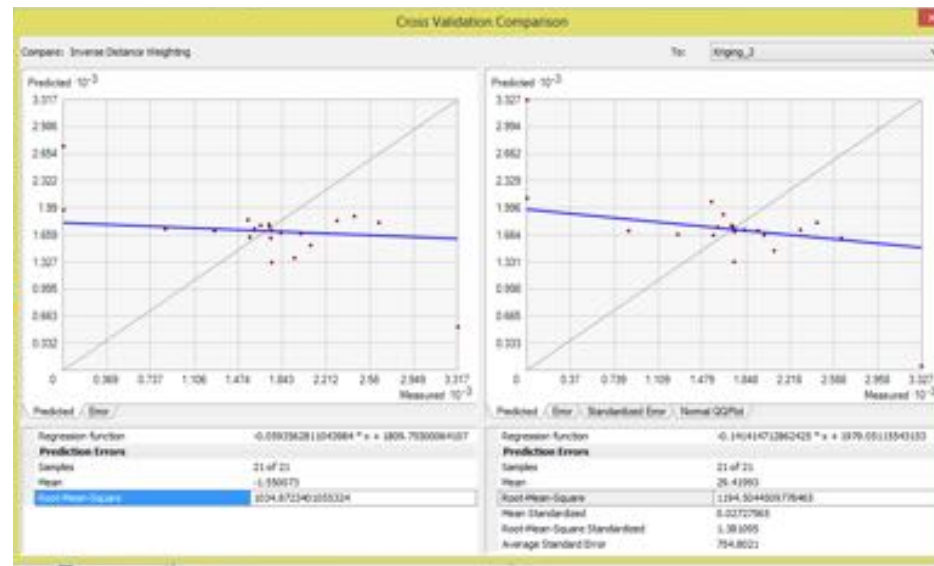
IDW vs. GPI 1989



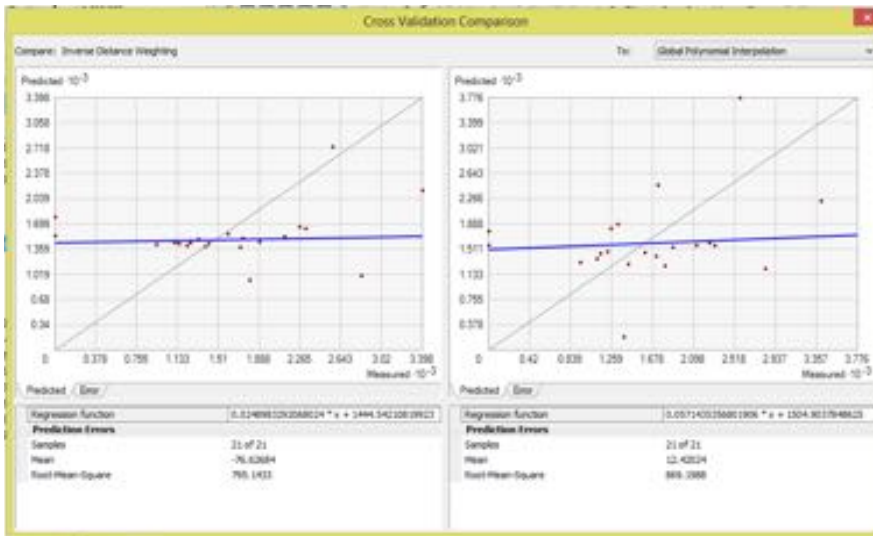
IDW Vs. GPI, año 1990



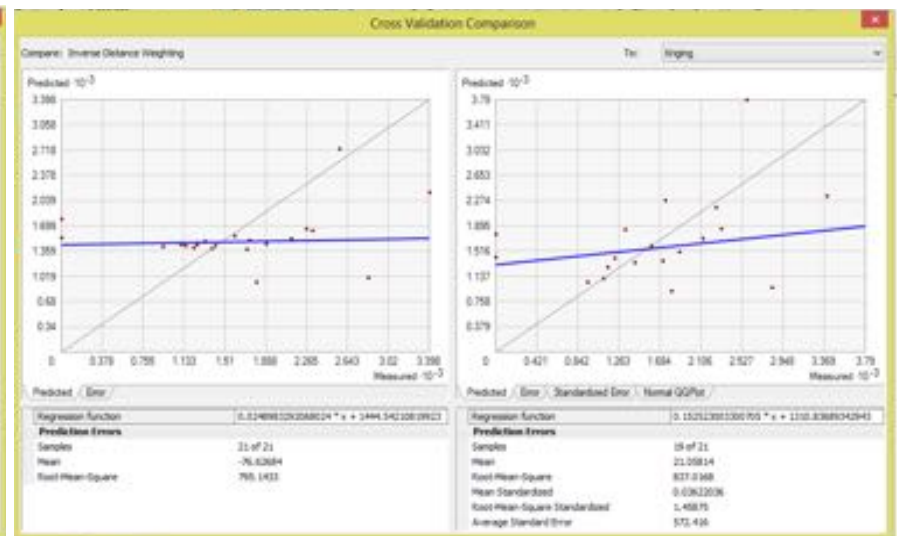
IDW Vs. Kriging Ordinario, año 1990



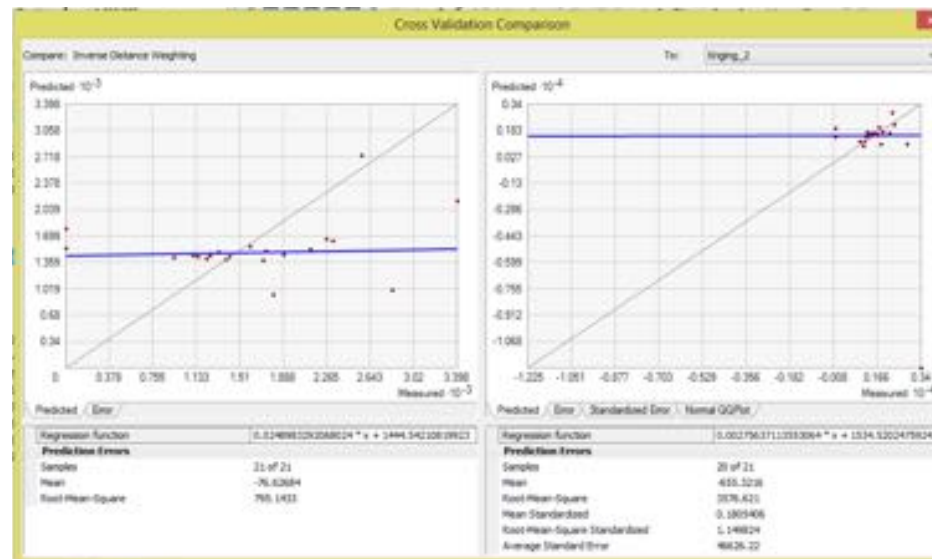
IDW VS. Kriging Universal, año 1990



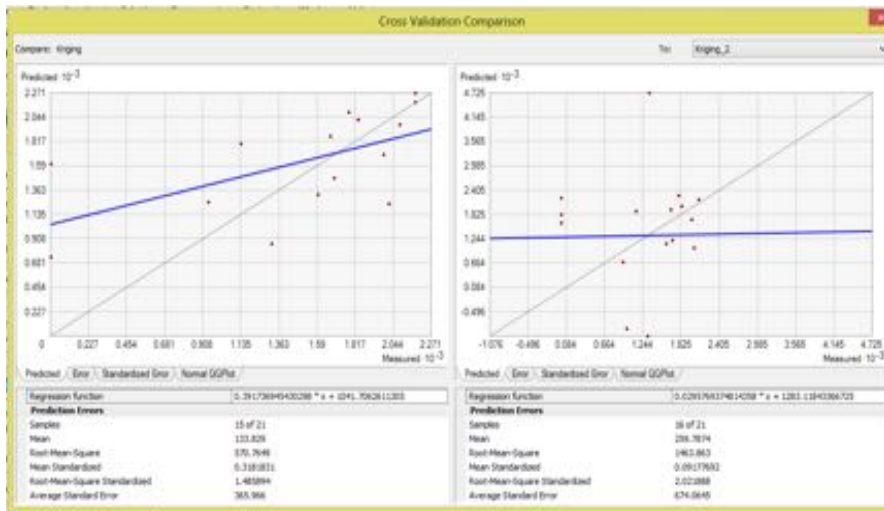
IDW vs. GPI año 1991



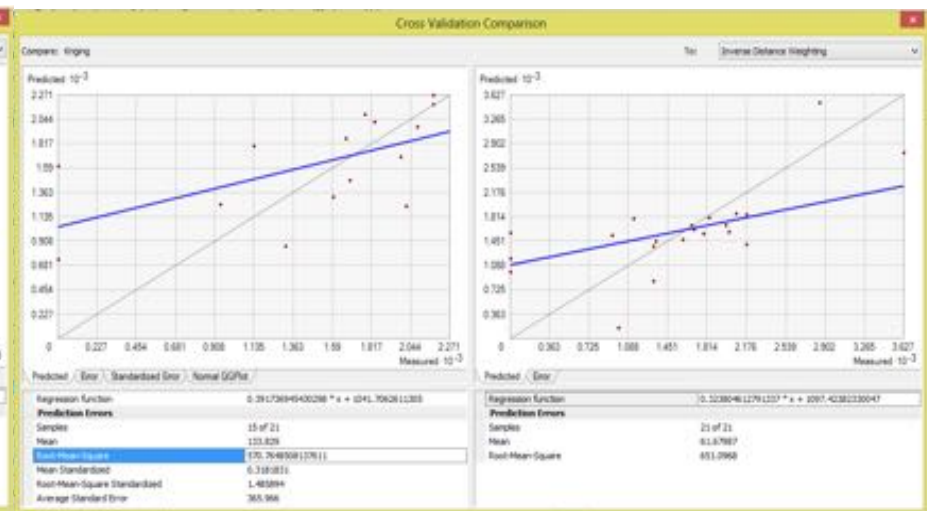
IDW Vs. Kriging Ordinario, año 1991



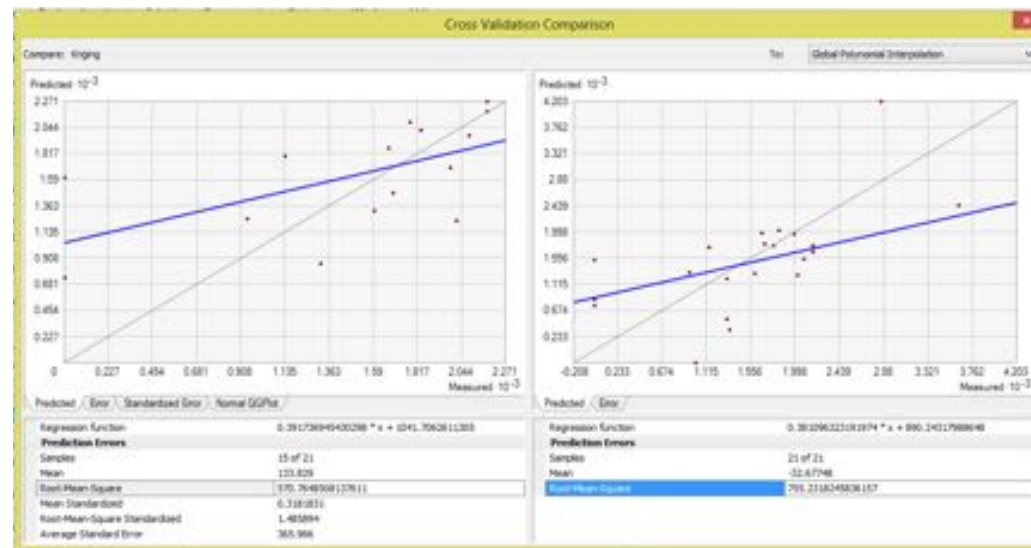
IDW VS. Kriging Universal, año 1991



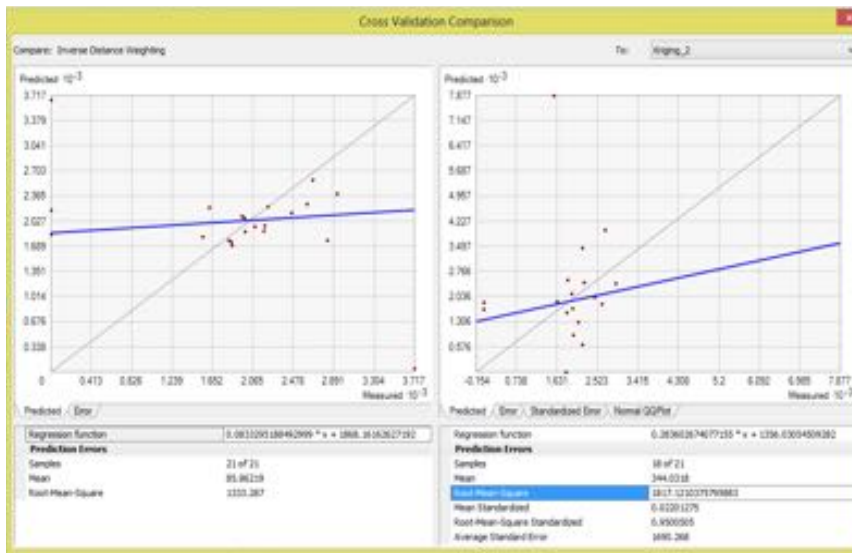
Kriging Ordinario Vs. Kriging Universal, año 1992



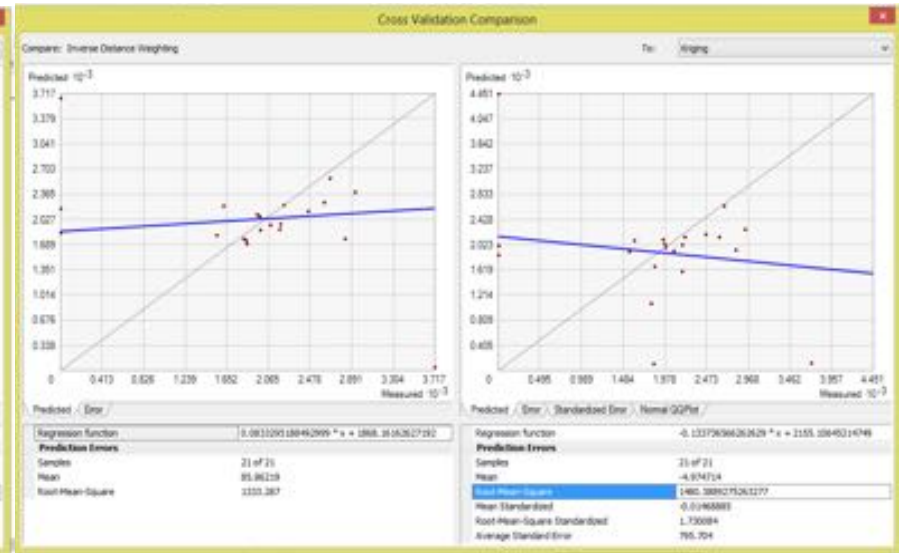
Kriging Ordinario Vs IDW, año 1992



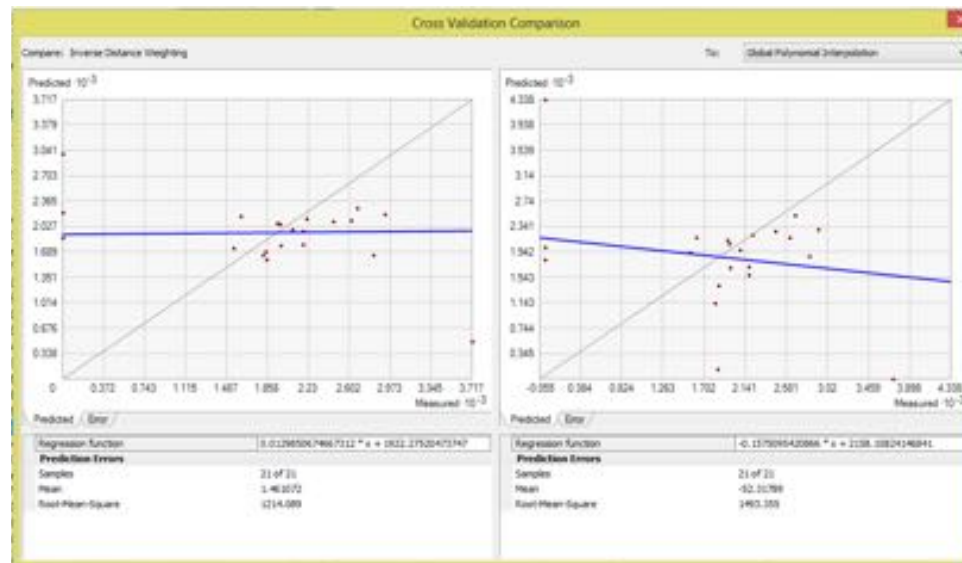
Kriging Ordinario Vs GPI, año 1992



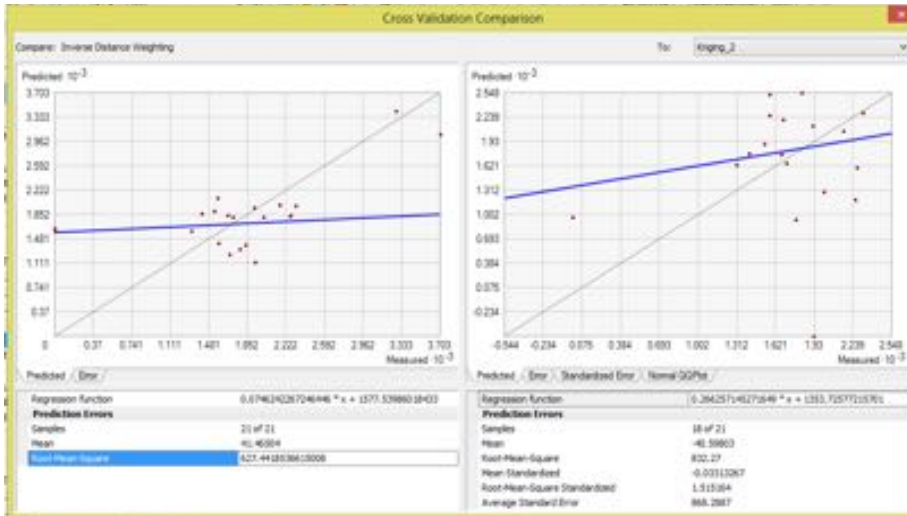
IDW VS. Kriging Universal, año 1993



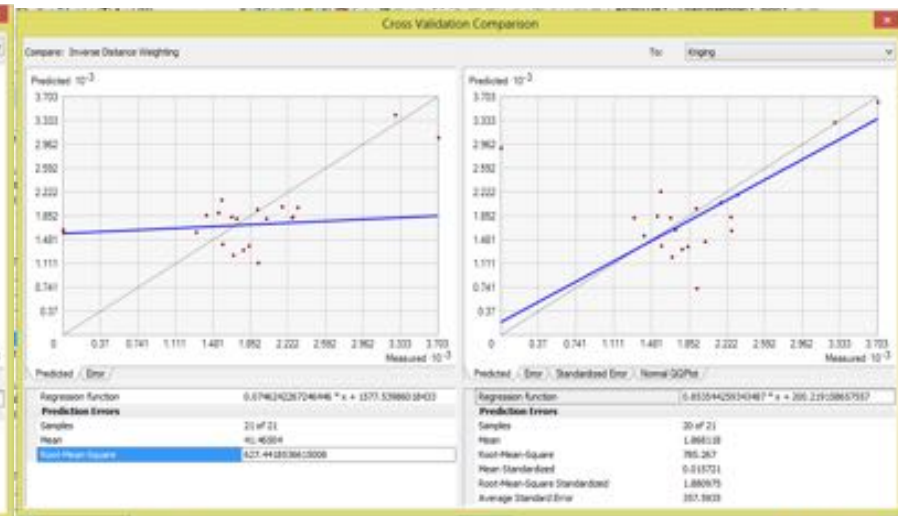
IDW Vs. Kriging Ordinario, año 1993



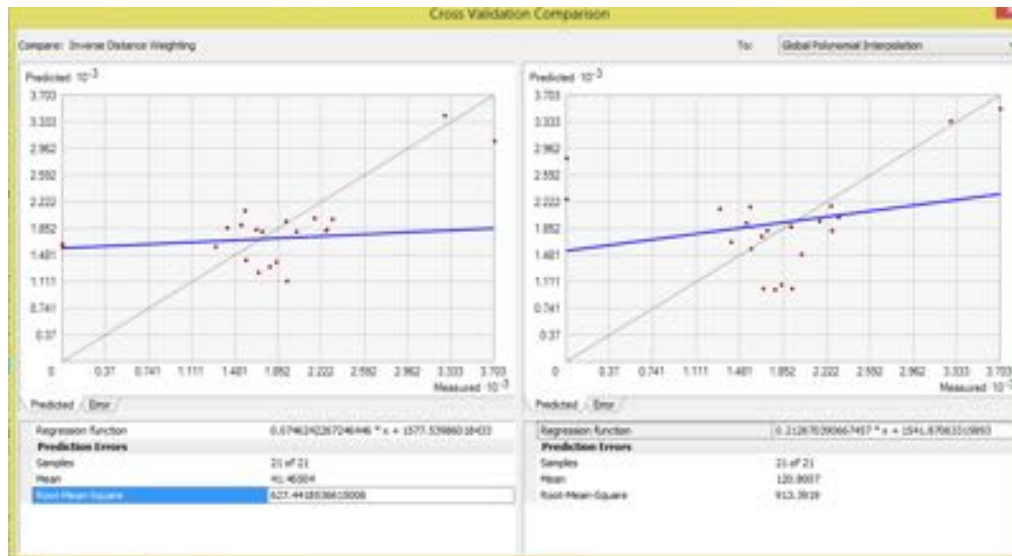
IDW vs. GPI 1993



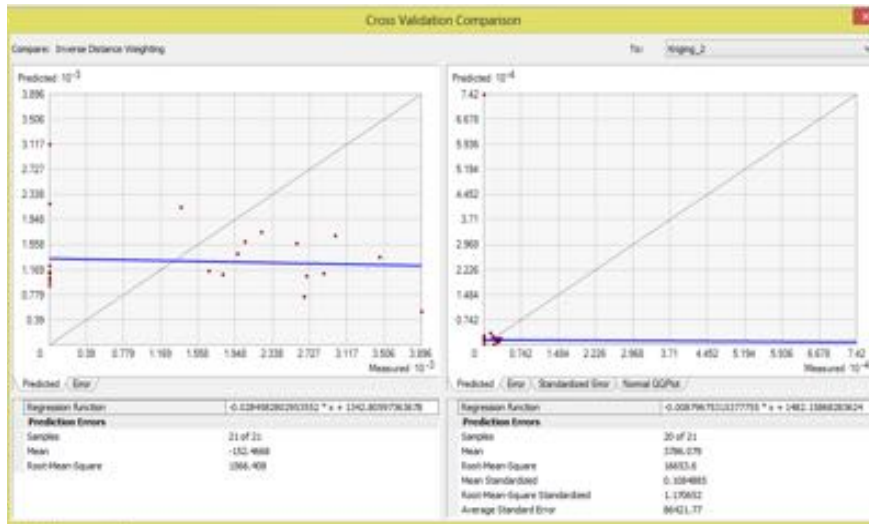
IDW Vs. Kriging Universal, año 1997



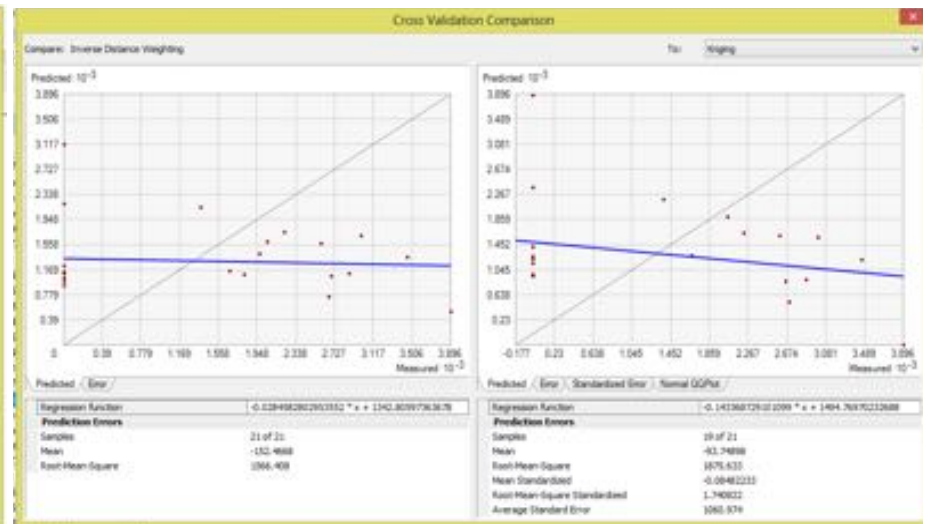
IDW vs. Kriging ordinario, año 1997



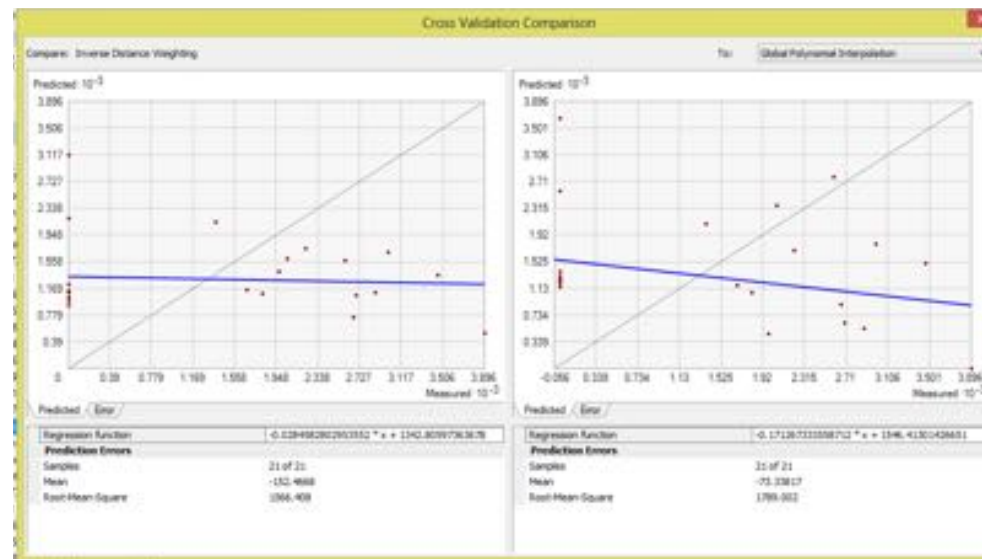
IDW VS. GPI, año 1997



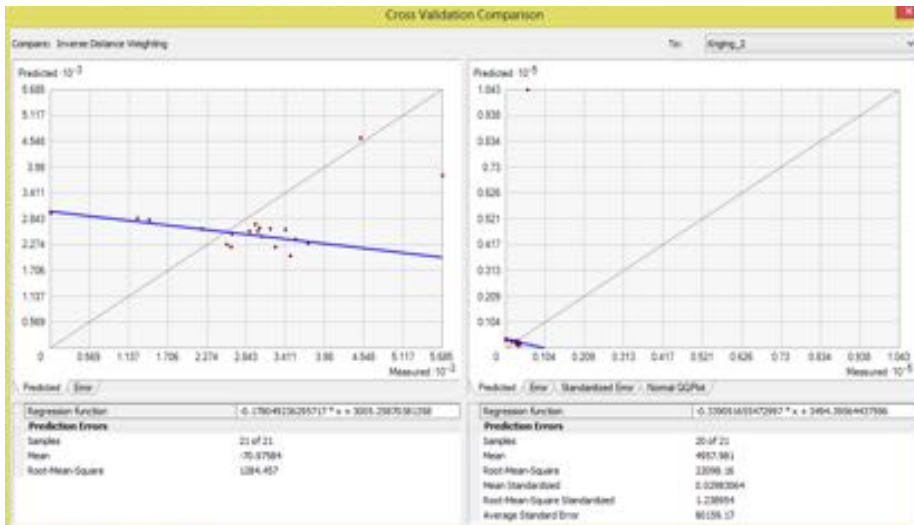
IDW vs. K Universal, año 1998



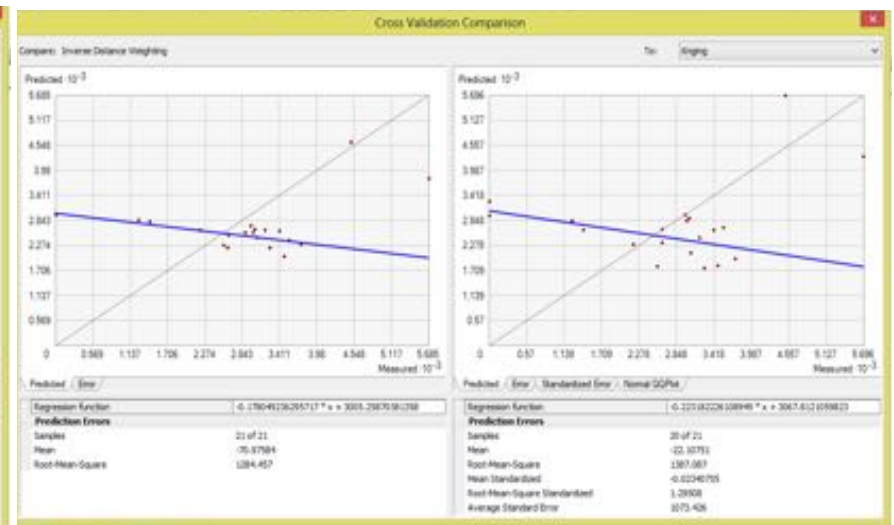
IDW vs. K. Ordinario, año 1998



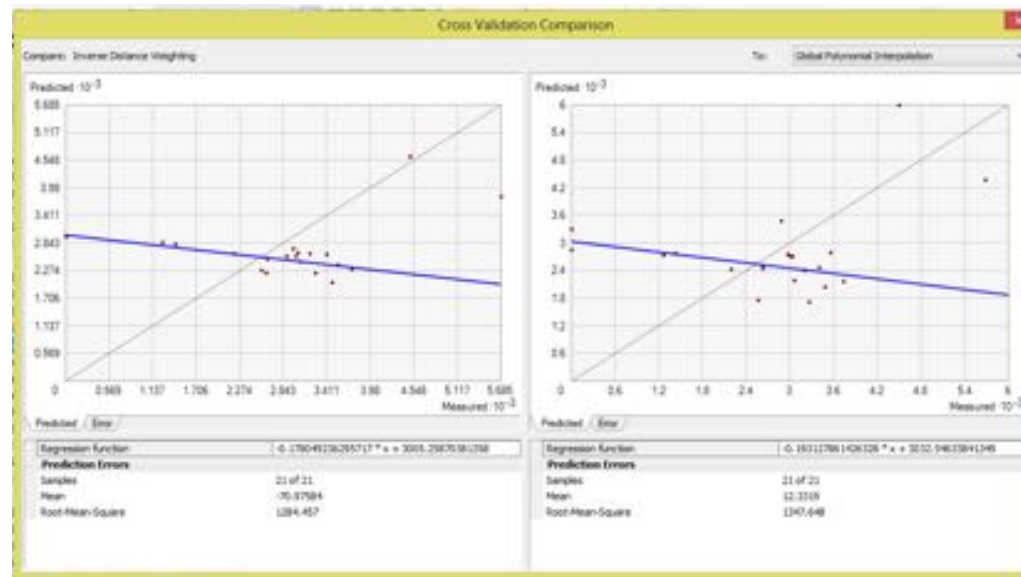
IDW vs. GPI, año 1998



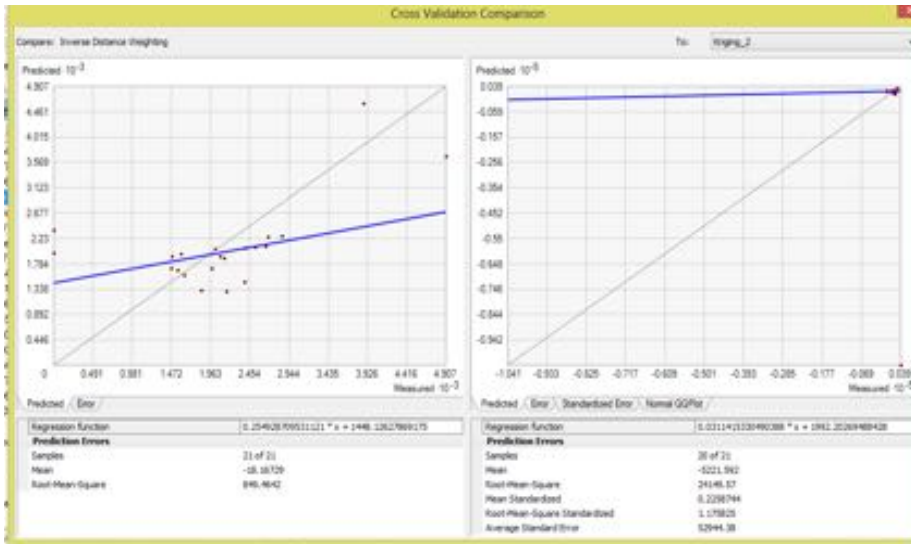
IDW vs K. Universal, año 1999



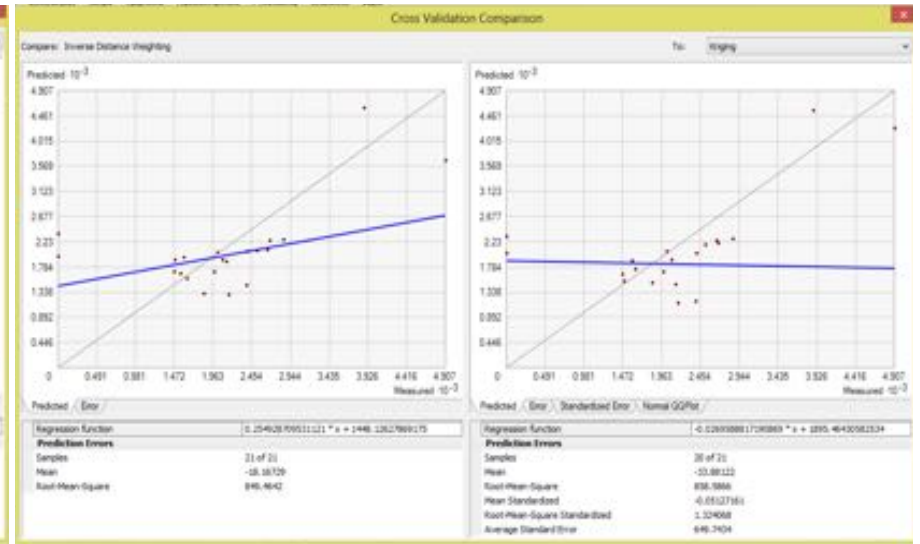
IDW vs. K Ordinario, año 1999



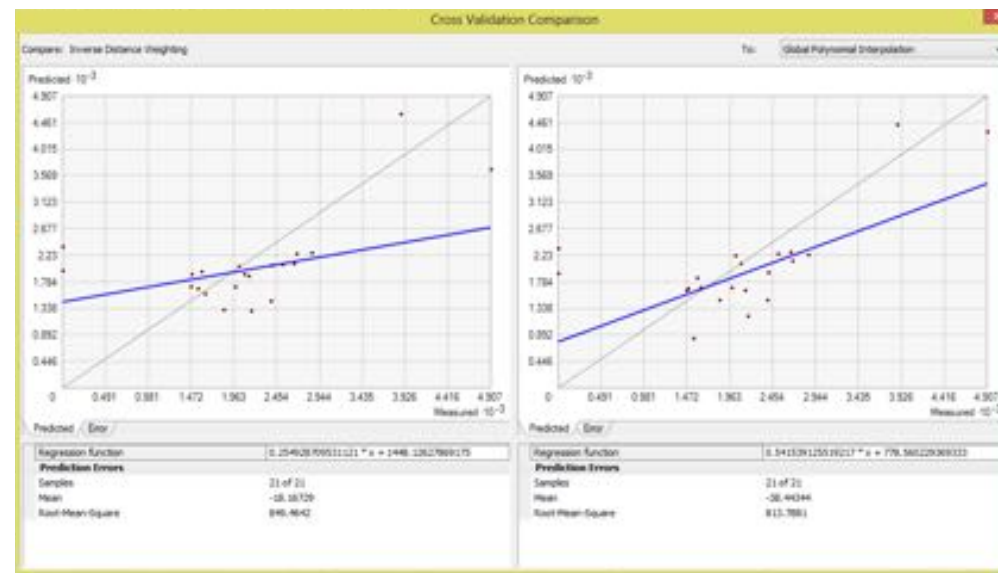
IDW vs. GPI, año 1999



IDW vs. K Universal, año 2001



IDW vs. K. Ordinario, año 2001



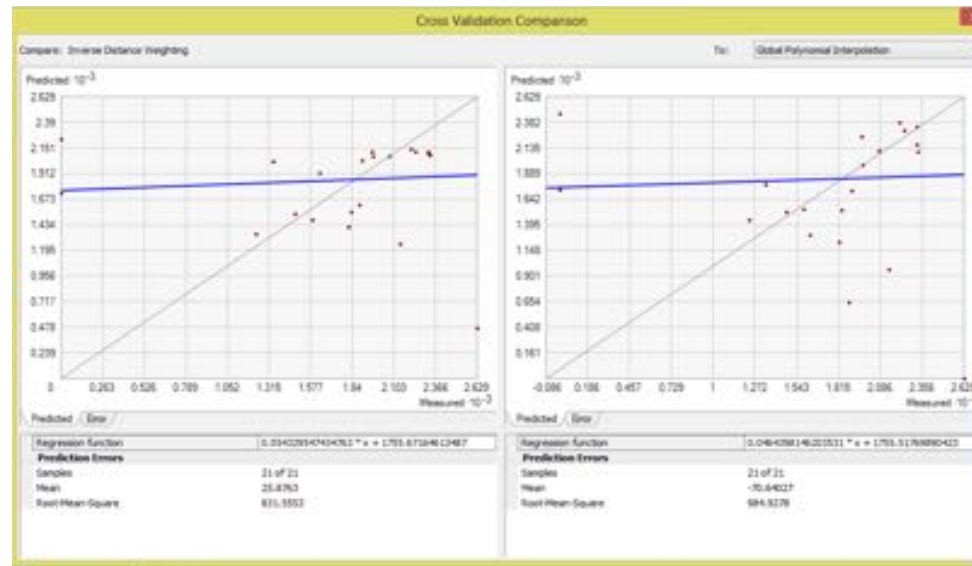
IDW vs. GPI, año 2001



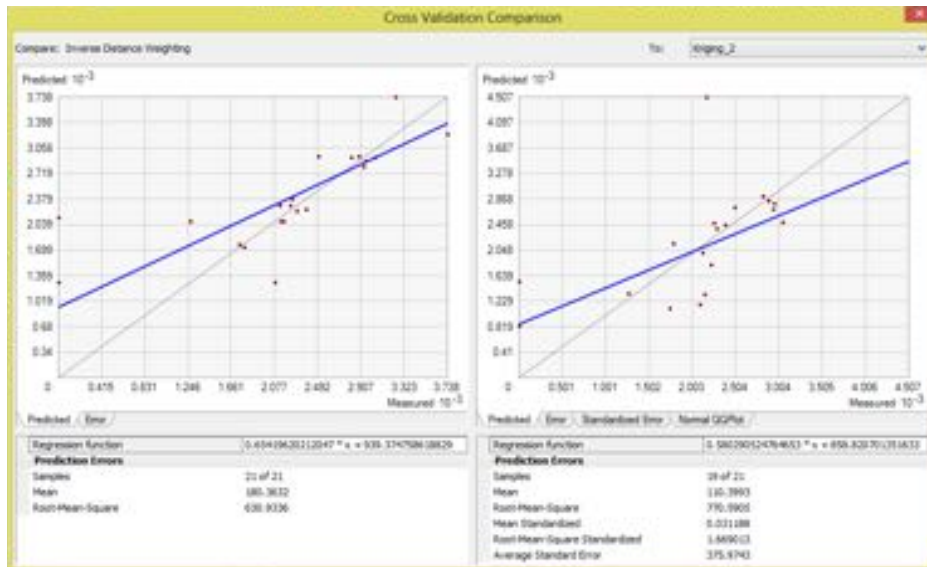
IDW vs. K. Universal, año 2002



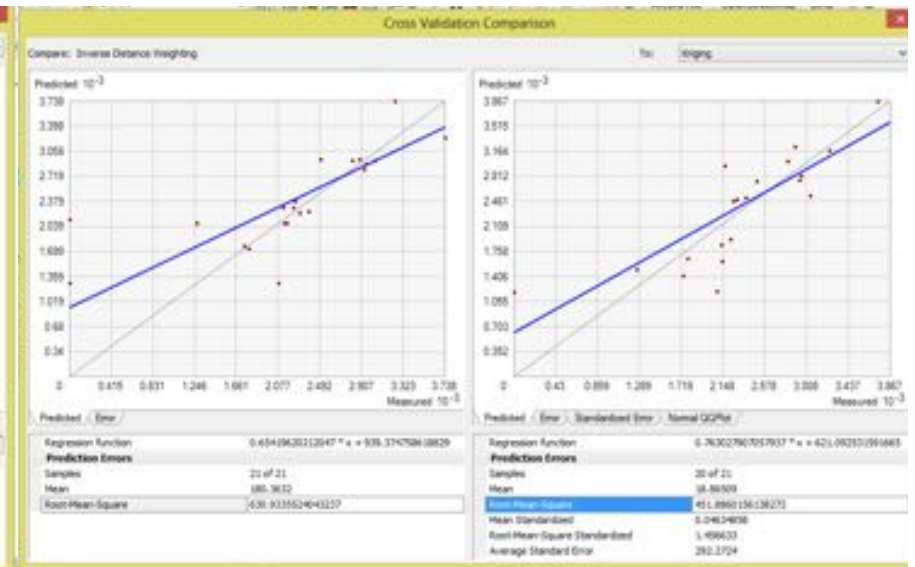
IDW vs. K. Ordinario, año 2002



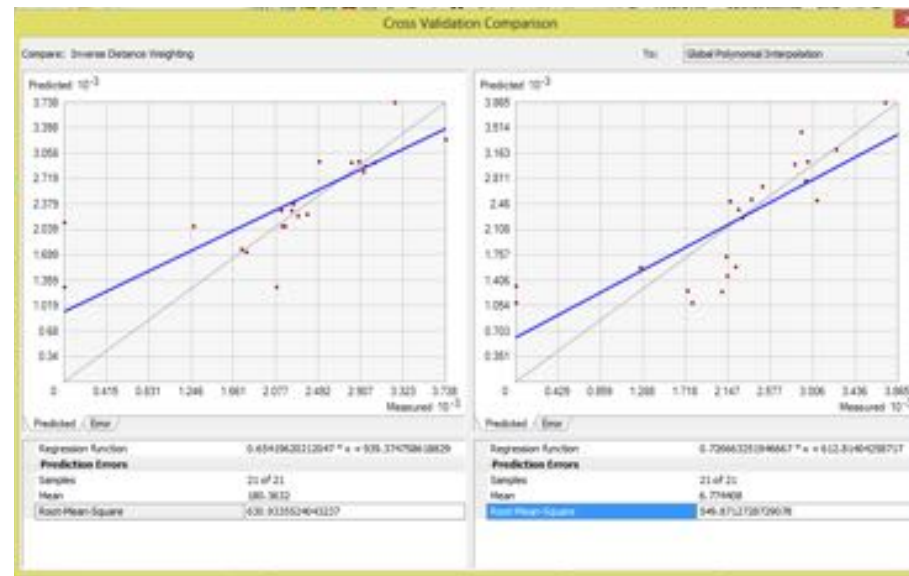
IDW vs. GPI, año 2002



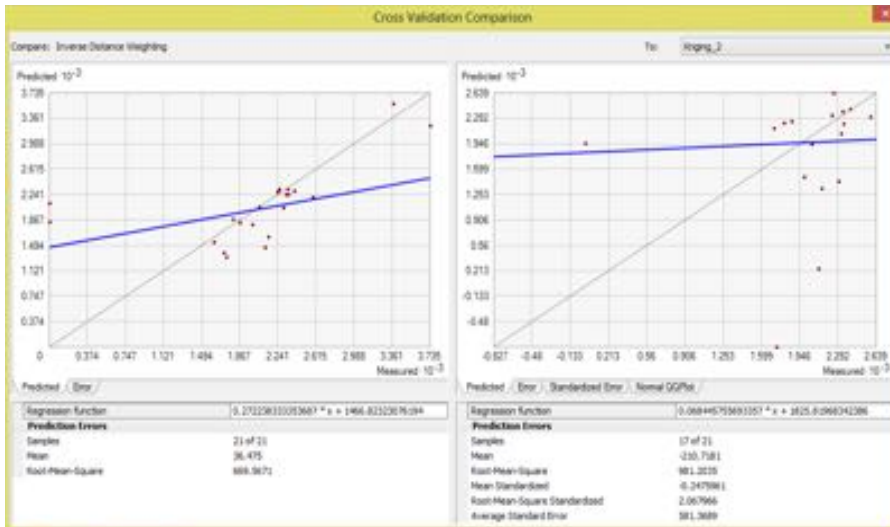
IDW VS. K Universal, año 2003



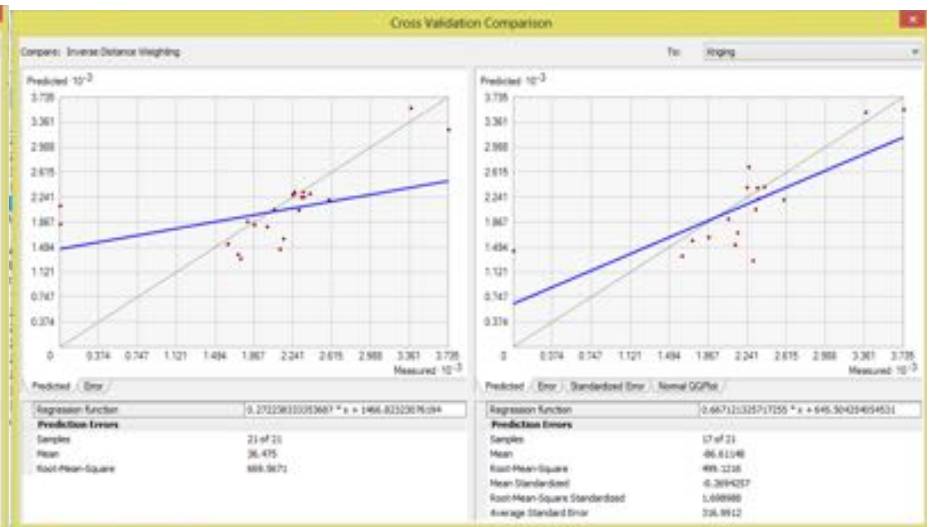
IDW VS. K Ordinario, año 2003



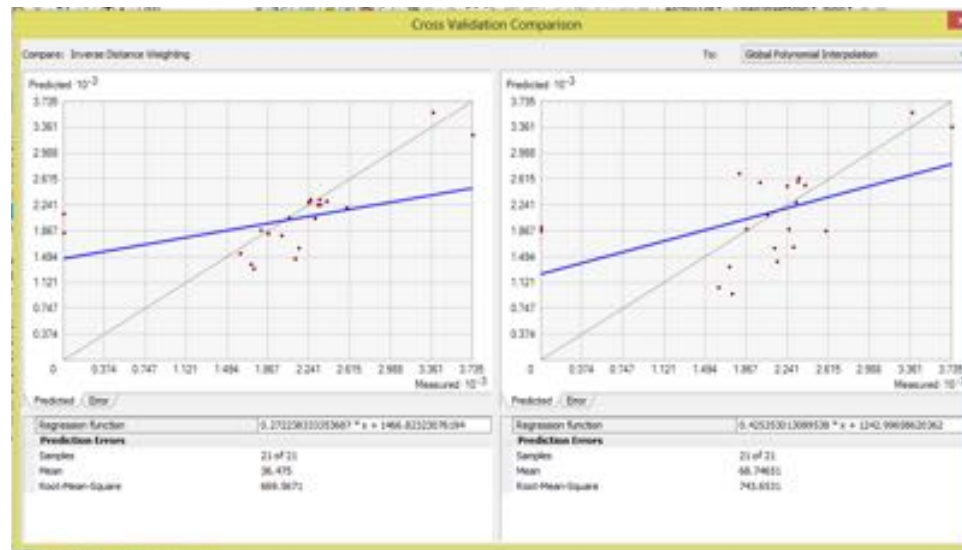
IDW vs. GPI, año 2003



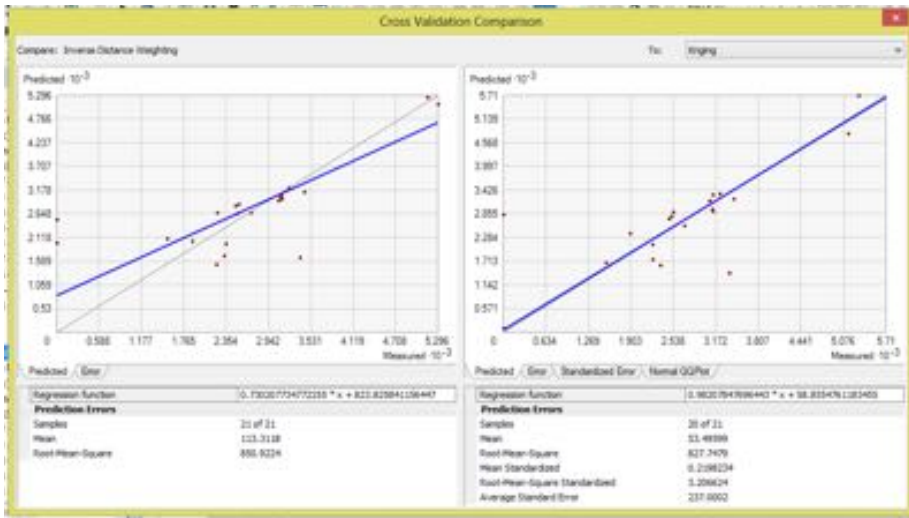
IDW VS. K Universal, año 2004



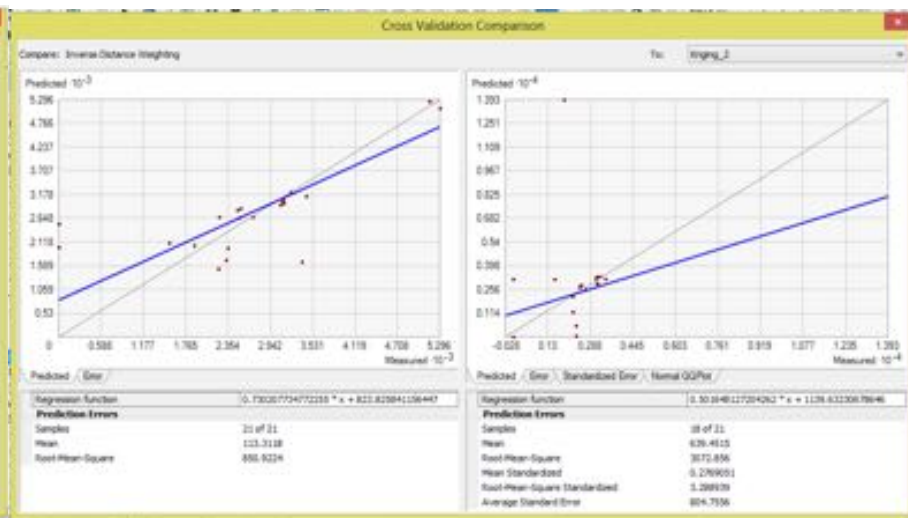
IDW VS. K. Ordinario, año 2004



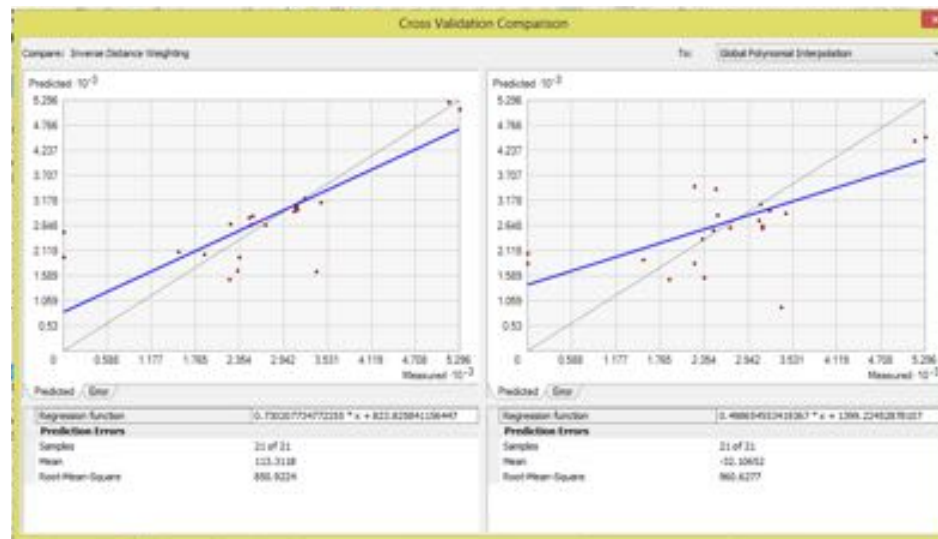
IDW vs. GPI, año 2004



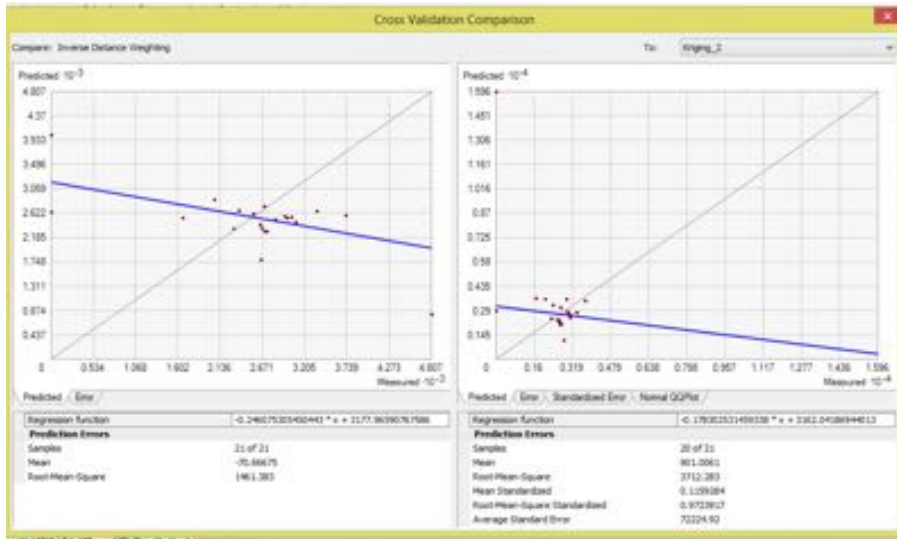
IDW vs. K. Ordinario, año 2005



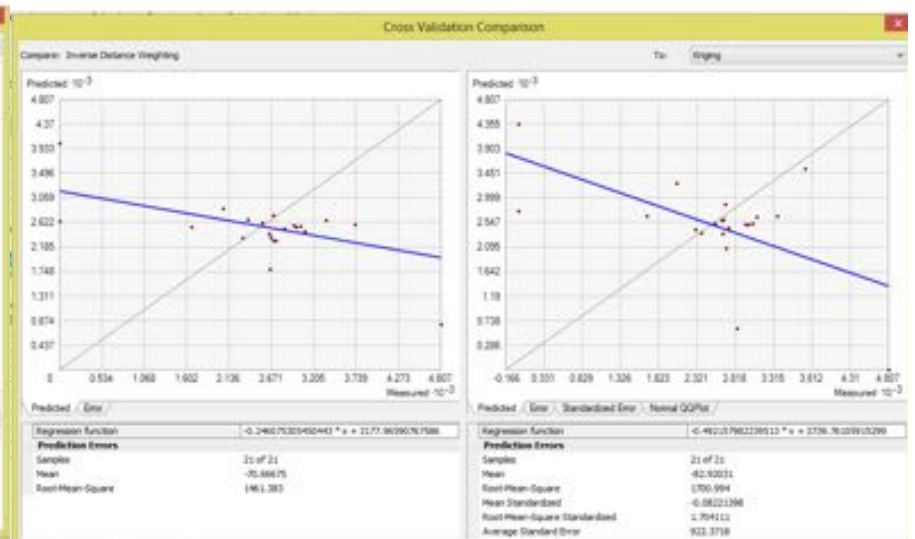
IDW vs. K. Universal, año 2005



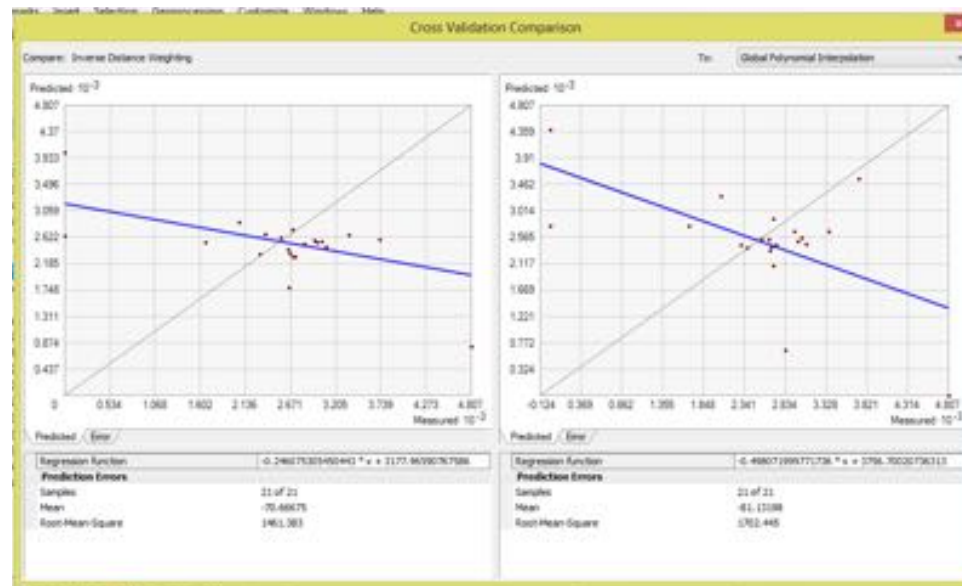
IDW vs. GPI 2005



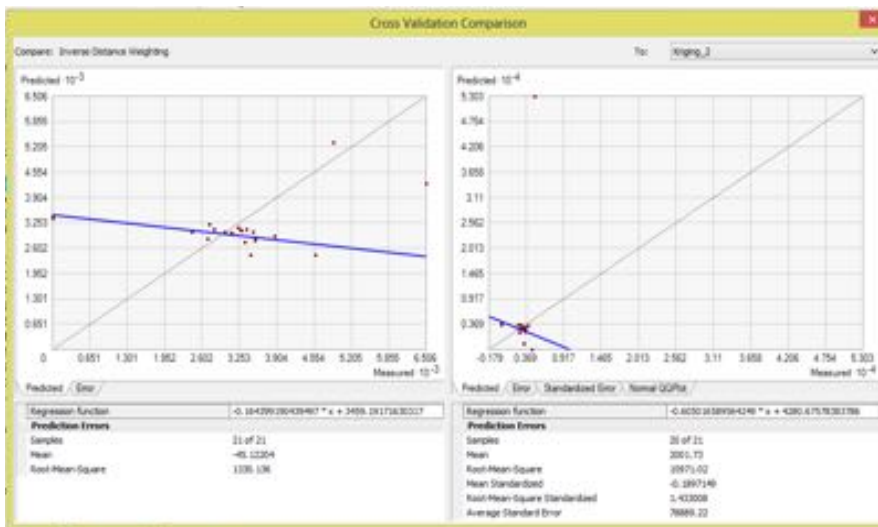
IDW vs. K. Universal, año 2007



IDW vs. K. Ordinario, año 2007



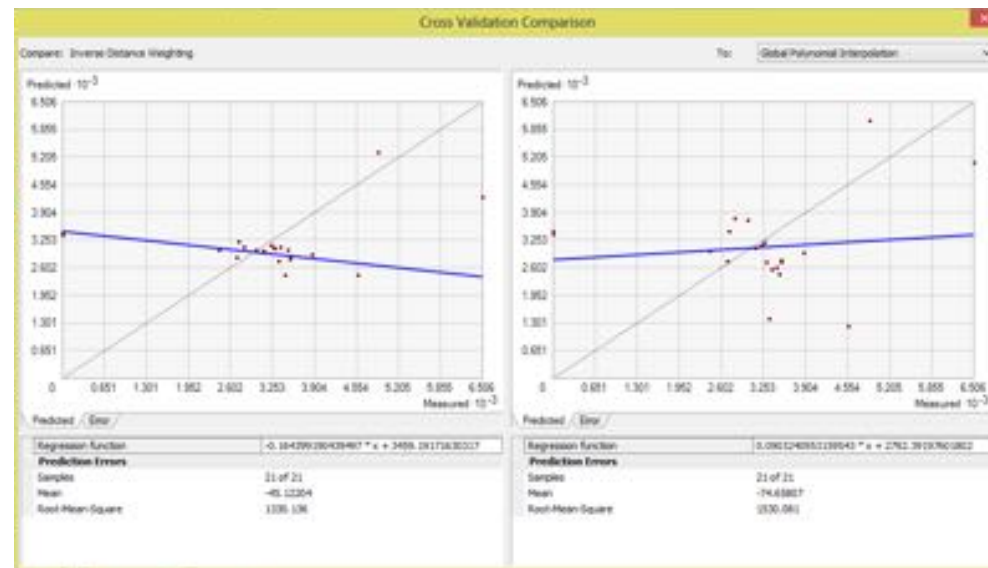
IDW vs. GPI, año 2007



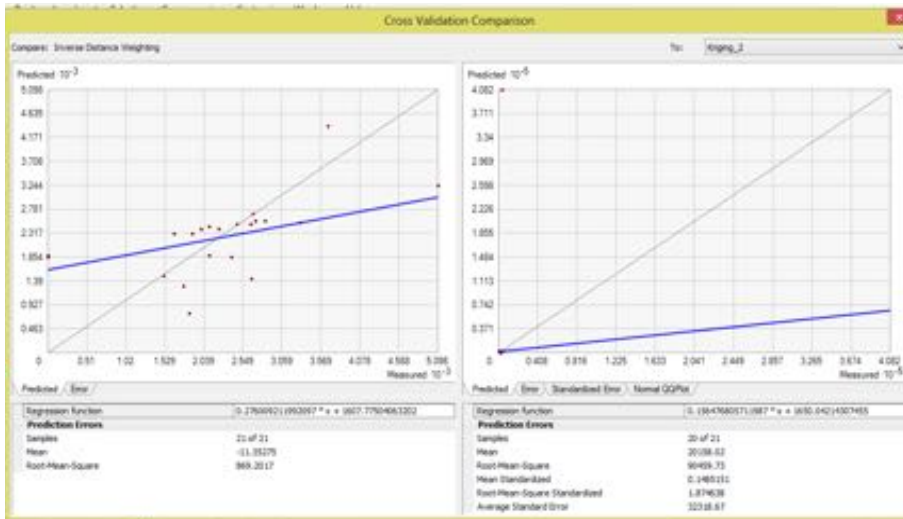
IDW vs. K. Universal, año 2008



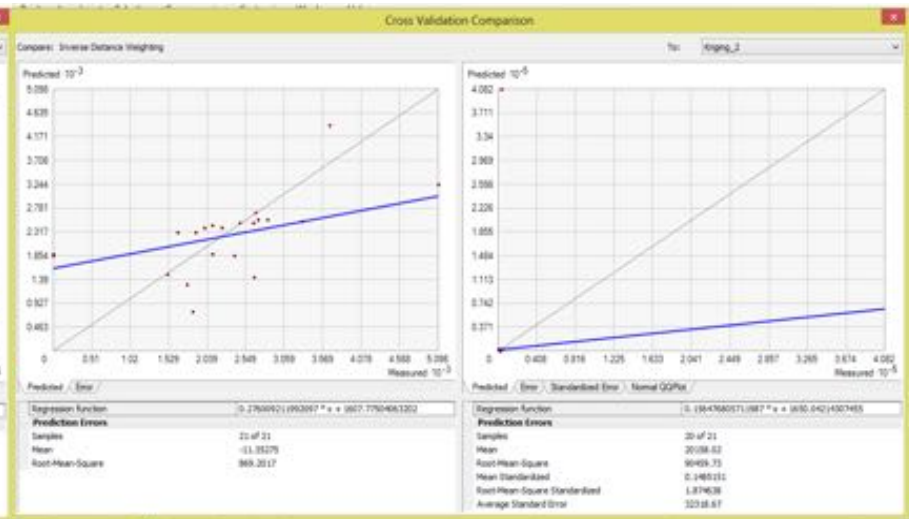
IDW vs. K. Ordinario, año 2008



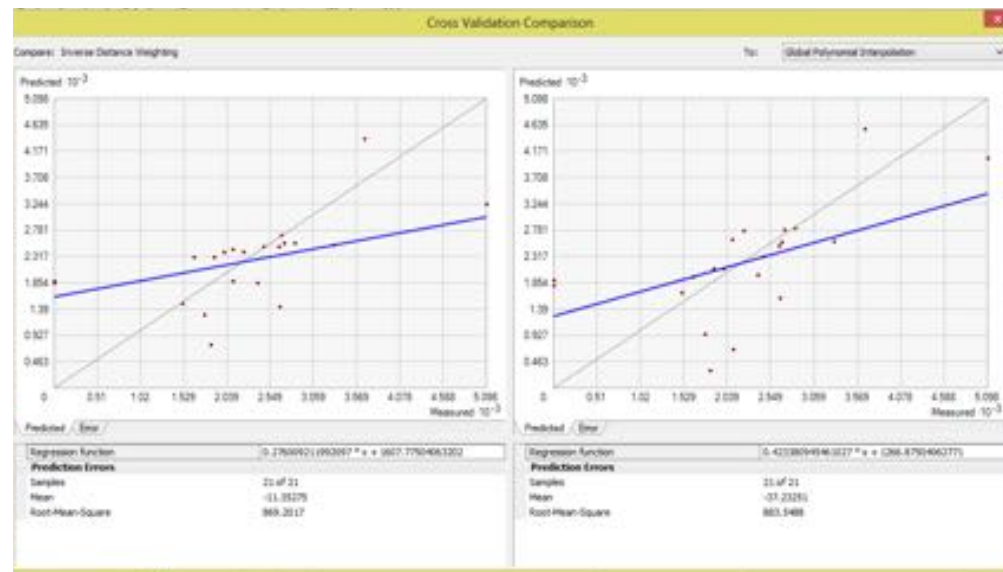
IDW vs. GPI 2008



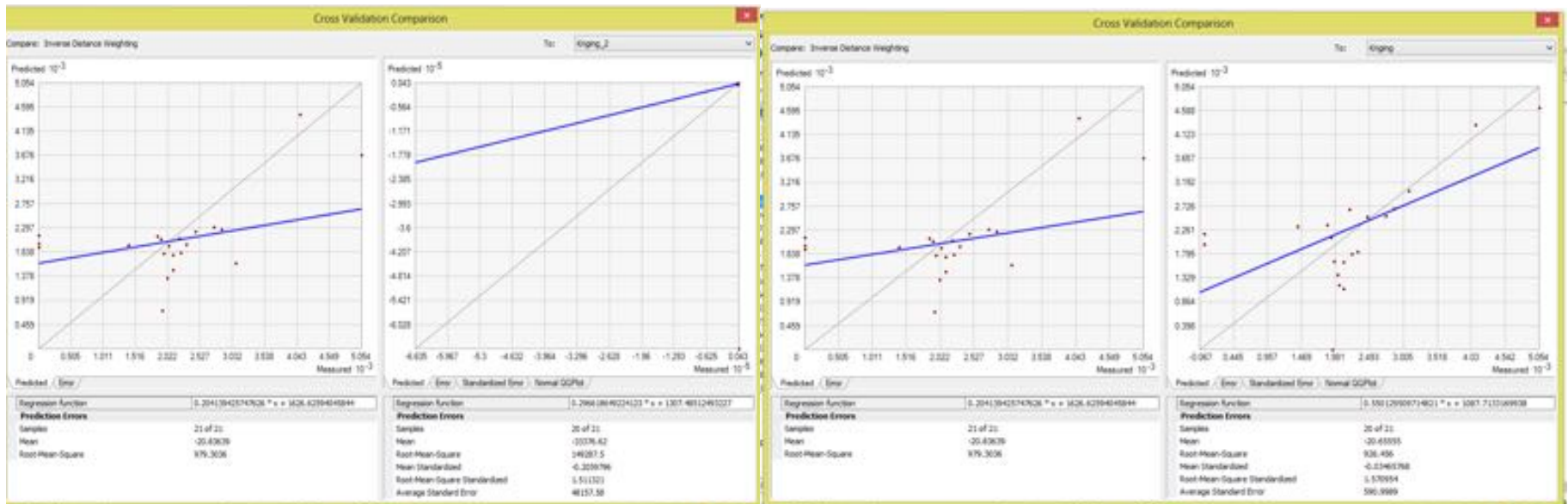
IDW vs. K. Universal, año 2009



IDW vs. K. Ordinario, año 2009

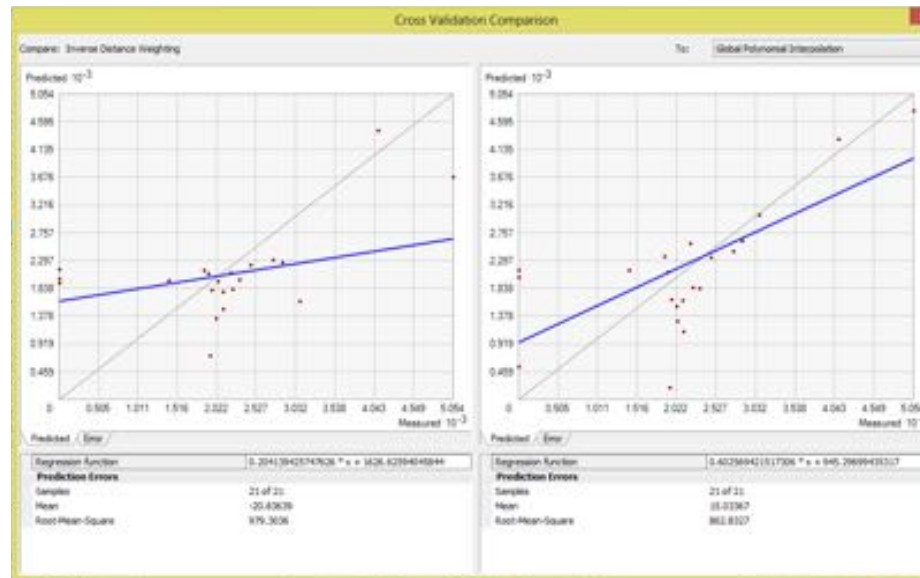


IDW vs. GPI, año 2009

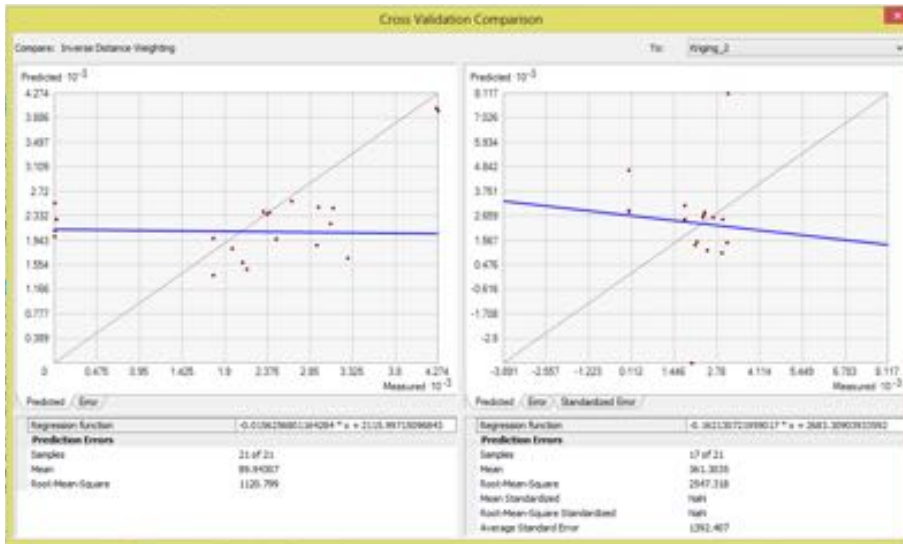


IDW vs. K. Universal, año 2012

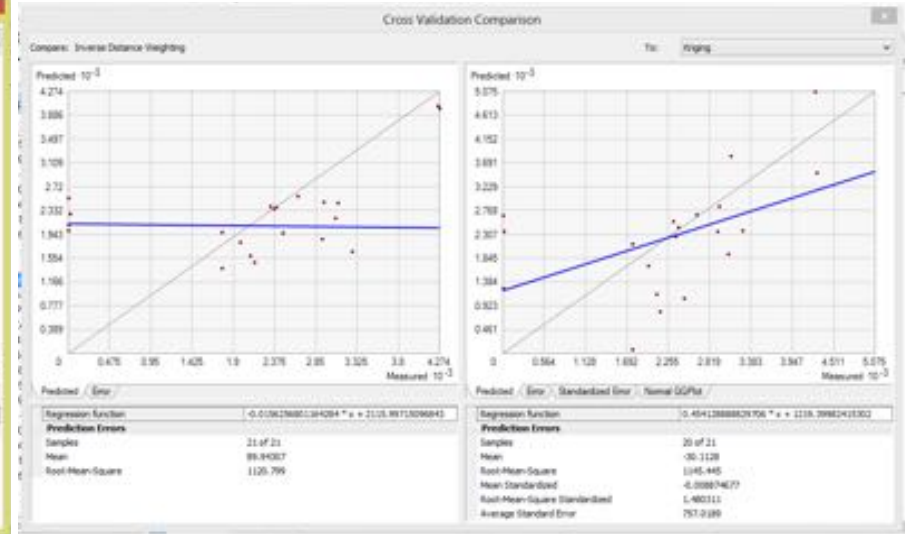
IDW vs. K. Ordinario, año 2012



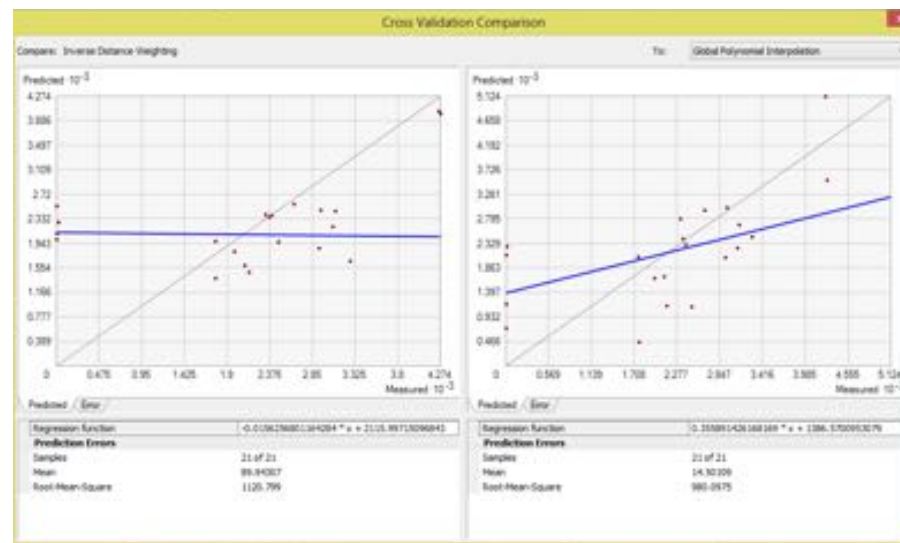
IDW vs. GPI, año 2012



IDW vs. K. Universal, año 2013



IDW vs. K. Ordinario, año 2013



IDW vs. GPI 2013