



**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.**

**Facultad de Ciencias Naturales y Matemáticas**

**“COMPARACIÓN DE METODOLOGÍAS ESTADÍSTICAS PARA  
INTERPOLAR LA PRECIPITACIÓN EN EL ECUADOR”**

Informe de

**MATERIA INTEGRADORA**

Previa a la obtención de Título de:

**INGENIERO EN ESTADÍSTICA INFORMÁTICA**

Presentada por

**GABRIEL FERNANDO MANTILLA SALTOS**

**GUAYAQUIL – ECUADOR**

**2016**

## AGRADECIMIENTOS

*A Dios*

*A la M.Sc. Sonia Reyes Ramos, por su colaboración y conocimientos entregados  
para poder desarrollar muy bien este trabajo de investigación*

*Al Ph.D. Francisco Vera Alcívar, por haberme brindado la oportunidad de recurrir  
a su capacidad y conocimiento científico*

*Al M.Sc. Raúl Mejía, por sus sugerencias en esta investigación, por su colaboración  
con la información de los datos meteorológicos a través del  
Instituto nacional de Meteorología e hidrología del Ecuador.*

## DEDICATORIA

*Este trabajo se lo dedico con todo el amor y cariño a mis padres, ya que sin ellos no pude haber culminado mi sueño de poder obtener un título académico, todo se lo debo a mis padres, ellos siempre confiaron y tuvieron fe en que yo podría lograr mis metas, y con su ayuda, paciencia y dedicación, logramos juntos este logro.*

*También dedico este trabajo a mis profesores y a mi querida Universidad, por haberme brindado la formación humana y profesional con la que ahora cuento, muchas gracias a mí segundo hogar ESPOL, por haberme abierto las puertas al conocimiento científico, y ahora poder llevar en alto su nombre a donde quiera que me dirija.*



ESCUELA SUPERIOR POLITECNICA DEL LITORAL

Facultad de Ciencias Naturales y Matemáticas

Ingeniería en Estadística Informática

**INFORME DEL PROFESOR DE LA MATERIA INTEGRADORA**

Habiendo sido nombrada PROFESOR DE LA MATERIA INTEGRADORA del señor

**GABRIEL FERNANDO MANTILLA SALTOS**

Con el tema del proyecto integrador “COMPARACIÓN DE METODOLOGÍAS ESTADÍSTICAS PARA INTERPOLAR LA PRECIPITACIÓN EN EL ECUADOR”, previa a la obtención del título de INGENIERO EN ESTADÍSTICA INFORMÁTICA, me permito informar que he leído el contenido del proyecto integrador; luego de lo cual indico que estoy de acuerdo en que el mismo se lo ha desarrollado conforme a los lineamientos de la Unidad de Titulación Especial de la ESPOL.

**M.Sc. Sonia Paola Reyes Ramos**  
**PROFESOR DE LA MATERIA INTEGRADORA**

Guayaquil, 22 de febrero del 2016

## **DECLARACION EXPRESA**

"La responsabilidad del contenido de esta Tesis de Grado, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la Escuela Superior Politécnica del Litoral".

ATENTAMENTE

-----  
GABRIEL MANTILLA SALTOS

## RESUMEN

Contrastar 4 metodologías estadísticas para definir cuál de ellas interpreta mejor el comportamiento de las precipitaciones en Ecuador. Los métodos usados son: Ponderación por Distancias Inversas, Kriging Universal, Cokriging y Kriging Bayesiano. Se utilizó como datos de estudio 78 estaciones meteorológicas distribuidas alrededor de la parte baja de la Cuenca del Río Guayas, en la parte alta de la región Costa y en la región Interandina. Las variables monitoreadas en las estaciones meteorológicas son: la Precipitación y la Humedad Relativa como variables de interés, y como variables de fijar de explicación: la Altitud y la Distancia de la estación con respecto al mar. Los datos son promedios de los 12 meses del año desde 1990 hasta el 2015. Se utilizaron medidas de ajuste como los Mínimos Cuadrados Ponderados para minimizar la Suma Cuadrática del Error (SCE), optimizando los parámetros de entrada para los modelos propuestos. Se realizó Validación Cruzada para medir el nivel de error de la salida de cada modelo y así poder comparar las metodologías. Basándonos en los Errores concluimos que para los meses secos de Junio a Noviembre el kriging Bayesiano tiene menor SCE con respecto a los otros métodos, mientras que el Kriging y el Cokriging en los meses lluviosos de Diciembre a Mayo tienen menor SCE con respecto a los otros métodos. En todos los meses de Enero a Diciembre la varianza del Cokriging resulta ser la más baja debido a que la Altitud condiciona a la precipitación, dando un aporte significativo a la predicción espacial, y de esta manera obtener mayor confianza en utilizar este método, al momento de interpretar sus resultados.

## INDICE GENERAL

	Pagina
AGRADECIMIENTO.....	I
DEDICATORIA.....	II
TRIBUNAL DE GRADUACION.....	III
DECLARACION EXPRESA.....	IV
RESUMEN.....	V
INDICE GENERAL.....	VI
INDICE DE TABLAS.....	VIII
INDICE DE GRÁFICOS.....	IX
INDICE DE ILUSTRACIONES.....	X
ABREVIATURAS.....	XI
SIMBOLOGÍA.....	XII
CAPITULO I.....	1
1. PLANTEAMIENTO DEL PROBLEMA.....	1
1.1 Definición del Problema.....	2
1.2 Objetivo General.....	2
1.3 Objetivos Específicos.....	2
1.4 Justificación del Problema.....	3
CAPITULO II.....	5
2. MARCO TEORICO.....	5
2.1. Área de Estudio.....	5
2.2. Variables de Estudio y Explicación.....	6
CAPITULO III.....	13
3. METODOLOGÍAS.....	13
3.1. METODOLOGÍA DESCRIPTIVA.....	13
3.1.1. Diagrama de Cajas.....	14
3.1.2. Mínimo.....	14
3.1.3. Mediana.....	14
3.1.4. Media.....	14
3.1.5. Máximo.....	15
3.2. METODOLOGÍA INFERENCIAL.....	15
3.2.1. Prueba de Shapiro Wilk.....	15
3.2.2. Grafico Quantil Quantil.....	16
3.2.3. Histograma de Frecuencia.....	16
3.3. METODOLOGÍA ESPECIALIZADA.....	16
3.3.1. Ponderación por Distancias Inversas.....	17
3.3.2. Kriging Universal.....	18
3.3.3. Cokriging.....	23
3.3.4. Kriging Bayesiano.....	25
3.3.5. Suma Cuadrática del Error.....	28
3.3.6. Mínimos Cuadrados Ponderados.....	28
3.3.7. Validación Cruzada.....	29

CAPITULO IV.....	30
4. RESULTADOS.....	30
4.1. Resultados Metodología Descriptiva.....	30
4.2. Resultados Metodología Inferencial.....	35
4.3. Resultados Metodología Especializada.....	40
CONCLUSIONES Y RECOMENDACIONES.....	89
BIBILOGRAFÍA.....	91
ANEXOS.....	94
A. Glosario de Términos.....	94
B. Código Fuente.....	96



## INDICE DE TABLAS

	Pagina
Tabla 2.1 Estaciones con altitud menor a 500 metros.....	9
Tabla 2.2. Estaciones con altitud mayor a 500 metros.....	10
Tabla 2.3. Estaciones con altitud mayor a 500 metros.....	11
Tabla 3.1. Variogramas teóricos.....	22
Tabla 4.1. Estadísticas Descriptivas, estaciones con altitud menor a 500 m...	31
Tabla 4.2. Estadísticas Descriptivas, estaciones con altitud mayor a 500 m...	34
Tabla 4.3. Parámetros de los Modelos para Enero.....	41
Tabla 4.4. Parámetros de los Modelos para Febrero.....	45
Tabla 4.5. Parámetros de los Modelos para Marzo.....	49
Tabla 4.6. Parámetros de los Modelos para Abril.....	53
Tabla 4.7. Parámetros de los Modelos para Mayo.....	57
Tabla 4.8. Parámetros de los Modelos para Junio.....	61
Tabla 4.9. Parámetros de los Modelos para Julio.....	65
Tabla 4.10. Parámetros de los Modelos para Agosto.....	69
Tabla 4.11. Parámetros de los Modelos para Septiembre.....	73
Tabla 4.12. Parámetros de los Modelos para Octubre.....	77
Tabla 4.13. Parámetros de los Modelos para Noviembre.....	81
Tabla 4.14. Parámetros de los Modelos para Diciembre.....	85

## INDICE DE GRÁFICOS

	Pagina
Grafico 3.1. Representación del Semivariograma experimental.....	22
Grafico 4.1. Diagrama de Cajas Estaciones con una altitud menor a 5 m.....	30
Grafico 4.2. Diagrama de Cajas Estaciones con una altitud mayor a 5 m.....	33
Grafico 4.3. Gráficos Quantil Quantil, Precipitación.....	36
Grafico 4.4. Histogramas de Frecuencia, Precipitación.....	37
Grafico 4.5. Gráficos Quantil Quantil, Humedad Relativa.....	38
Grafico 4.6. Histogramas de Frecuencia, Humedad Relativa.....	39
Grafico 4.7. Medidas de Ajuste, Normal de Enero.....	41
Grafico 4.8. Validación cruzada Enero.....	42
Grafico 4.9. Medidas de Ajuste, Normal de Febrero.....	45
Grafico 4.10. Validación cruzada Febrero.....	46
Grafico 4.11. Medidas de Ajuste, Normal de Marzo.....	49
Grafico 4.12. Validación cruzada Marzo.....	50
Grafico 4.13. Medidas de Ajuste, Normal de Abril.....	53
Grafico 4.14. Validación cruzada Abril.....	54
Grafico 4.15. Medidas de Ajuste, Normal de Mayo.....	57
Grafico 4.16. Validación cruzada Mayo.....	58
Grafico 4.17. Medidas de Ajuste, Normal de Junio.....	61
Grafico 4.18. Validación cruzada Junio.....	62
Grafico 4.19. Medidas de Ajuste, Normal de Julio.....	65
Grafico 4.20. Validación cruzada Julio.....	66
Grafico 4.21. Medidas de Ajuste, Normal de Agosto.....	69
Grafico 4.22. Validación cruzada Agosto.....	70
Grafico 4.23. Medidas de Ajuste, Normal de Septiembre.....	73
Grafico 4.24. Validación cruzada Septiembre.....	74
Grafico 4.25. Medidas de Ajuste, Normal de Octubre.....	77
Grafico 4.26. Validación cruzada Octubre.....	78
Grafico 4.27. Medidas de Ajuste, Normal de Noviembre.....	81
Grafico 4.28. Validación cruzada Noviembre.....	82
Grafico 4.29. Medidas de Ajuste, Normal de Diciembre.....	85
Grafico 4.30. Validación cruzada Diciembre.....	86

## INDICE DE ILUSTRACIONES

	Pagina
Ilustración 2.1. Estaciones a diferentes altitudes.....	8
Ilustración 2.2. Distancia de los 88956 puntos al borde costero del Ecuador.	12
Ilustración 4.1. Varianzas de los interpoladores para Enero.....	43
Ilustración 4.2. Interpolación de los métodos para Enero.....	44
Ilustración 4.3. Varianzas de los interpoladores para Febrero.....	47
Ilustración 4.4. Interpolación de los métodos para Febrero.....	48
Ilustración 4.5. Varianzas de los interpoladores para Marzo.....	51
Ilustración 4.6. Interpolación de los métodos para Marzo.....	52
Ilustración 4.7. Varianzas de los interpoladores para Abril.....	55
Ilustración 4.8. Interpolación de los métodos para Abril.....	56
Ilustración 4.9. Varianzas de los interpoladores para Mayo.....	59
Ilustración 4.10. Interpolación de los métodos para Mayo.....	60
Ilustración 4.11. Varianzas de los interpoladores para Junio.....	63
Ilustración 4.12. Interpolación de los métodos para Junio.....	64
Ilustración 4.13. Varianzas de los interpoladores para Julio.....	67
Ilustración 4.14. Interpolación de los métodos para Julio.....	68
Ilustración 4.15. Varianzas de los interpoladores para Agosto.....	71
Ilustración 4.16. Interpolación de los métodos para Agosto.....	72
Ilustración 4.17. Varianzas de los interpoladores para Septiembre.....	75
Ilustración 4.18. Interpolación de los métodos para Septiembre.....	76
Ilustración 4.19. Varianzas de los interpoladores para Octubre.....	79
Ilustración 4.20. Interpolación de los métodos para Octubre.....	80
Ilustración 4.21. Varianzas de los interpoladores para Noviembre.....	83
Ilustración 4.22. Interpolación de los métodos para Noviembre.....	84
Ilustración 4.23. Varianzas de los interpoladores para Diciembre.....	87
Ilustración 4.24. Interpolación de los métodos para Diciembre.....	88

## **ABREVIATURAS**

INAMHI = Instituto Nacional de Meteorología e Hidrología del Ecuador.

SNMH = Servicio Nacional de Meteorología e Hidrología del Ecuador.

SNGR = Secretaria Nacional de Gestión de Riesgos.

IWD = Método de la distancia Inversa.

SCE = Suma Cuadrática de los Errores.

MCP = Mínimos Cuadrados Ponderados.

VC = Validación Cruzada.

m = Metros.

km = Kilómetros.

## SIMBOLOGÍAS

$\gamma$  Semivariograma Empírico.

$h$  Distancia de separación.

$N(h)$  Número total de pares de muestras separados por un intervalo de distancia  $h$ .

$\hat{\gamma}$  Es el semivariograma estimado.

$\gamma_{\theta}$  Es el semivariograma empírico.

$n$  Número de observaciones.

$Z$  Valores observados de una variable de estudio.

$\lambda$  Ponderadores constantes.

$Z_0$  Valores observados a estimar.

$x_0$  Observación conocida de la coordenada a predecir.

$m_0, m_1, m_2$  Multiplicadores de Lagrange.

# CAPITULO I

## 1. PLANTEAMIENTO DEL PROBLEMA

Los datos históricos muestran que la meteorología en el Ecuador se registra a partir del año 1890, en donde las primeras observaciones meteorológicas se realizaron en este año en las ciudades de Ambato y Quito, consolidando los estudios y mediciones meteorológicas a partir de los años 60's dando la creación al SNMH, que ahora en la actualidad gracias al decreto Ejecutivo 391 toma el nombre de estando adscrito a la SNGR desde el año 2010<sup>[1]</sup>.

Es el SNMH creado por Ley, como una necesidad y un derecho fundamental de la comunidad, con capacidad y la obligación de suministrar información vital sobre el tiempo, el clima y los recursos hídricos del pasado, presente y futuro, que necesita conocer el país para la protección de la vida humana y los bienes materiales<sup>[2]</sup>. Todo el esfuerzo realizado por los que estudian el tiempo, clima y agua, que pueden ayudar a mitigar los efectos negativos de desastres naturales basándonos en el Plan del Buen Vivir vigente en el actual gobierno 2013 – 2017<sup>[3]</sup>, siendo la población ecuatoriana, los beneficiarios directos de la información generada por la red de estaciones hidrometeorológicas con la que cuenta la institución.

### **1.1 Definición del Problema.**

El enfoque de este estudio es poder estimar el comportamiento de la lluvia de modo espacial, en zonas de difícil acceso donde no se ha podido registrar datos climáticos, por la forma de cómo fue diseñada la red de estaciones hidrometeorológicas.

### **1.2 Objetivo General.**

Contrastar que métodos de interpoladores espaciales tienen mejor potencial predictivo para explicar el comportamiento de las lluvias en el Ecuador, validando los modelos con la suma cuadrática de los Errores, y realizando una validación de los pronósticos espaciales con respecto a los valores observados, utilizando la validación cruzada.

### **1.3 Objetivos Específicos**

- ✓ Describir los datos de las estaciones con estadísticas de tipo continuas, como el diagrama de cajas, mínimo, mediana, máximo y media.
- ✓ Describir ciertas propiedades de la variable de estudio, utilizando Histogramas de frecuencia, Pruebas de Normalidad y Grafico Quantil-Quantil
- ✓ Utilizar cuatro metodologías distintas como la ponderación de distancias inversas, Kriging universal, cokriging y el kriging bayesiano.
- ✓ Realizar una validación de los modelos con estadísticas de tipo continuas, como el la suma cuadrática de los errores.
- ✓ Realizar una validación de los pronósticos con respecto a los valores observados, utilizando la validación cruzada de la suma cuadrática de los errores.

- ✓ Realizar mapas temáticos con las salidas de los modelos anteriormente definidos, para poder analizar sus diferencias de modo gráfico.
- ✓ Discernir los resultados del proyecto en una conferencia con entendedores del tema.

#### **1.4. Justificación del Problema.**

Una estación meteorológica en nuestro país cuesta aproximadamente de treinta mil a cincuenta mil dólares muy aparte del costo de manteniendo, una estación climatológica ordinaria tiene una cobertura de 50 kilómetros <sup>[4]</sup>. Aplicando el conocimiento geoestadístico se puede hacer inferencia a esas zonas donde no hay cobertura de las estaciones como las zonas boscosas, montañosas por las pendientes pronunciadas o en zonas donde simplemente no se ha hecho un estudio que demuestre y amerite implementar una estación debido a que no hay densidad poblacional.

Según el Plan del Buen Vivir 2013 – 2017 es de interés estudiar el clima para poder tomar decisiones sobre los posibles riesgos y desastres de origen natural que podrían darse en zonas vulnerables a inundaciones, para de este modo prevenir estos efectos negativos que causan pérdidas de infraestructura o humanas producidas por estos eventos extremos, disminuyendo así la incertidumbre en dichas zonas exponiendo mapas ilustrativos que representen la variabilidad espacial de la lluvia en zonas donde no ha sido posible medir la atmósfera, beneficiando también a los agricultores al momento de tomar decisiones, para invertir en un sembrío o no, ante la llegada de un evento adverso, y a las autoridades de salud para poder prevenir brotes de dengue o



cualquier otra enfermedad. También aprovechando el análisis de los modelos estadísticos definidos, se podrá contrastar cuál de ellos es el más óptimo.

## CAPITULO II

### 2.1 MARCO TEORICO

Para la realización de esta investigación, es necesario contar con información espacial de las variables de estudio, y las variables predictoras, las variables de estudio son las precipitaciones y la humedad en el Ecuador, sin contar con la información del clima en el oriente, y en las islas Galápagos. Se tiene una base de datos con un periodo de tiempo desde el año 1990 hasta el 2015 de 78 estaciones hidrometeorológicas y meteorológicas, todas ellas cuentan con el valor de la precipitación en milímetros y la humedad relativa en términos de 0 a 1, también está la información de su elevación con respecto al mar en metros. Pero por problemas técnicos hay ciertos meses con valores faltantes, para no rellenar estos valores faltantes, se calculó el promedio de cada mes desde 1990 hasta el 2015, para cada estación. Con esto tenemos 12 valores completos para los 12 meses del año.

#### 2.1. Área de Estudio

De modo ilustrativo mostraremos a continuación 2 mapas, representado el lugar donde se encuentra cada estación, y la distancia a la que se encuentra cada uno de los puntos que voy a predecir. Para esto se tienen 88956 puntos que voy a predecir en todo el Ecuador. Esta base de datos contiene la elevación y la distancia al mar en metros en cada uno de esos 88956 puntos. Estos puntos se los obtuvo de una imagen raster, describiendo la elevación cada 200 metros aproximadamente entre puntos. Y con un algoritmo desarrollado en la plataforma de programación estadística R, se obtuvo la

distancia mínima de cada uno de los 88956 puntos con respecto al borde costero del Ecuador en metros.

También mostraremos 3 tablas con los datos de la posición geográfica en coordenadas decimales de cada estación,

## **2.2. Variables de Estudio y Explicación.**

Refiriéndonos a la variable de estudio (Precipitación), es la cantidad de agua registrada en milímetros por cada metro cuadrado, mostraremos dos grupos de estaciones. El primer grupo tiene 29 estaciones con una altitud no mayor a 500 metros, estas estaciones se encuentran situadas alrededor de la Cuenca Baja del Río Guayas, el segundo grupo de 49 estaciones con una altitud mayor a 500 metros, se ubican entre la parte alta de la costa y la Región Interandina. Es de interés diferenciarlas de esa manera, para poder describir el comportamiento de las estaciones a diferentes altitudes. En la parte norte de la costa, encontramos bosque húmedo tropical con exuberante vegetación, gracias a que está ubicada en el corredor Ecológico de Choco (bosque tropical húmedo que se distribuye en la provincia de Esmeraldas, Manabí, Carchi, Imbabura, Pichincha), en el centro de la costa encontramos remanente de bosques húmedo tropical, en cambio en la parte sur encontramos bosque seco. En la parte alta de la costa y la Región Interandina encontramos bosques humedecidos por las lluvias y neblina donde se forman las nubes.

En la región costa la precipitación tiene un comportamiento unimodal (un máximo de precipitación siendo el mes de marzo), la época lluviosa empieza en finales de diciembre y se extiende aproximadamente hasta mediados de mayo, y la época seca

empieza a mediados de mayo y se extiende aproximadamente hasta mediados de diciembre. La región interandina tiene un comportamiento bimodal (dos picos de precipitación bien definidos que son en abril y octubre), la primera época lluviosa empieza en enero y termina en mayo, y la segunda época lluviosa empieza a mediados de septiembre y se extiende hasta noviembre, la época seca empieza en junio y termina en agosto. Algunas estaciones que se encuentran dentro de valles bajos, presentando un clima frío aunque no están a gran altitud se consideran de la sierra.

Refiriéndonos a la variable humedad relativa, representada en porcentaje, es la relación que tiene la presión parcial de vapor de agua en un cuerpo de aire con la presión de saturación de agua a la temperatura del cuerpo de aire, a mayor elevación podemos encontrar menor porcentaje de humedad, y a menor elevación tenemos mayor porcentaje de humedad. La Humedad relativa puede ser expresada perfectamente en términos de 0 a 1, pero generalmente es expresarla en porcentajes de 0 a 100, en esta investigación se la uso en términos de 0 a 1 para su interpretación al momento de aplicar el cokriging, dado que el pronóstico de esta metodología trata de explicar el comportamiento de ambas variables. Pero como todos los mapas están en milímetros, por ejemplo si sumamos una humedad relativa de 100 a 450 milímetros, no podríamos decir que han precipitado 550 milímetros en una estación, esto no ocurre si sumamos 0.1 a 450 milímetros siendo 450.1. De tal manera que no podríamos relacionar el mapa resultante del pronóstico del cokriging, con los mapas resultantes de los pronósticos de los otros interpoladores.

Mapa del Ecuador, posición geográfica de las estaciones y elevación

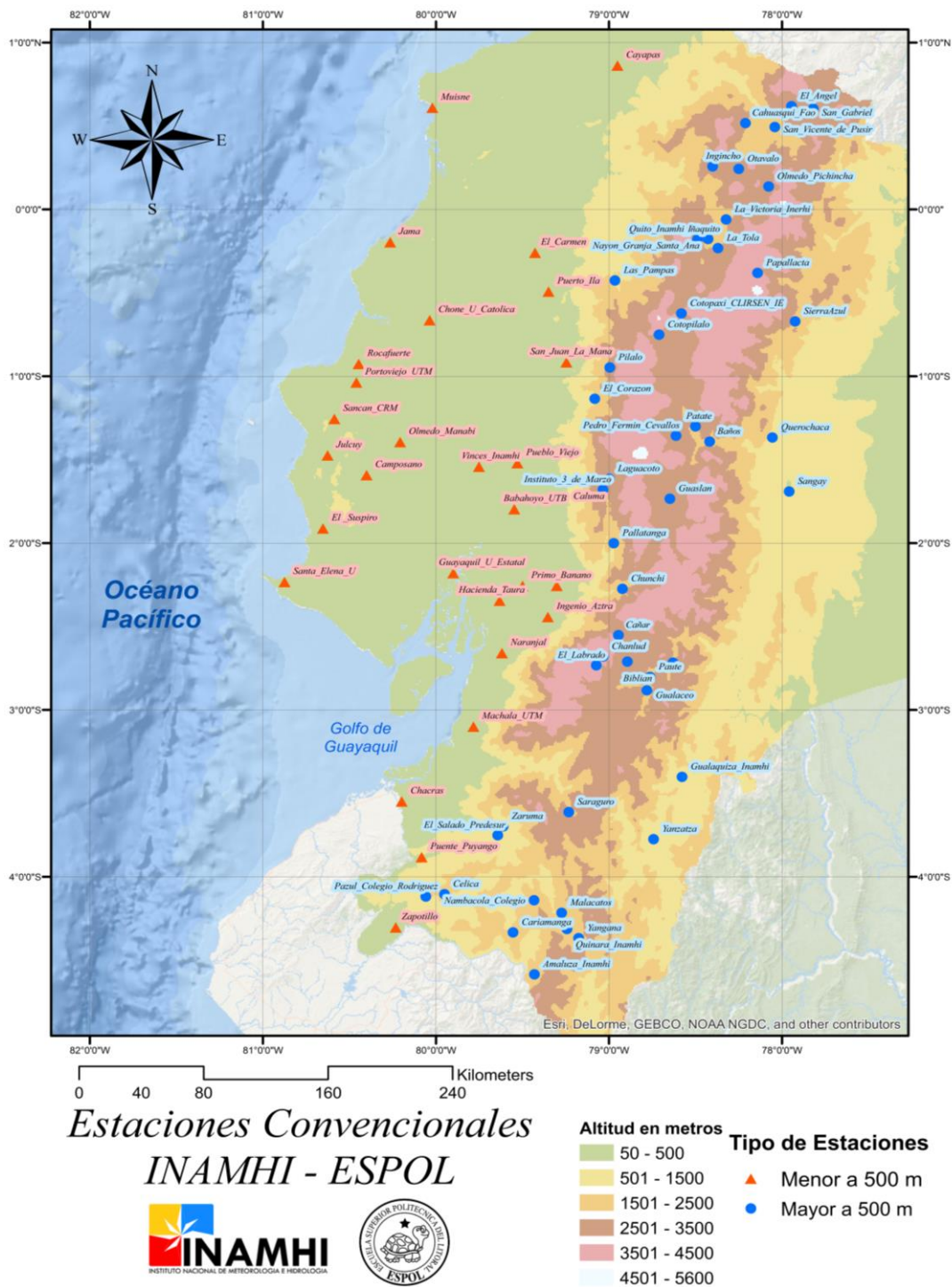


Ilustración 2.1. Estaciones a diferentes altitudes.  
Elaborado por: Gabriel Mantilla.

Estaciones con una altitud menor a los 500 metros.

Estación	Longitud	Latitud	Elevación
<b>El Suspiro</b>	-80,6530556	-1,9127778	0
<b>Muisne</b>	-80,0208333	0,6095556	5
<b>Guayaquil Univ. Estatal</b>	-79,8999056	-2,1805167	6
<b>Babahoyo UTB</b>	-79,5472222	-1,7969444	7
<b>Hacienda Taura</b>	-79,6318917	-2,3455111	12
<b>Machala UTM</b>	-79,7833333	-3,1	13
<b>Santa Elena Univ.</b>	-80,875125	-2,2336111	13
<b>Vinces Inamhi</b>	-79,7513333	-1,5423611	14
<b>Pueblo Viejo</b>	-79,5287778	-1,52125	19
<b>Rocafuerte</b>	-80,4467778	-0,9267222	20
<b>Ingenio San Carlos</b>	-79,3013889	-2,2566667	35
<b>Chone Univ. Católica</b>	-80,0365	-0,6641389	39
<b>Jama</b>	-80,2647222	-0,1969444	46
<b>Portoviejo UTM</b>	-80,4597222	-1,0375	46
<b>Primo Banano</b>	-79,4991667	-2,2569444	49
<b>Ingenio Aztra</b>	-79,3525528	-2,4432667	50
<b>Naranjal</b>	-79,6180556	-2,6599167	50
<b>Olmedo Manabí</b>	-80,2069444	-1,3946944	50
<b>Cayapas</b>	-78,95075	0,8622222	55
<b>Chacras</b>	-80,1978333	-3,5486667	60
<b>Camposano</b>	-80,4011667	-1,5928333	156
<b>San Juan La Mana</b>	-79,2455556	-0,9163889	215
<b>Zapotillo</b>	-80,2324444	-4,3036944	223
<b>Sancan CRM</b>	-80,5873278	-1,2561083	229
<b>El Carmen</b>	-79,4269444	-0,2597222	260
<b>Julcuy</b>	-80,6261667	-1,4745278	263
<b>Puente Puyango</b>	-80,0829167	-3,8825556	305
<b>Puerto Ila</b>	-79,3488889	-0,4927778	319
<b>Caluma</b>	-79,2558333	-1,6302778	350

Tabla 2.1. Estaciones con altitud menor a 500 metros.

Elaborado por: Gabriel Mantilla

Estaciones con una altitud mayor a los 500 metros.

<b>Estación</b>	<b>Longitud</b>	<b>Latitud</b>	<b>Elevación</b>
<b>El Salado Predesur</b>	-79,6430556	-3,7510833	570
<b>Gualaquiza Inamhi</b>	-78,5775833	-3,4009444	750
<b>Yanzatza</b>	-78,7423056	-3,7744167	830
<b>Sangay</b>	-77,9586111	-1,6908333	880
<b>Zaruma</b>	-79,6115556	-3,6975278	1100
<b>Malacatos</b>	-79,2733333	-4,2161111	1453
<b>El Corazón</b>	-79,0819444	-1,1350833	1471
<b>Pallatanga</b>	-78,9722222	-2,0004722	1523
<b>Quinara Inamhi</b>	-79,2425	-4,3131111	1559
<b>Las Pampas</b>	-78,965	-0,4255556	1583
<b>Amaluza Inamhi</b>	-79,4305556	-4,5847222	1672
<b>Baños</b>	-78,4180556	-1,3913889	1695
<b>Pazul Colegio Rodriguez</b>	-80,0589556	-4,1183528	1739
<b>Nambacola Colegio</b>	-79,4331167	-4,1405583	1835
<b>Yangana</b>	-79,1750556	-4,3679167	1835
<b>San Vicente de Pusir</b>	-78,0413889	0,4941667	1891
<b>Celica</b>	-79,9509722	-4,1044167	1904
<b>Cariamanga</b>	-79,5543333	-4,3333889	1950
<b>Chunchi</b>	-78,9222222	-2,2741667	2177
<b>Paute</b>	-78,7627778	-2,8024722	2194
<b>Patate</b>	-78,5	-1,3002778	2220
<b>Gualaceo</b>	-78,7805556	-2,8819444	2230
<b>Sierra Azul</b>	-77,9241778	-0,6708417	2240
<b>La Victoria Inerhi</b>	-78,3230556	-0,06	2262
<b>Mira FAO Granja la Portada</b>	-78,0275	0,5530556	2275
<b>Cahuasqui Fao</b>	-78,2111111	0,5180556	2335
<b>Nayon Granja Santa Ana</b>	-78,4271861	-0,1770167	2395
<b>Palmas Azuay</b>	-78,6297222	-2,7161111	2400
<b>La Tola</b>	-78,3702778	-0,2316667	2480
<b>Instituto_3_de_Marzo</b>	-79,0346528	-1,6798389	2504

Tabla 2.2. Estaciones con altitud mayor a 500 metros.

Elaborado por: Gabriel Mantilla

Estaciones con una altitud mayor a los 500 metros.

<b>Estación</b>	<b>Longitud</b>	<b>Latitud</b>	<b>Elevación</b>
<b>Pilalo</b>	-78,995	-0,9472222	2504
<b>Saraguro</b>	-79,2322222	-3,6122222	2525
<b>Otavalo</b>	-78,25	0,2433333	2550
<b>Laguacoto</b>	-78,9983389	-1,6143778	2608
<b>Biblian</b>	-78,8936111	-2,7102778	2640
<b>Quito Inamhi Ñaquito</b>	-78,4877778	-0,1783333	2789
<b>Guaslan</b>	-78,6477778	-1,7340278	2850
<b>San Gabriel</b>	-77,8194444	0,6041667	2860
<b>Querochaca</b>	-78,0555	-1,3671111	2865
<b>Pedro Fermin Cevallos</b>	-78,6127778	-1,3561111	2910
<b>El Angel</b>	-77,9448333	0,6189444	3000
<b>Cañar</b>	-78,9452778	-2,5519444	3083
<b>Olmedo Pichincha</b>	-78,0775	0,1377778	3120
<b>Ingincho</b>	-78,4008333	0,2583333	3140
<b>Papallacta</b>	-78,1413889	-0,3806111	3150
<b>Cotopilalo</b>	-78,7104472	-0,7508056	3250
<b>El Labrado</b>	-79,0730556	-2,7327778	3335
<b>Chanlud</b>	-79,0339528	-2,6800639	3336
<b>Cotopaxi CLIRSEN IE</b>	-78,5813889	-0,6233333	3510

Tabla 2.3. Estaciones con altitud mayor a 500 metros.

Elaborado por: Gabriel Mantilla



Mapa de distancias al mar, en metros.

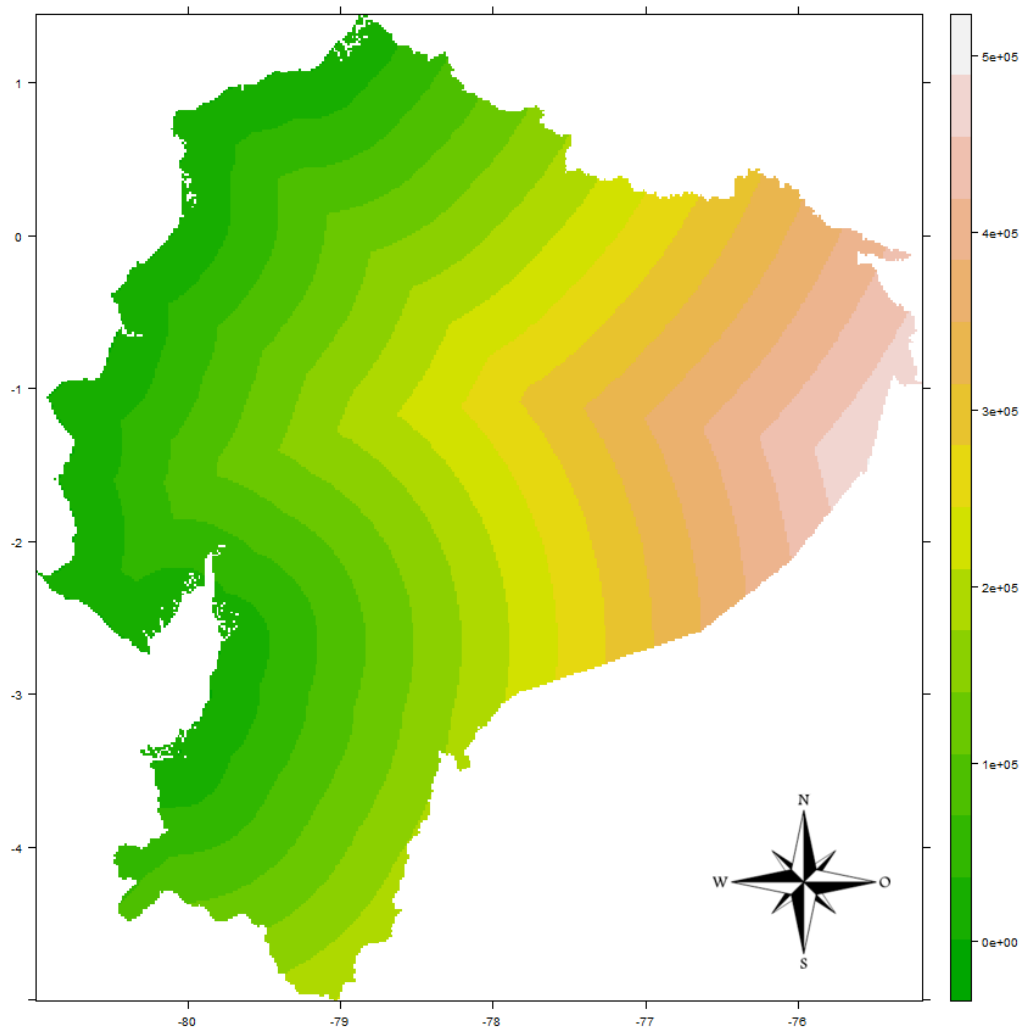


Ilustración 2.2. Distancia de los 88956 puntos al borde costero del Ecuador.  
Elaborado por: Gabriel Mantilla.

La ilustración No. II, representa la distancia a la que se encuentra cada zona con respecto al borde costero, siendo el color verde la menor distancia, y el color naranja con una distancia mayor a 400.000 metros.

## **CAPITULO III**

### **3. METODOLOGÍAS**

En este capítulo, se expondrá las metodologías estadísticas a usarse en esta investigación, primero está la estadística descriptiva, que nos introducirán de forma breve los datos con los que contamos. Luego está la estadística inferencial, con la que podemos entender más la naturaleza de las distribuciones. Por último esta la estadística especializada, en donde se explicaran las ecuaciones de los interpoladores, y las estadísticas con las que vamos a medir el rendimiento de estos interpoladores.

#### **3.1. METODOLOGÍA DESCRIPTIVA**

Esta metodología describe los datos utilizados en este estudio, siendo el mínimo, media, mediana y máximo mostrados en los resultados de la metodología descriptiva, estos valores también pueden ser representados en los gráficos de diagramas de caja. Los valores atípicos son los máximos y mínimos mostrados alrededor de los bigotes del diagrama de caja.

Para la realización del gráfico del diagrama de caja de todos los meses del año, se calculó el promedio de cada estación por mes, en las cajas se observan puntos, siendo cada punto los promedios calculados anteriormente, y el rango definido de años o climatología es desde 1990 hasta el 2014. Para todos estos análisis se crearon hojas de cálculo en Excel, que fueron luego guardadas como archivos de texto plano, que luego fueron utilizadas en los respectivos cálculos estadísticos.

### **3.1.1. Diagrama de Cajas.**

Se realizó un diagrama de cajas para describir el comportamiento de las precipitaciones en las estaciones por cada uno de los meses, un diagrama de cajas muestra la información de los datos ordenados en forma creciente, el 25% de los datos ordenados es el primer cuartil, el segundo cuartil es el 50% de los datos ordenados también llamado mediana, y el tercer cuartil es el 75% de los datos ordenados'

Los bigotes de el diagrama de caja representan el valor máximo y mínimo, además podemos observar los valores atípicos que están fuera del rango intercuartílico que es la diferencia entre el primer y tercer cuartil, este grafico nos enseña la dispersión de los datos con respecto a las estadísticas mencionadas anteriormente.

### **3.1.2. Mínimo.**

El mínimo es el valor numérico más bajo del conjunto de datos, este valor representa la estación con las más baja precipitación con respecto a las demás estaciones, para cada uno de los 12 meses.

### **3.1.3. Mediana.**

Este valor numérico en milímetros representa el 50% de los datos ordenados de menor a mayor, esta estadística es bien similar a la media aritmética, dado que la media representa el valor central de los datos.

### **3.1.4. Media.**

Para un conjunto de datos cuantitativos, se denomina media aritmética a la suma de todo el conjunto de datos dividido para el número total de datos, esta estadística también es conocida como el promedio de una muestra aleatoria de una población

conocida. La media representa el valor de la estación promedio de cada uno de los 12 meses.

### **3.1.5. Máximo.**

El máximo es el valor numérico más alto de un conjunto de datos, este valor representa la estación con la más alta precipitación con respecto a las demás estaciones, para cada uno de los 12 meses.

## **3.2. METODOLOGÍA INFERENCIAL**

Se realizó la Prueba de Shapiro Wilk para cada mes con su respectivo Grafico QQ para poder analizar si los datos provienen de una distribución normal. También se realizaron histogramas de frecuencia para ver la dispersión de los datos de cada mes y poder contrastar con la prueba de Shapiro Wilk la normalidad de los datos, con este análisis podremos conocer mejor nuestra variable de estudio, y así ajustar mejor los modelos de interpolación.

### **3.2.1. Prueba de Shapiro Wilk.**

Se plantea la hipótesis Nula que la muestra proviene de una distribución normal, la muestra son los promedios de cada mes de todas las estaciones, si el estadístico p es menor que 0.1 entonces se concluye que la hipótesis nula es falsa por tanto se la rechaza, caso contrario si el valor p es mayor a 0.1 se acepta la hipótesis nula concluyendo que los datos son gaussianos.

### **3.2.2. Grafico Quantil-Quantil.**

Este método gráfico es la comparación entre la distribución de probabilidad de dos conjuntos de datos, el primer conjunto es la muestra o datos utilizados en este análisis que es el promedio de cada mes de todas las estaciones, y el segundo conjunto de datos son números aleatorios de una distribución normal, si los datos provienen de una distribución normal, los puntos se centraran a lo largo de la línea diagonal de color rojo del gráfico, caso contrario si no se parecen se concluye que los datos no son gaussianos.

### **3.2.3. Histograma de Frecuencia**

Es un gráfico de barra ordenado por clases o categorías, cada barra representa el conteo de los valores que están dentro de los intervalos o clases previamente definidos de un conjunto de datos, con este gráfico podemos ver la tendencia de los datos, el grado de variabilidad y la dispersión que toman las clases.

## **3.3. METODOLOGÍA ESPECIALIZADA.**

Unos de los objetivos de este proyecto, es experimentar con modelos estadísticos para modelar espacialmente el comportamiento de la precipitación, obtener estadísticas que describan el potencial predictivo de cada uno de ellos, y poder concluir cuál de ellos es el más óptimo, la propuesta es utilizar el kriging universal, cokriging, kriging bayesiano y un interpolador menos complejo como el método de las ponderaciones de las distancias inversas.

### 3.3.1. Ponderación por Distancias Inversas.

Por experiencia en aplicar estos métodos, proponemos utilizar este método como referencia a los métodos kriging, al momento de graficar el mapa utilizando este método, no sobre estima los valores observados ni los modifica, mostrando en el mapa los valores exactos de las observaciones, Romina Mezher y Pablo Mercuri <sup>[5]</sup> manifiesta que este método no suaviza ni sobrestima los valores manteniendo sus valores observados, una consideración de este método es que a medida que uno se aleja más de la coordenada observada, la incertidumbre aumenta, por eso es recomendable tener muchos puntos para utilizar este método, esto quiere decir que este método predice bien cerca de las observaciones más no las zonas que están más lejos, otra consideración es que al ser un modelo más simple, nos brinda menor complejidad de cálculo, este método utiliza el algoritmo basado en distancias <sup>[6]</sup>.

$$\hat{Z}_j = \sum_{i=1}^n \lambda_i * Z_i \quad (1)$$

Sujeto a:

$$\lambda_{i,j} = \frac{\frac{1}{d(S_i, S_j)}}{\sum_i^n \frac{1}{d(S_i, S_j)}} \quad (2)$$

Donde n es el número de observaciones,  $S_i$  son las coordenadas de las observaciones,  $\hat{Z}_j$  es el valor de la observación en la coordenada desconocida que queremos estimar, y los  $Z_i$  son las observaciones de las coordenadas  $S_i$  conocidas.

$$Var(\hat{Z}_j) = 2 * c^t * L - L^t * B * L \quad (3)$$

$\lambda_{i,j}$  son los pesos que multiplican a las precipitaciones  $Z_i$  para  $i=1,2,\dots,78$  y  $j=1,2,\dots,88956$ ,  $\hat{Z}_j$  es el valor a predecir,  $d(S_i, S_j)$  es la distancia entre las coordenadas observadas y la coordenada a predecir,  $c$  es un vector resultante del cálculo del semivariograma teórico exponencial entre las observaciones y las predicciones  $\gamma(\hat{Z}_j - Z_i)$ .  $B$  es una matriz cuadrada resultante del cálculo del semivariograma teórico exponencial de las observaciones  $\gamma(Z_i - Z_i)$ ,  $L$  es la matriz de pesos  $\lambda_{i,j}$ .

### 3.3.2. Kriging Universal.

Este método se basa en calcular una coordenada de interés  $Z_0$ , teniendo como base una observación conocida, apoyándonos de covariables fijas como la altitud, la distancia al mar, etc. Utilizando las coordenadas  $Z_i$  conocidas, esto implica que cada una de esas coordenadas deben tener el valor de la variable de interés que queremos estimar, y el valor de las variables explicativas que estén correlacionadas con la variable de interés. Este método estadístico desarrollado en la década de los sesenta por Daniel G. Krige que luego fue formalizado por Georges Matheron, es denominado como el mejor estimador lineal insesgado, es lineal porque el modelo está representado como la combinación lineal de las observaciones por los ponderadores, e insesgado porque la media de los errores es igual a cero. Para la implementación de este método, no es necesario contar con datos que provengan de una distribución normal, pero un supuesto que si se debe de cumplir, es que las observaciones estén correlacionadas espacialmente, esto no se cumple si muy cerca de la coordenada no hubiera presencia de precipitación, una estación meteorológica cubre un radio de 30 km, es probable que

menor a esa distancia existan otras estaciones meteorológicas en donde se ha evidenciado precipitaciones, cumpliendo con el supuesto de que existe correlación espacial entre las coordenadas vecinas, y las observaciones que estén más juntas tienden a ser más parecidas a las que están más separadas. En la geoestadística la idea de autocorrelación se cuantifica a través de la función llamada semivariograma denotado por el símbolo  $\gamma$ .

### **Semivariograma Empírico.**

Es el semivariograma experimental de las coordenadas observadas, representa en función el proceso de dependencia espacial entre las coordenadas, a mayor distancia mayor semivarianza. Es calculado en base a la distancia en kilómetros a la que se encuentran las coordenadas geográficas según criterios de tolerancia, las librerías desarrolladas en R que contemplan el análisis del semivariograma empírico, contienen funciones que por default según las coordenadas observadas más distantes, crean un plano, trazan una diagonal entre los extremos del plano, y la tercera parte de esa distancia se utiliza como referencia para encontrar conjuntos de estaciones que se encuentran aproximadamente a esa distancia, el semivariograma cumple con las siguientes propiedades mencionadas a continuación.

1.-  $\gamma(h) = \gamma(-h)$ , siendo  $h$  la distancia de separación, esto quiere decir que

$$\text{var}(Z(s) - Z(v)) = \text{var}(Z(v) - Z(s)).$$

2.-  $\gamma(h) = 0$ , esto quiere decir que  $\text{var}(Z(s) - Z(s)) = 0$



$$3.- \lim_{\|h\| \rightarrow \infty} \frac{\gamma(h)}{\|h\|^2} = 0$$

4.-  $\gamma$  debe ser condicionalmente negativa, esto quiere decir que dado  $n$  coordenadas  $S_1, S_2, S_3, S_4, \dots, S_n$  y dado  $n$  números reales  $a_1, a_2, a_3, a_4, \dots, a_n$  tales que  $a_1 + a_2 + a_3 + a_4 + \dots + a_n = 0$ . Entonces  $\sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma(s_i - s_j) \leq 0$ . Esta condición es similar a la condición de definida positiva para las función de covarianza y la matriz de varianza y covarianza, asegurando que todas las varianzas no son negativas.

5.- Si el proceso espacial es isotrópico  $\gamma(\mathbf{h}) \equiv \gamma(h)$ , donde  $h \equiv \|\mathbf{h}\|$ .

El estimador empírico se lo define de la siguiente ecuación.

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Z(S_i) - Z(S_{i+h}))^2 \quad (4)$$

Ecuaciones

$$\hat{Z}_j = \sum_{i=1}^n \lambda_i * Z_i \quad (5)$$

Sujeto a:

$$\sum_{i=1}^n \lambda_i = 1 \quad (6)$$

$$\sum_{i=1}^n \lambda_i x_i = x_0 \quad (7)$$

$$\sum_{i=1}^n \gamma(S_i - S_j) + m_0 + x_i * m_1 = \gamma(S_j - S_i) \quad (8)$$

Varianza;

$$\sum_{i=1}^n \lambda_i \gamma(S_i - S_0) + m_2 \quad (9)$$

En donde  $N(h)$  es el número total de pares de muestras separados por un intervalo de distancia  $h$ ,  $n$  es el número de observaciones,  $Z_i$  son observaciones (precipitación, humedad relativa) para  $i=1,2,\dots,78$ ,  $\lambda_i$  son los ponderadores,  $\hat{Z}_j$  es la observación a estimar para  $j=1,2,\dots,88956$ ,  $x_i$  son observaciones (variables como elevación, distancia al mar),  $x_j$  es la observación conocida de la coordenada a predecir (variables antes mencionadas),  $m_0, m_1, m_2$  son multiplicadores de Lagrange desconocidos que tenemos que obtenerlos a partir de los datos y  $\gamma$  es el semivariograma teórico ajustado a los semivariogramas experimentales en base a la curva de distancias [7].

### **Parámetros definidos en los semivariogramas teóricos.**

#### **Nugget**

Definido por el símbolo  $C_0$ , también conocida como efecto pepita, se lo atribuye al error de medición de una muestra de coordenadas.

#### **Silo**

Definido por el símbolo  $(C_0+C)$ . Es el valor en donde la distancia se vuelve constante, también es conocido como la varianza asintótica del proceso.

#### **Rango**

Definido por el símbolo  $a$ , y representa la distancia por el cual los valores de la variable dejan de estar correlacionadas.

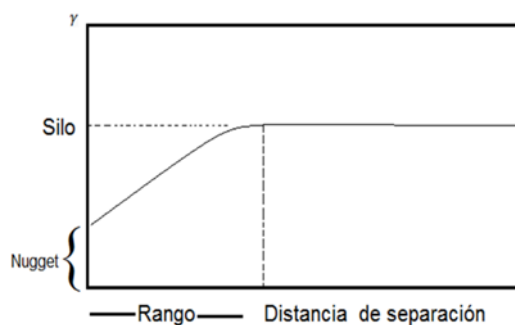


Gráfico 3.1. Representación del Semivariograma experimental.  
Elaborado por: Gabriel Mantilla.

### Semivariogramas teóricos

Modelos	$\gamma(h)$	Rango
Esférico	$C_o + C \left[ \frac{1.5 * \left(\frac{h}{a}\right) + 0.5 \left(\frac{h}{a}\right)^3}{1} \right]$	$0 \leq h \leq a$ $h > a$
Exponencial	$C_o + C \left[ 1 - e^{-\left(\frac{h}{a}\right)} \right]$	$h \Rightarrow 0$
Circular	$C_o + C \left[ \frac{\frac{2h}{\pi a} \sqrt{1 - \left(\frac{h}{a}\right)^2} + \frac{2}{\pi} \arcsin\left(\frac{h}{a}\right)}{1} \right]$	$0 \leq h \leq a$ $h > a$
Pentaesférico	$\frac{15h}{8a} - \frac{5}{4} \left(\frac{h}{a}\right)^3 + \frac{3}{8} \left(\frac{h}{a}\right)^5$	$0 \leq h \leq a$ $h > a$
Gaussiano	$C_o + C \left[ 1 - e^{-\left(\frac{h}{a}\right)^2} \right]$	$h \Rightarrow 0$
K-Bessel	$C_o + C \left[ 1 - \frac{1}{2^{\alpha-1} \Gamma(\alpha)} \left(\frac{h}{a}\right)^{\alpha} k_{\alpha} \frac{h}{a} \right]$	$h > 0$
Logarítmico	$C_o + C \left[ \log\left(\frac{h}{a}\right) \right]$	$h > 0$

Tabla 3.1. Variogramas teóricos.  
Elaborado por: Gabriel Mantilla.

### 3.3.3. Cokriging.

Este método es similar al kriging universal, la diferencia es que es una especie de modelo multivariado, en donde la variable de respuesta se modela con una variable basada en su misma información, y variables explicativas que están correlacionadas espacialmente con la variable de respuesta. Se ajusta con los mismos modelos semivariogramas definidos anteriormente.

Ecuaciones

#### Semivariograma Cruzada.

$$\gamma_{v_1 v_2}(h) = \frac{1}{2N(h)} \sum_{j=1}^{N(h)} \{Z_{v_1}(S_i - S_j)\} \{Z_{v_2}(S_i - S_j)\} \quad \text{para } i = 1, 2, 3, \dots, n \quad (10)$$

#### Modelo Lineal de Corregionalización para dos variables.

$$\gamma_{v_1}(h) = \alpha_0 \gamma_0(h) + \dots + \alpha_m \gamma_m(h)$$

$$\gamma_{v_2}(h) = \beta_0 \gamma_0(h) + \dots + \beta_m \gamma_m(h)$$

$$\gamma_{v_1 v_2}(h) = \delta_0 \gamma_0(h) + \dots + \delta_m \gamma_m(h) \quad (11)$$

Un modelo lineal de Corregionalización asume que todos los semivariogramas simples y cruzados pueden expresarse como una combinación lineal de los mismos modelos teóricos [8]. En la formula como ejemplo solo se consideraron dos variables correlacionadas, donde  $\gamma_{v_1}(h)$  y  $\gamma_{v_2}(h)$  son los semivariogramas simples y  $\gamma_{v_1 v_2}(h)$  es el semivariograma cruzado,  $\gamma_1(h), \gamma_2(h), \gamma_3(h), \dots, \gamma_m(h)$  son los modelos semivariogramas teóricos y  $\alpha_i, \beta_i$  y  $\delta_i$ , con  $i = 1, 2, \dots, m$ , son constantes.

$$\hat{Z}_{v1}(x_0) = \sum_{i=1}^{n1} a_i Z_{v1}(x_i) + \sum_{j=1}^{n2} b_j Z_{v2}(x_j) \quad (12)$$

Sujeto a.

$$\sum_{i=1}^{n1} a_i = 1 \quad (13)$$

$$\sum_{j=1}^{n2} b_j = 0 \quad (14)$$

En donde  $N(h)$  es el número total de pares de muestras separados por un intervalo de distancia  $h$ ,  $\hat{Z}_{v1}(x_0)$  son las observaciones desconocidas a predecir de la variable de interés (precipitación), los  $a_i$  son las ponderaciones que multiplican a los  $Z_{v1}$  observaciones conocidas (precipitación),  $b_i$  son las ponderaciones que multiplican a los  $Z_{v2}$  observaciones conocidas (humedad), también se pueden introducir  $v_p$  variables al modelos, con el supuesto que están fuertemente correlacionadas espacialmente con la precipitación.

#### **Varianza de predicción del método cokriging.**

$$\sigma^2_k = cov(Z_{v1}(x_0), Z_{v2}(x_0)) + u_1 + \sum_{i=1}^n a_i cov(Z_{v1}(x_i), Z_{v1}(x_0)) + \sum_{j=1}^m b_j cov(Z_{v2}(x_j), Z_{v2}(x_0)) \quad (15)$$

### 3.3.4. Kriging Bayesiano.

El kriging Bayesiano es un método que ajusta el semivariograma empírico conocido anteriormente, de modo automático mediante simulaciones, tomando como valor inicial los parámetros de entrada que son probabilidades a priori. En la práctica, uno ajusta de modo manual el semivariograma empírico, apoyándonos en modelos teóricos, a partir de la posición geográfica de los datos, y luego utilizando esta estimación realizamos predicciones en ubicaciones desconocidas, que por teoría deberían estar cerca a los valores observados. El kriging bayesiano asume implícitamente que la semivarianza experimental es el verdadero semivariograma para las zonas a pronosticar <sup>[9]</sup>.

#### En el Marco Bayesiano.

En la práctica, los parámetros utilizados son desconocidos, en la inferencia bayesiana utiliza a estos parámetros como variables aleatorias. Basándonos en esta incertidumbre, con simulaciones, se ajustan mejor estos parámetros, que luego serán usados en las predicciones.

Considerando un vector aleatorio  $Y$  con probabilidad de distribución dado por la función  $pr(y/\vartheta)$ , fijado por un vector de parámetros  $\vartheta$ . Considerando que de una muestra  $Y = y$  puede ser escrita como  $L(\vartheta|y) \equiv pr(y|\vartheta)$ ,  $L$  es una función del parámetro  $\vartheta$  y que es llamada, función de verosimilitud <sup>[10]</sup>.

Considerando la distribución de  $Y$  cómo.

$$(Y|\beta, \sigma^2, \varnothing, \tau^2) \sim N(X\beta; \tau^2 I + V_y(\sigma^2, \varnothing)) \quad (16)$$

La verosimilitud es una función de  $\vartheta = (\beta, \sigma^2, \emptyset, \tau^2)'$

$$L(\vartheta|y) = |\tau^2 I + V_y(\sigma^2, \emptyset)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (y - X\beta)' (\tau^2 I + V_y(\sigma^2, \emptyset))^{-1} (y - X\beta) \right\} \quad (17)$$

La estadística Bayesiana, el parámetro  $\vartheta$  y variable  $Y$  son considerados como cantidades aleatorias distribución de probabilidad conjunta

$pr(y, \vartheta) = pr(y/\vartheta)pr(\vartheta)$ . La información sobre los parámetros externos del modelo de los datos es reflejado en la distribución A priori  $pr(\vartheta)$ . El Teorema de Bayes combina la información A priori y la probabilidad de tal forma que el conocimiento A priori de los parámetros se van actualizando, después de recoger la muestra, siendo esta de la forma.

$$pr(\vartheta|Y) = pr(\vartheta)pr(Y|\vartheta) \quad (18)$$

La distribución  $pr(\vartheta|Y)$  es conocida como la distribución A posteriori en base a la inferencia Bayesiana. Para el Modelo A posteriori.

$$pr(\beta, \sigma^2, \emptyset, \tau^2|y) = |\tau^2 I + V_y(\sigma^2, \emptyset)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (y - X\beta)' (\tau^2 I + V_y(\sigma^2, \emptyset))^{-1} (y - X\beta) \right\} \quad (19)$$

De las distribuciones A prioris que conllevan A Posterioris se llaman distribuciones previas conjugadas. Esas distribuciones previas pueden ser ajustadas de manera sencilla con el poder computacional de la actualidad.

Dos casos extremos para ajustar estas distribuciones son:

- 1) Cuando los parámetros son conocidos
- 2) Cuando el conocimiento de los parámetros es aproximado, los A priori son llamados no informativos, o A priori planos.

Para el último caso se utiliza la estadística Bayesiana clásica para ajustarlos.

Las bases para el pronóstico Bayesiano es también llamado como distribución predictiva  $pr(y_o|y)$ . La distribución predictiva considera a los parámetros desconocidos como el promedio de los parámetros resultantes de la distribución condicional  $pr(y_o|y, \vartheta)$ , con pesos dados por la distribución A posteriori por los parámetros del modelo  $pr(\vartheta|y)$ .

$$pr(y_o|y) = \int pr(y_o, \vartheta|y) d\vartheta \quad (20)$$

$$pr(y_o|y) = \int pr(y_o|y, \vartheta) * pr(\vartheta|y) d\vartheta \quad (21)$$

La distribución predictiva también se puede escribir en base a su función de verosimilitud dada anteriormente.

$$pr(y_o|y) = \int \frac{pr(y_o|y, \vartheta) * pr(\vartheta)}{\int pr(y|\vartheta) pr(\vartheta) d\vartheta} d\vartheta \quad (22)$$

En la práctica, lo que realiza esta metodología, es primero estimar un semivariograma empírico a partir de un subconjunto de los coordenadas observados, teniendo este semivariograma como base, se simulan los puntos nuevos a predecir dentro del mismo subconjunto, se estiman semivariogramas nuevos a partir del subconjunto pronosticado, este proceso se repite un número determinado de veces, hasta terminar de pronosticar todos los puntos a pronosticar.



### 3.3.5. Suma Cuadrática del Error.

La suma de los errores al cuadrado representa la variación aleatoria del error en cada periodo de medición, en la teoría se encuentra el error cuadrático medio <sup>[11]</sup>, siendo suma cuadrática de los errores dividida para el numero de observaciones, la suma cuadrática de los errores esta expresada de la siguiente forma.

$$SCE = \sum_{i=1}^n (Z_{obs} - \hat{Z}_{est})^2 \quad (23)$$

### 3.3.6. Mínimos Cuadrados Ponderados.

Esta medida de ajustes es utilizada para medir que modelo paramétrico se ajusta mejor al semivariograma empírico, de tal manera que a las observaciones que se encuentran más lejos, se tendrán menor peso de ponderación al aumentar su varianza, que las observaciones que se encuentren más cerca tendrán mayor ponderación al disminuir su varianza <sup>[12]</sup>.

$$Min(Q) = \sum_{i=1}^n \frac{N(h)}{\gamma_{\theta}(h_i)^2} (\hat{\gamma}(h_i) - \gamma_{\theta}(h_i))^2 \quad (24)$$

En donde  $N(h)$  es el número total de pares de muestras separados por un intervalo de distancia  $h$ ,  $\hat{\gamma}$  es el semivariograma estimado,  $\gamma_{\theta}$  es el semivariograma empírico verdadero, con valores inicial dados (Nugget, Silo, Rango), itera un número determinado de veces hasta encontrar el modelo óptimo.

### **3.3.7. Validación Cruzada.**

En este enfoque se utilizó la validación cruzada (Leave one out) dejando uno fuera, este algoritmo fue propuesto por Craven y Wahba en 1979<sup>[13]</sup>, y aplicación del método lo hicieron K.W. Lau y Q.H. Wu<sup>[14]</sup> para medir el ancho de banda óptimo de un proceso computacional. Este método consiste en evaluar los resultados de un análisis estadístico, en nuestro enfoque mide el nivel de predictibilidad de los interpoladores propuestos en las coordenadas observadas. Del total de observaciones  $n$  se saca un dato, siendo esa coordenada el dato de prueba, y los demás datos son los datos de entrenamiento, con los datos de entrenamiento se predice en la posición del punto de prueba, se itera  $n$  veces cada modelo.

## CAPITULO IV

### 4. RESULTADOS

A continuación se presentara los resultados de las metodologías anteriormente expuestas. En la metodología especializada se mostrada un resultado por cada uno de los 12 meses del año.

#### 4.1. Resultados Metodología Descriptiva.

En el siguiente gráfico podemos evidenciar el comportamiento pluviométrico mensual de todas las estaciones involucradas en el cálculo, así observamos la dispersión promedia mensual entre estaciones que nos ayudara a conocer mejor la variable de estudio.

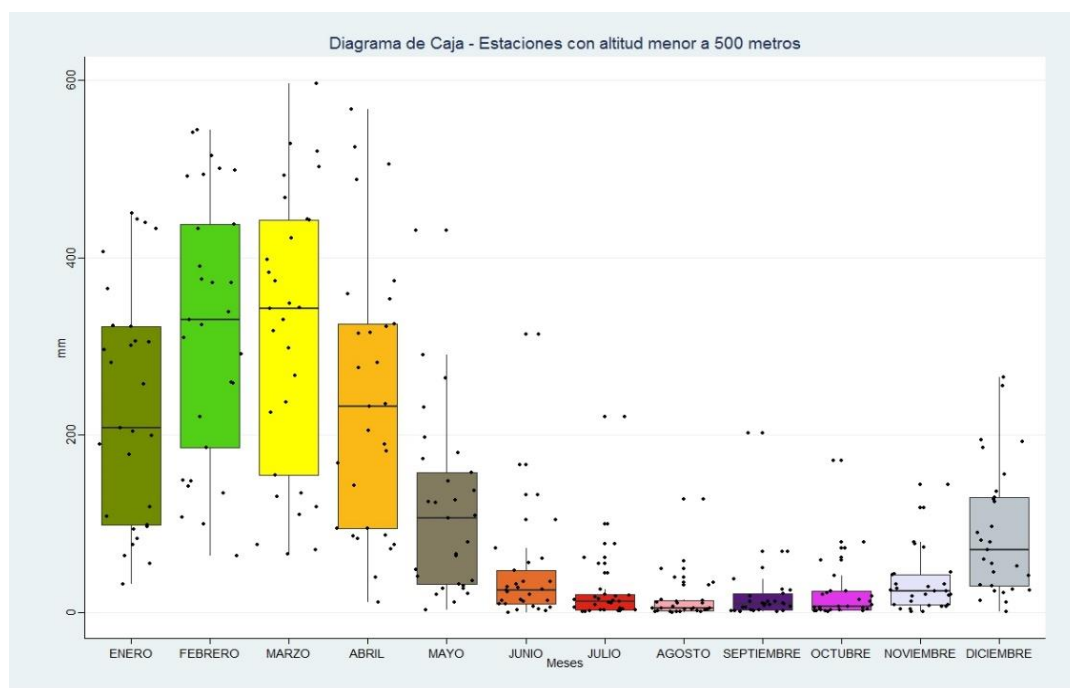


Gráfico 4.1. Diagrama de Cajas para las estaciones con una altitud menor a 5 m  
Elaborado por: Gabriel Mantilla.

El grafico No.2 nos muestra los cuartiles para los 12 meses del año, para las 29 estaciones que están a una altitud no mayor a 500 metros, podemos ver que la época lluviosa empieza a partir de Diciembre y termina en el mes de mayo para luego entrar a la época seca, los datos están dispersos alrededor de las cajas y también presentan valores atípicos máximos y mínimos. La siguiente tabla muestra las estadísticas descriptivas de las 29 estaciones con una altura menor a los 500 metros.

Meses	Máximo	Mediana	Mínimo	Media
Enero	San Juan La Mana	Machala UTM	Santa Elena Univ.	232.1
	449.9	208.3	21.7	
Febrero	Caluma	Guayaquil Univ. Estatal	Santa Elena Univ.	321.6
	544.2	329.9	63.8	
Marzo	San Juan La Mana	Vinces Inamhi	El Suspiro	315.2
	596.4	343	65.8	
Abril	San Juan La Mana	Primo Banano	Santa Elena Univ.	241.5
	567.6	231.8	11.1	
Mayo	Cayapas	Hacienda Taura	Santa Elena Univ.	113.9
	430.9	105.9	2.2	
Junio	Cayapas	Olmedo Manabí	Santa Elena Univ.	44.6
	313.2	25.4	0.1	
Julio	Cayapas	Olmedo Manabí	Zapotillo	26.6
	220.7	21.6	0.3	
Agosto	Cayapas	Vinces Inamhi	Zapotillo	15.7
	127.9	4.9	0.1	
Septiembre	Cayapas	Ingenio San Carlos	Hacienda Taura	22
	202.1	9.6	0.3	
Octubre	Cayapas	Vinces Inamhi	Santa Elena Univ.	23.4
	171.6	6.9	1.1	
Noviembre	El Carmen	Chone Univ. Católica	Hacienda Taura	32.8
	144.5	24.3	0.9	
Diciembre	El Carmen	Chone Univ. Católica	Santa Elena Univ.	90.3
	264.8	70.5	0.3	

Tabla 4.1. Estadísticas Descriptivas, estaciones con altitud menor a 500 metros.  
Elaborado por: Gabriel Mantilla.

De la tabla No V. podemos ver que el promedio más alto se dio en el mes de Febrero con 321.6 milímetros, y el máximo registrado se dio en el mes de Marzo de la estación de San Juan en la Mana.

La estación de San Juan la Mana es la estación con el promedio más alto de precipitación registrado en los meses de Enero, Marzo y Abril. La estación de Cayapas registro el promedio máximo de precipitación desde Mayo a Octubre. La estación de la Universidad de Santa Elena registro el promedio mínimo de precipitación para los meses de Enero, Febrero, Abril, Mayo, Junio, Octubre y Diciembre.

El siguiente grafico tiene en resumen el comportamiento de las 49 estaciones que se encuentran a una altitud mayor a 500 metros, del mismo modo la época lluviosa empieza a partir del mes de noviembre, terminándose en el mes de mayo, y la época seca empieza en el mes de junio, terminando en octubre.

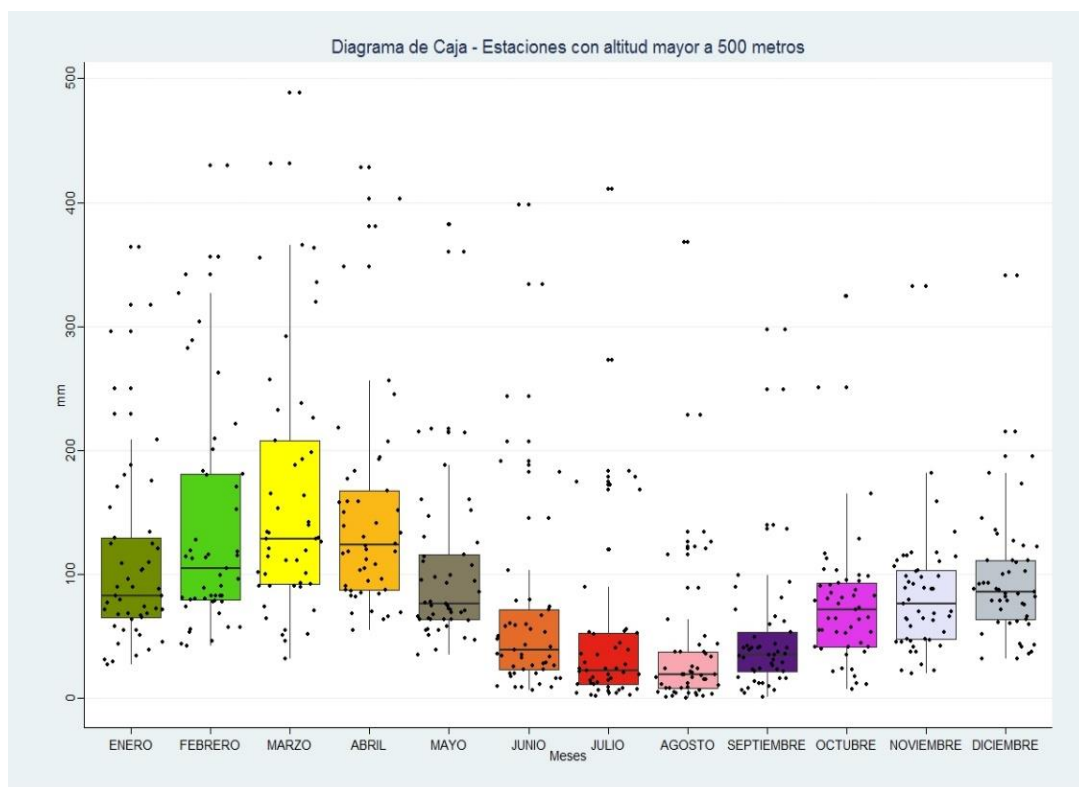


Gráfico 4.2. Diagrama de Cajas para las estaciones con una altitud mayor a 5 m  
Elaborado por: Gabriel Mantilla.

La siguiente tabla muestra las estadísticas descriptivas de los promedio de cada estación por mes, para las estaciones 49 con una altura mayor a los 500 metros.

Meses	Máximo	Mediana	Mínimo	Media
Enero	El Corazón	Nayon Granja Santa Ana	San Vicente de Pusir	110.9
	364	82.9	27	
Febrero	El Corazón	Nayon Granja Santa Ana	Pedro Fermín Cevallos	140.6
	429.5	105.1	42.2	
Marzo	El Corazón	Malacatos	San Vicente de Pusir	165.5
	488.5	128.5	32.2	
Abril	Sangay	Nambacola Colegio	San Vicente de Pusir	148.9
	428.5	123.5	54.9	
Mayo	Sierra Azul	Patate	Chunchi	105.3
	382.5	76.5	35.2	
Junio	Sierra Azul	Otavaló	Celica	70.4
	398.2	39	6.6	
Julio	Sierra Azul	Ingincho	El Salado Predesur	53.8
	410.4	22.2	1.9	
Agosto	Sierra Azul	Olmedo Pichincha	Celica	41.9
	368.2	18.9	0.6	
Septiembre	Sierra Azul	El Angel	Pazul Colegio Rodríguez	49.8
	297.2	35.3	1.1	
Octubre	Sangay	Paute	Pazul Colegio Rodríguez	76.6
	324.3	71.7	7.4	
Noviembre	Sangay	Amaluza Inamhi	Pazul Colegio Rodríguez	83.3
	332.5	76	19.9	
Diciembre	Sangay	Cahuasqui FAO	Chunchi	96.9
	341.3	86	31.9	

Tabla 4.2. Estadísticas Descriptivas, estaciones con altitud mayor a 500 metros.  
Elaborado por: Gabriel Mantilla.

La tabla anterior muestra que la estación de El Corazón en el mes de Marzo registra el máximo promedio de precipitación con 488.5 milímetros, el mínimo lo registra la estación de Celica con 0.6 milímetros en el mes de Agosto y la media más alta se dio en el mes de Marzo con 165 milímetros.

La estación del Corazón registro en promedio el máximo de precipitación para los meses de Enero, Febrero y Marzo. La Estación de Sierra Azul registro en máximo el máximo de precipitación para los meses de Mayo a Septiembre. La estación de San Vicente de Pusir registro en promedio el mínimo de precipitación para los meses de Enero, Marzo y Abril, del mismo modo la estación de Pazul Colegio Rodríguez registro el promedio mínimo de precipitación para los meses de Septiembre, Octubre y Noviembre.

#### **4.2. Resultados Metodología Inferencial.**

A continuación mostrados 4 gráficos, representando los gráficos Quantil Quantil y el histograma de frecuencia de cada mes, para la variable precipitación y la humedad relativa, con el estadístico p de la Prueba de Shapiro Wilk.



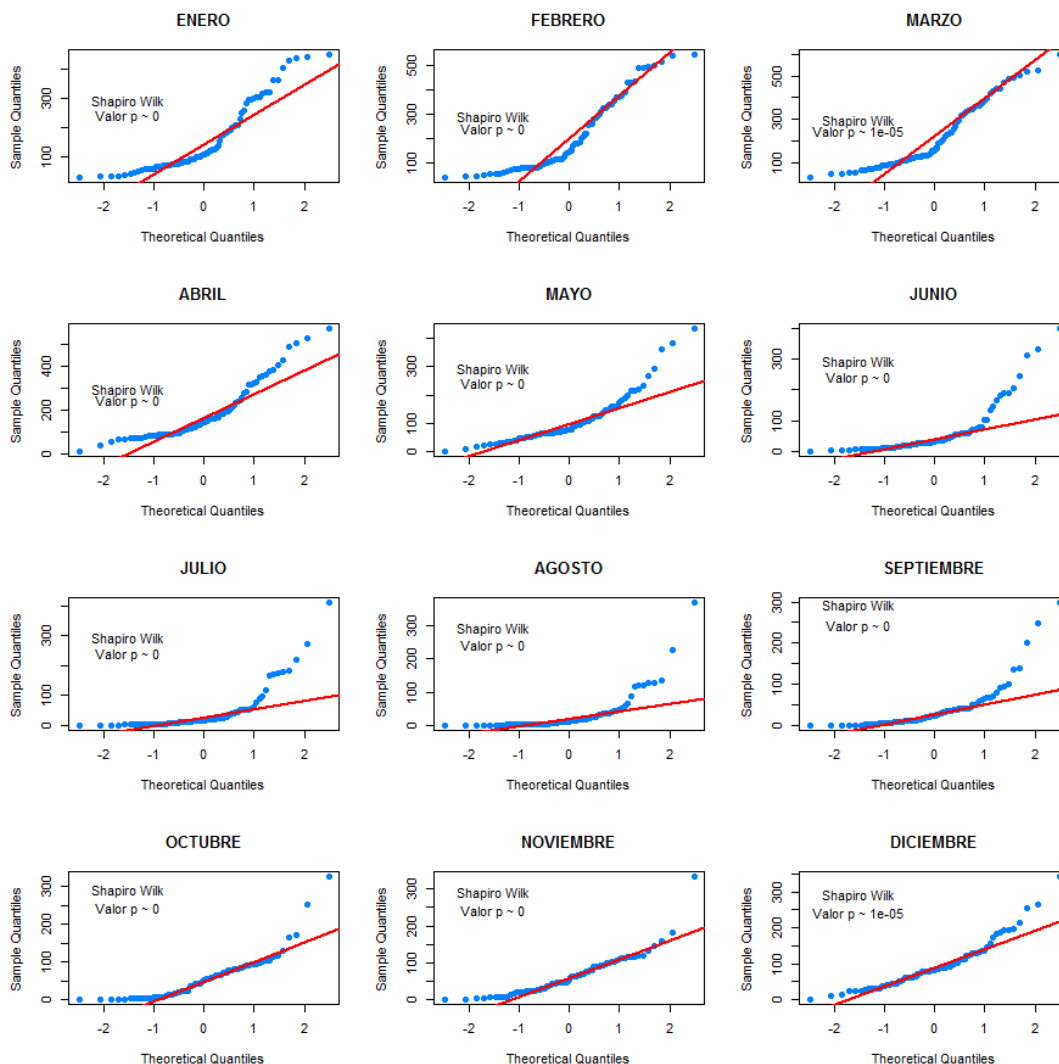


Grafico 4.3. Gráficos Quantil Quantil, Precipitación.  
Elaborado por: Gabriel Mantilla.

El grafico No.4 nos muestra que en ningún de los meses existe igual tendencia entre los puntos azules y la línea roja, también el valor  $p$  siempre es aproximadamente 0, dándonos una idea de que los datos de precipitación no provienen de distribuciones normales. En los valores mayores 0 vemos que los puntos azules se dispersan mucho más de la diagonal de normalidad.

A continuación mostramos los histogramas de frecuencia de cada mes para la variable Precipitación.

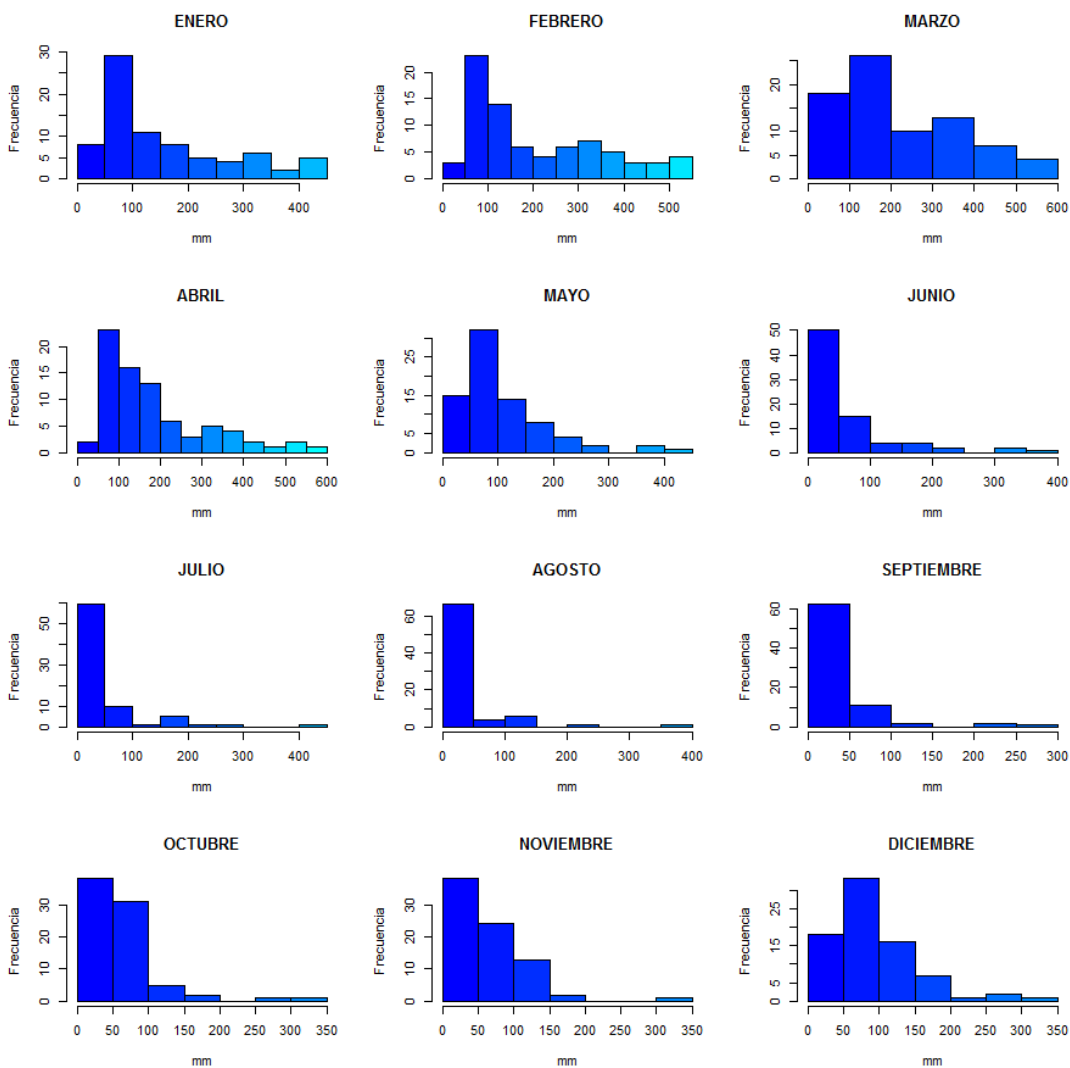


Grafico 4.4. Histogramas de Frecuencia, Precipitación.  
Elaborado por: Gabriel Mantilla.

Podemos ver que los datos no provienen de una distribución Normal dado que los bloques presentan un sesgo positivo centrado en 0, si se aplica una transformación simple a los datos para lograr normalidad, se podrían ajustar mejor los modelos a los

semivariogramas teóricos. Pero en ninguno de los casos los datos presentan normalidad, necesariamente nos vemos forzados a no usar kriging ordinario, donde se sugiere que los datos pertenezcan de una distribución normal.

A continuación mostramos los gráficos Cuantil Cuantil para la variable humedad relativa de los 12 meses del año.

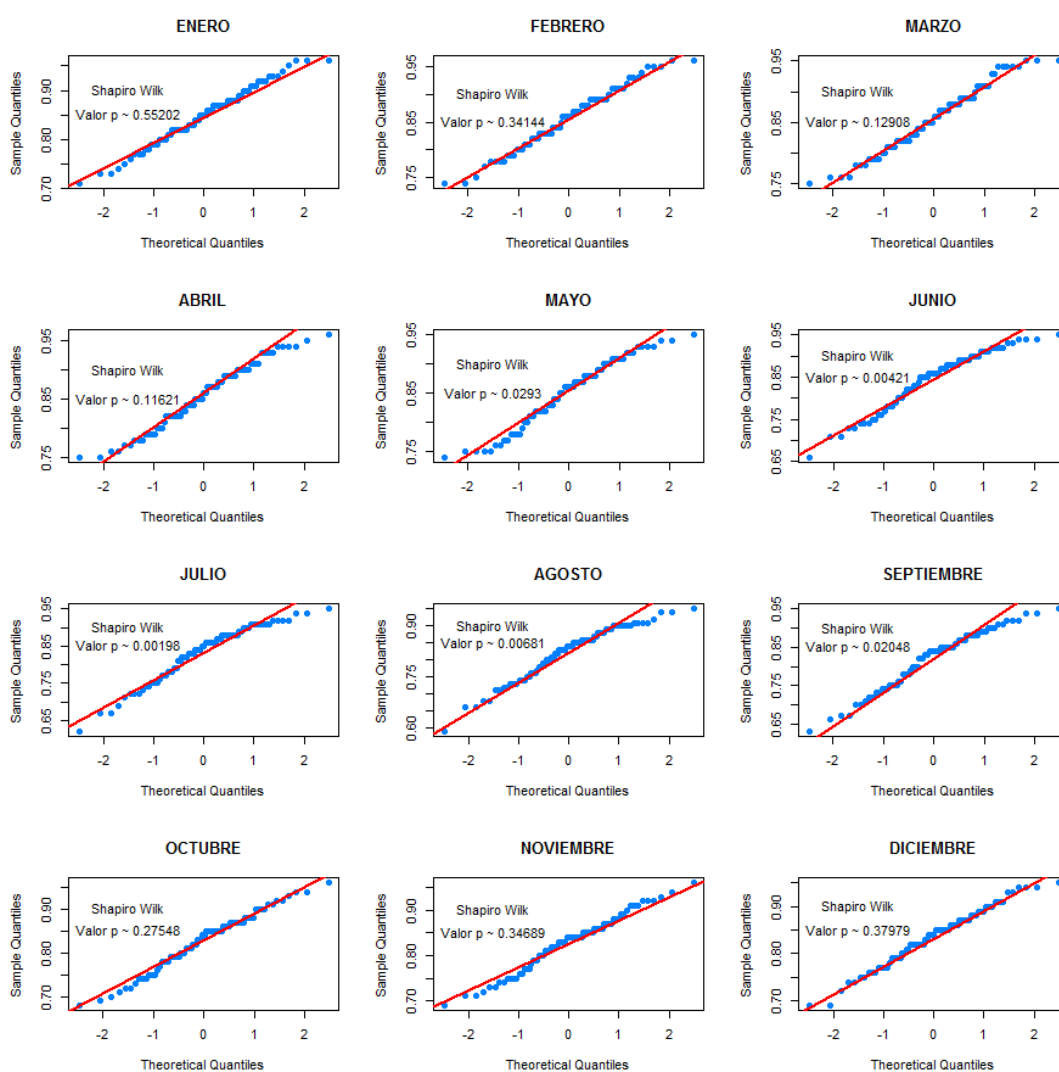


Grafico 4.5. Gráficos Cuantil Cuantil, Humedad Relativa.  
Elaborado por: Gabriel Mantilla.

El grafico No. 6, nos muestra que la variable humedad relativa presenta normalidad en sus datos desde el mes de octubre hasta abril, con un valor p mayor a 0.1, de los meses de mayo a septiembre el valor p está en un rango de 0 a 0.1, presentado datos con poca presencia de normalidad. En todos los casos los puntos azules se asemejan mucho a la línea roja.

A continuación mostramos los histogramas de frecuencia de cada mes para la variable Humedad Relativa expresada de 0 a 1.

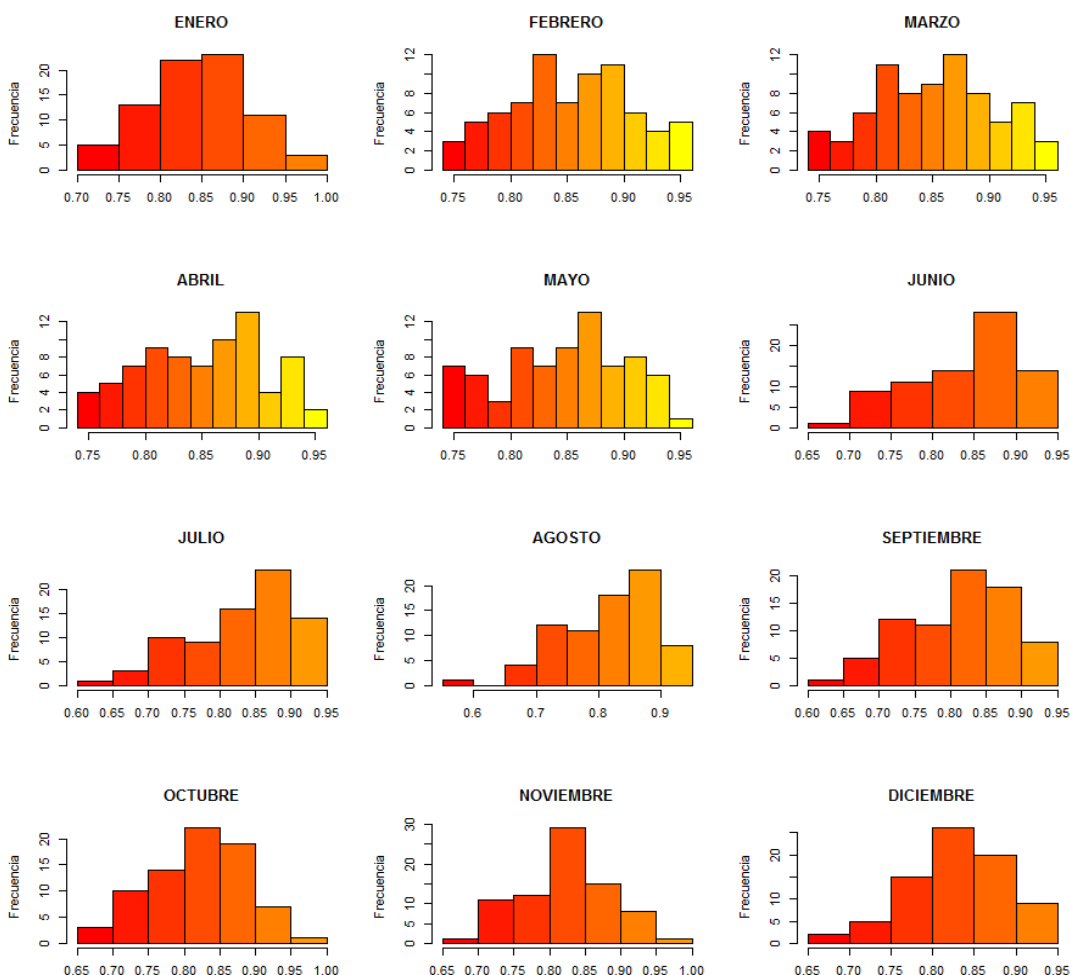


Grafico 4.6. Histogramas de Frecuencia, Humedad Relativa. Elaborado por: Gabriel Mantilla.

El gráfico No.7 nos muestra que de los meses de octubre a abril el sesgo es simétrico alrededor de 0.8 presentando normalidad en los datos, mientras que en los meses de mayo a septiembre el sesgo es positivo alrededor de 0.9.

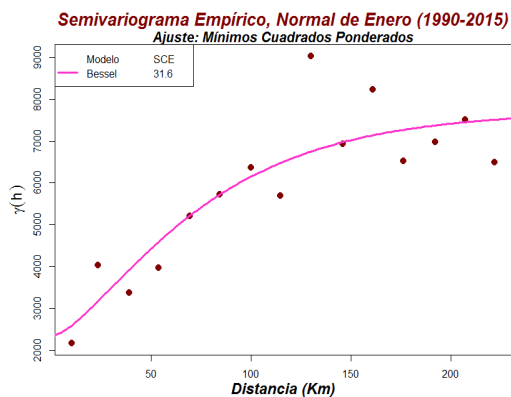
### **4.3. Resultados Metodología Especializada**

Los siguientes gráficos, tablas y mapas, mostrarán el resultado de implementar las cuatro metodologías antes descritas, con sus medidas de ajuste, también está incluido el análisis de validación cruzada para cada mes. Las cuatro metodologías por cada gráfico son; el cokriging el mapa de la izquierda superior, kriging bayesiano el de la esquina superior derecha, ponderación de distancias inversas el de la esquina inferior izquierda, y el kriging universal el de la esquina inferior derecha. La gama de colores para los mapas de precipitación representa las precipitaciones de menor a mayor desde el color blanco al azul. Mientras que la gama de colores para los mapas de varianza representa la variabilidad de menor a mayor desde el color azul hasta el rojo. Los números representan las coordenadas geográficas angulares en latitud y longitud donde se encuentra el Ecuador.

En meteorología es conocido el término “Normal”, al promedio mensual de una serie de tiempo en un periodo determinado mayor a 10 años, en los siguientes gráficos usaremos la palabra Normal, para definir el promedio de cada mes, con un periodo de tiempo desde el 1990 hasta el 2015.

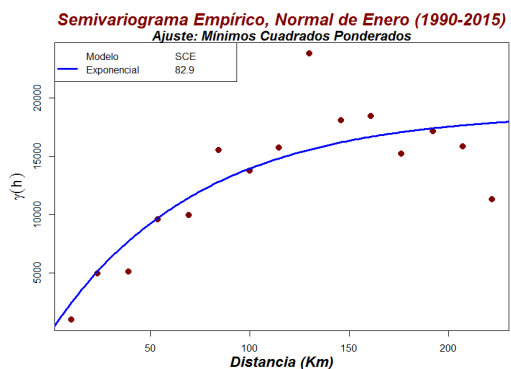
Resultados del mes de Enero

Cokriging



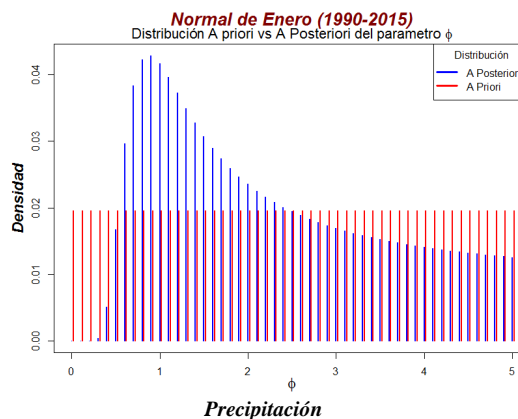
*Precipitación + Humedad\_Relativa ~ Elevación + Distancia\_Mar + Elevación\*Distancia\_Mar*

IDW

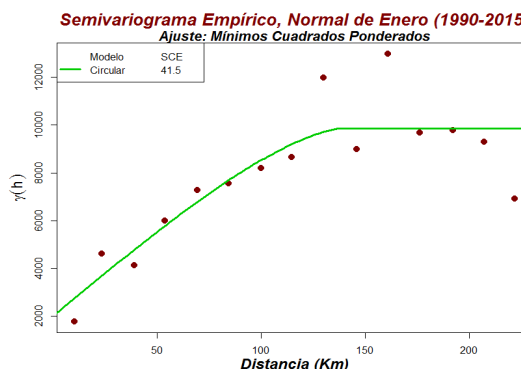


*Precipitación~I*

Kriging Bayesiano



Kriging Universal



*Precipitación ~ Elevación + Distancia\_Mar*

Gráfico 4.7. Medidas de Ajuste, Normal de Enero.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	18795.35	0	73.67	Exponencial	82.9
Kriging Universal	7855.29	2004.494	138.75	Circular	41.5
Cokriging	5365.72	2348.214	51.31	K-Bessel	31.6

Tabla 4.3. Parámetros de los Modelos para Enero.

Elaborado por: Gabriel Mantilla.

En el método del kriging Bayesiano, se usó un vector de probabilidades con un valor constante de 0.02 como distribución a priori, para cincuenta y un  $\emptyset$ .

De la Tabla No. VII, podemos ver los parámetros óptimos con los que los modelos se ajustan mejor, disminuyendo la suma cuadrática de los errores con respecto a los demás modelos, en el caso del IDW se usó el modelo exponencial para aproximar la varianza del método, dado que en la librería “spatstat” no ajusta por default la varianza. A continuación mostramos la validación cruzada del mes de Enero, iterando cada modelo descrito anteriormente, con un punto de prueba afuera, para luego comparar el pronóstico con su valor observado real, teniendo el cokriging el menor SCE.

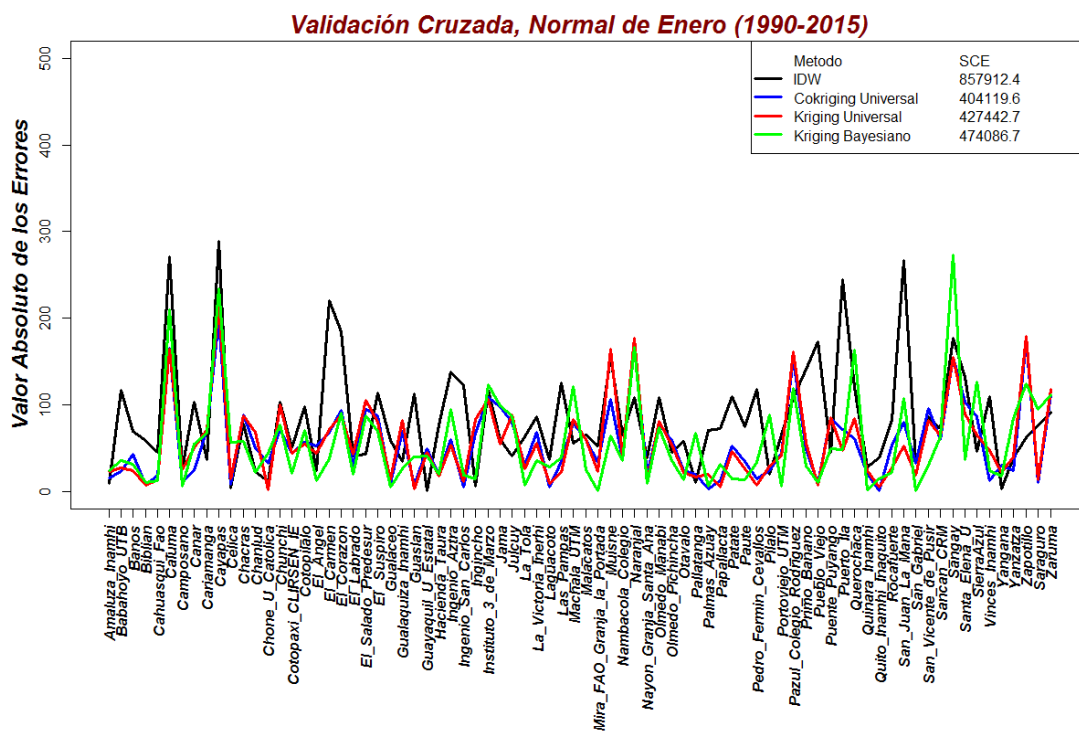


Grafico 4.8. Validación cruzada Enero.  
Elaborado por: Gabriel Mantilla.

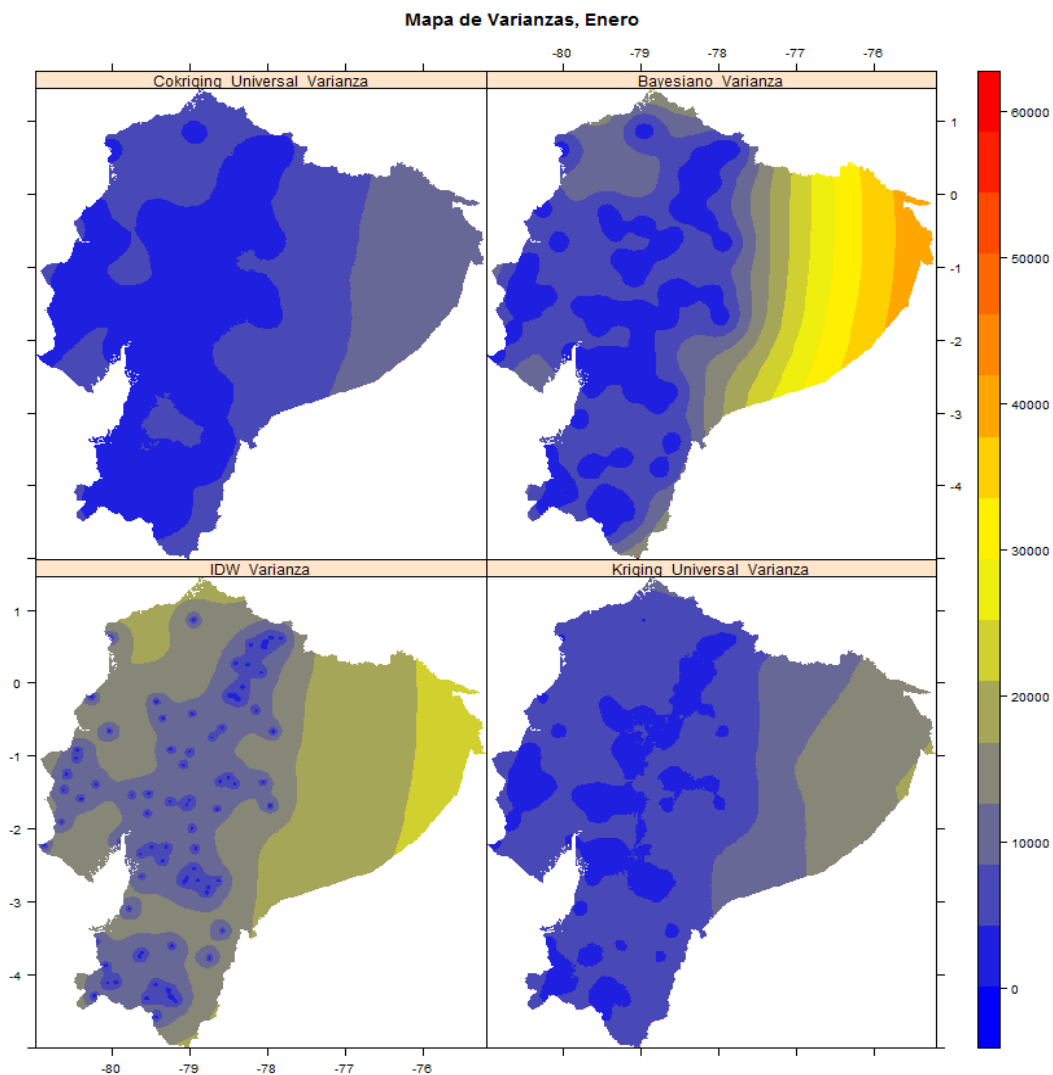


Ilustración 4.1. Varianzas de los interpoladores para Enero.  
Elaborado por: Gabriel Mantilla.

De la ilustración 4.1 vemos que el cokriging tiene una varianza más baja incluso en donde no tiene puntos observados, un comportamiento similar presenta el kriging universal, teniendo su varianza a incrementar en el oriente. El kriging bayesiano muestra también varianzas altas en donde no hay observaciones, superior a los otros interpoladores, el IDW muestra varianza pequeñas en los puntos observados, a medida que uno se aleja la varianza incrementa.



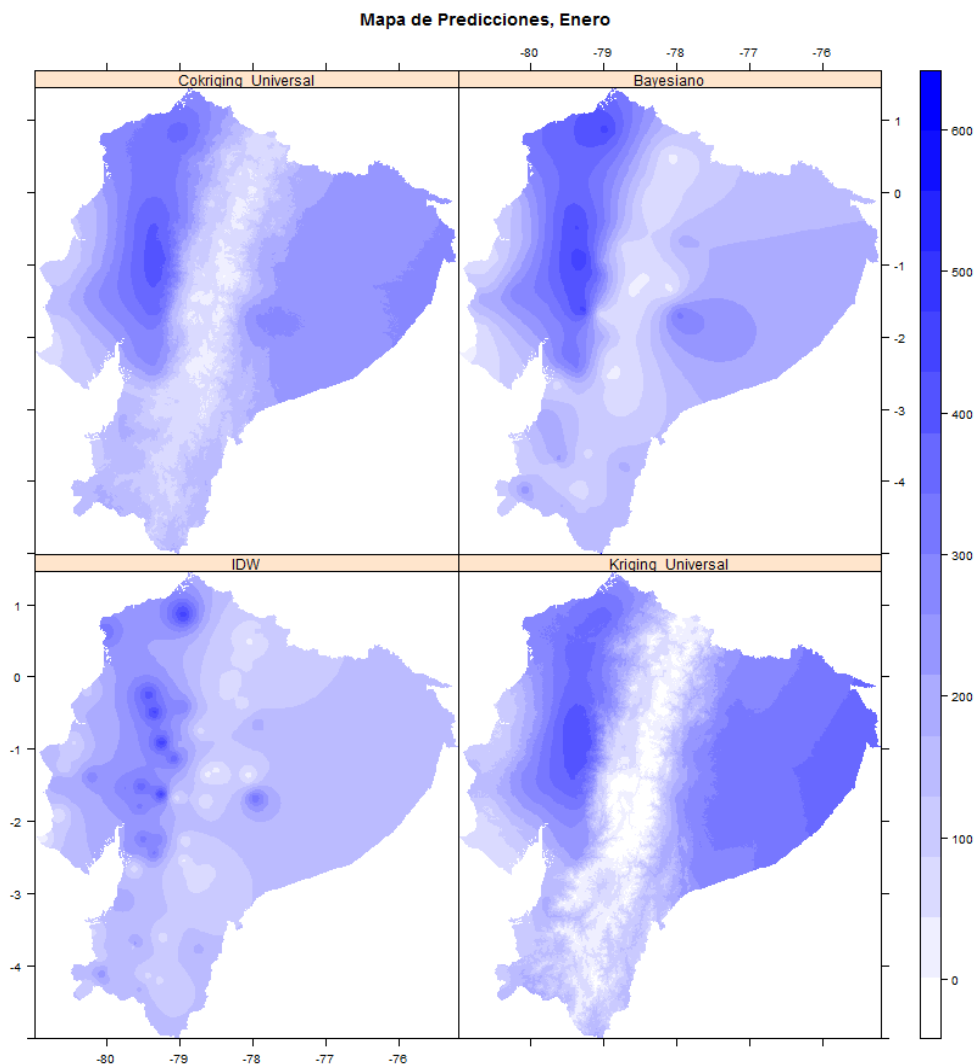
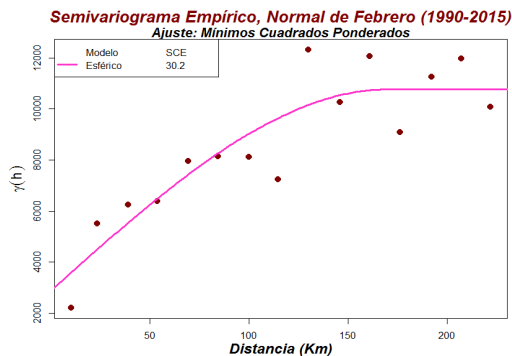


Ilustración 4.2. Interpolación de los métodos para Enero.  
Elaborado por: Gabriel Mantilla.

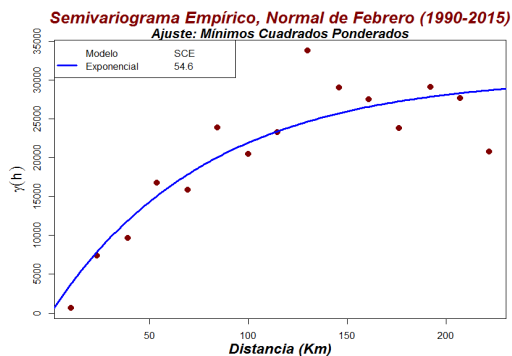
Vemos de la ilustración 4.2 que el cokriging no suaviza mucho los valores observados y detalla mejor las zonas en donde se encuentran las estaciones, el IDW presenta los valores reales utilizándolo como base para los otros interpoladores, el kriging universal tiene un alto nivel de detalle dándonos mayores precipitaciones en el Oriente, El kriging bayesiano sin ninguna variable explicativa hace un buen trabajo similar a los kriging, solo ocupando un mayor nivel computacional.

Resultados del mes de Febrero

Cokriging

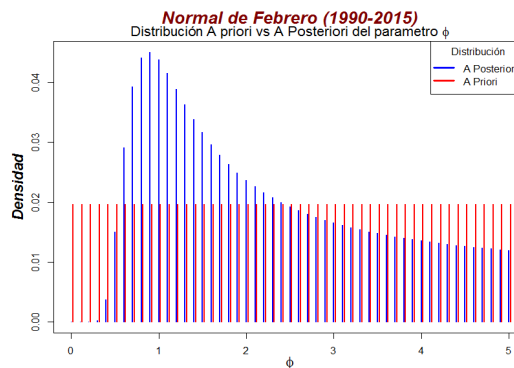


Precipitación + Humedad\_Relativa ~ Elevación + Distancia\_Mar + Elevación\*Distancia\_Mar  
IDW

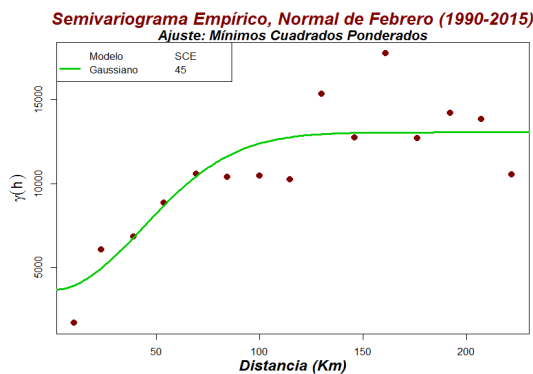


Precipitación~1

Kriging Bayesiano



Precipitación  
Kriging Universal



Precipitación ~ Elevación +Distancia\_Mar

Grafico 4.9. Medidas de Ajuste, Normal de Febrero.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	30523.8	0	78.94	Exponencial	54.6
Kriging Universal	9368.33	3679.531	61.48	Gaussiano	45
Cokriging	7909.7	2860.2	170.3	Esférico	30.2

Tabla 4.4. Parámetros de los Modelos para Febrero.

Elaborado por: Gabriel Mantilla.

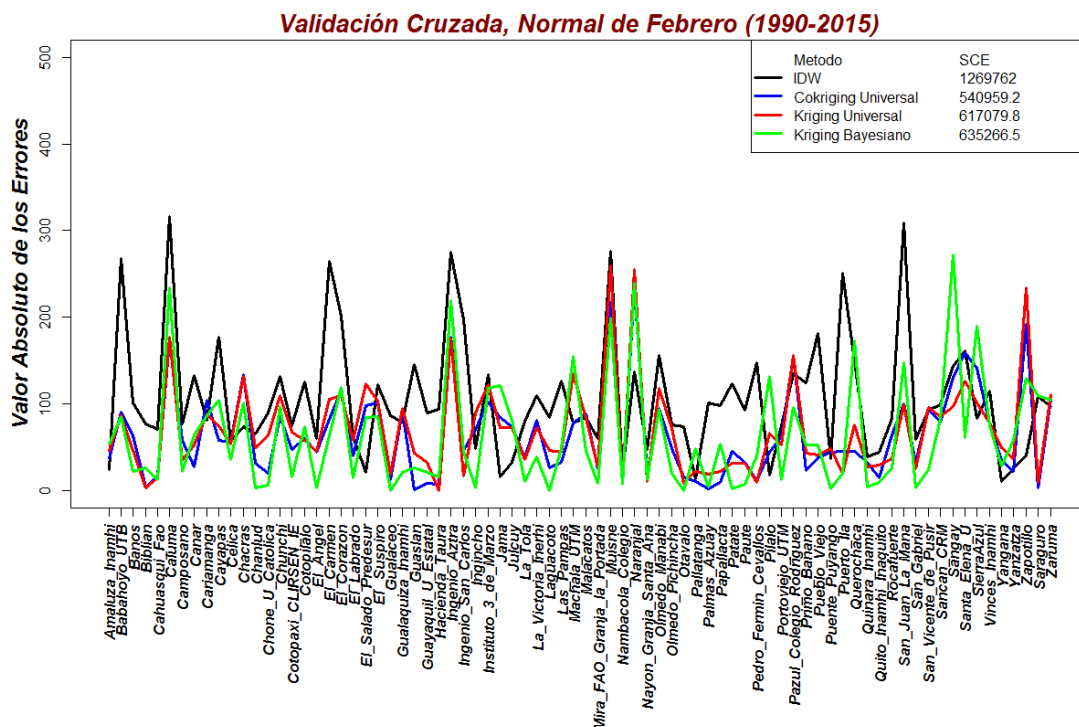


Gráfico 4.10. Validación cruzada Febrero.  
Elaborado por: Gabriel Mantilla.

El gráfico 4.10, muestra que el cokriging (línea azul) tiene el SCE más bajo con respecto a los otros modelos con 540.959, y el IDW (línea negra) el SCE mas alto con 1'269.762 para el mes de febrero,

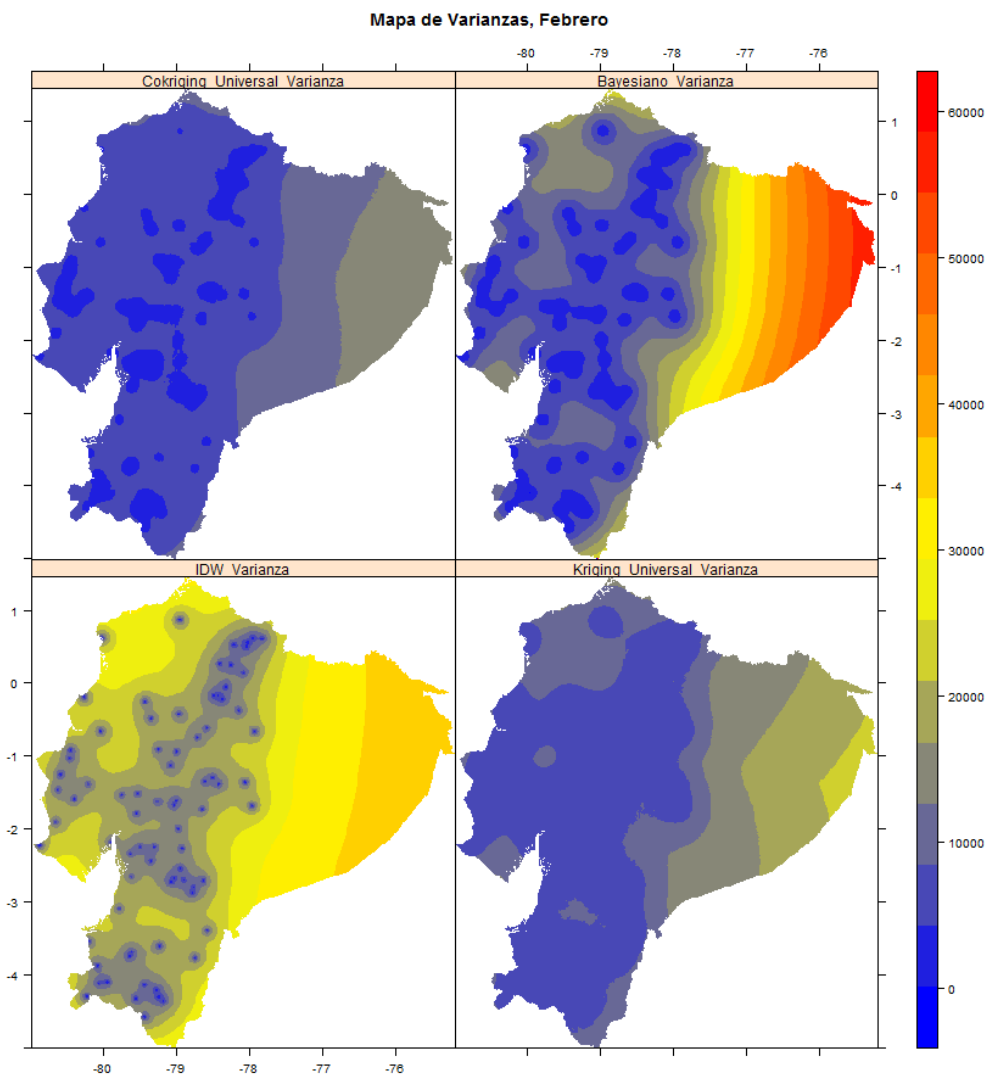


Ilustración 4.3. Varianzas de los interpoladores para Febrero.  
Elaborado por: Gabriel Mantilla.

La ilustración 4.3 nos muestra que las varianzas aumentan con respecto al mes anterior, esto es razonable debido a la variabilidad del mes de Febrero es más alta, debido a que es un mes más lluvioso.

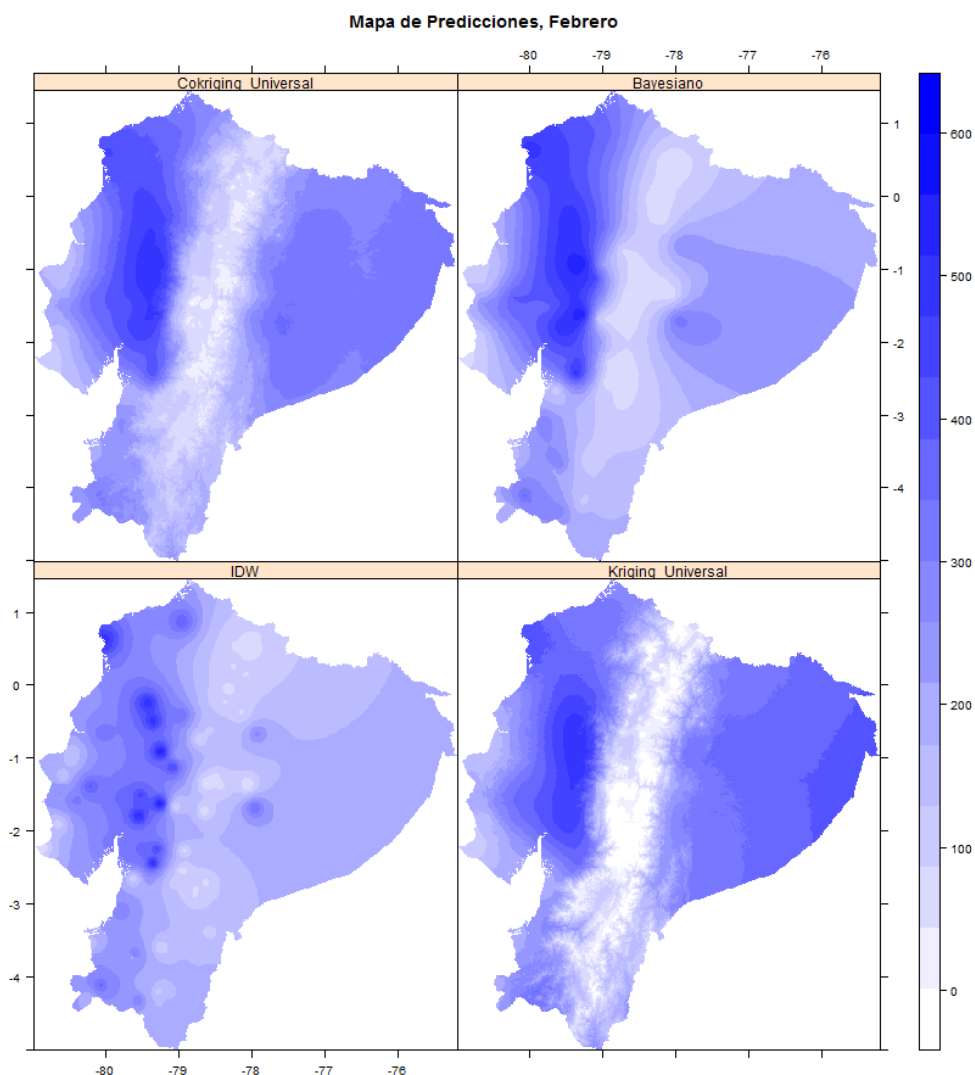


Ilustración 4.4. Interpolación de los métodos para Febrero.  
Elaborado por: Gabriel Mantilla.

La ilustración 4.4 muestra que el kriging universal presenta una gran diferencia con el respecto al kriging bayesiano, teniendo zonas más secas en la zona interandina, y zonas más húmedas en la parte oriental. También existe similitud en la interpolación con respecto al mapa del cokriging.

Resultados del mes de Marzo

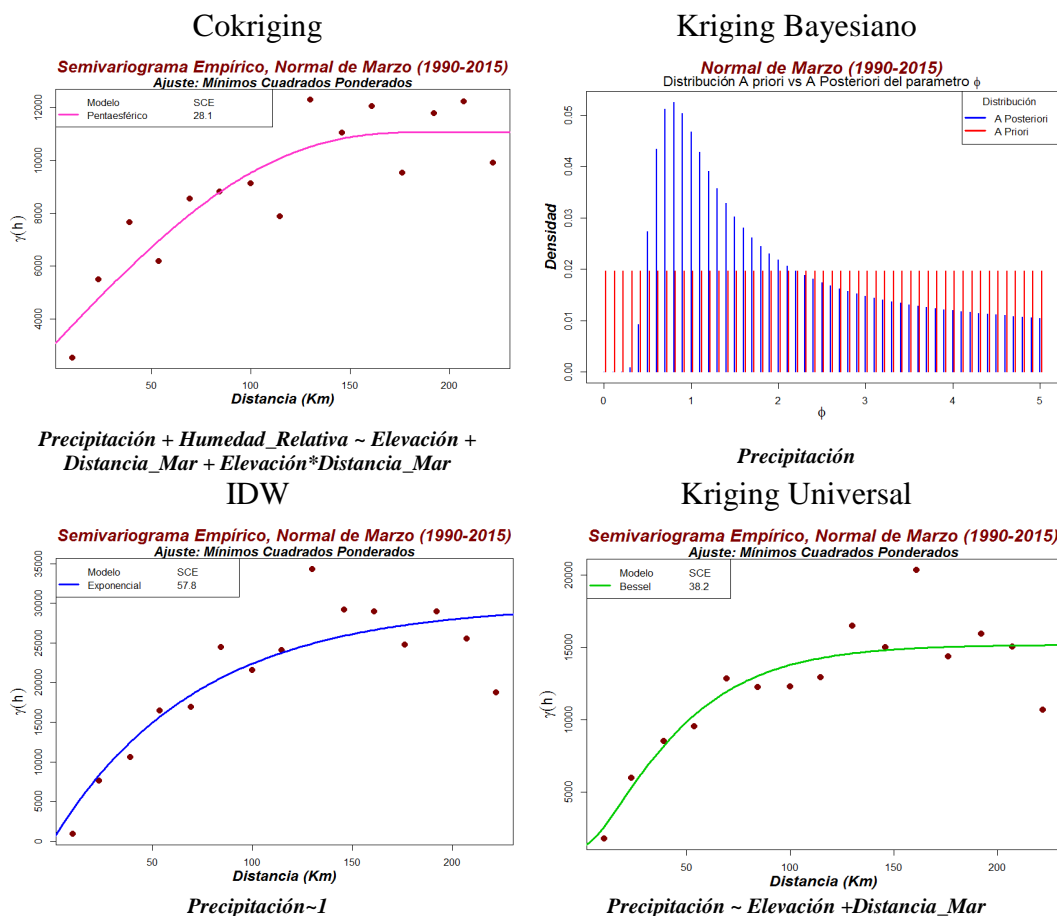


Gráfico 4.11. Medidas de Ajuste, Normal de Marzo.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	29838.87	0	72.10	Exponencial	57.8
Kriging Universal	13913.07	1290.189	31.224	K-Bessel	38.2
Cokriging	8135.92	2944.72	193.74	Pentaesférico	28.1

Tabla 4.5. Parámetros de los Modelos para Marzo.

Elaborado por: Gabriel Mantilla.

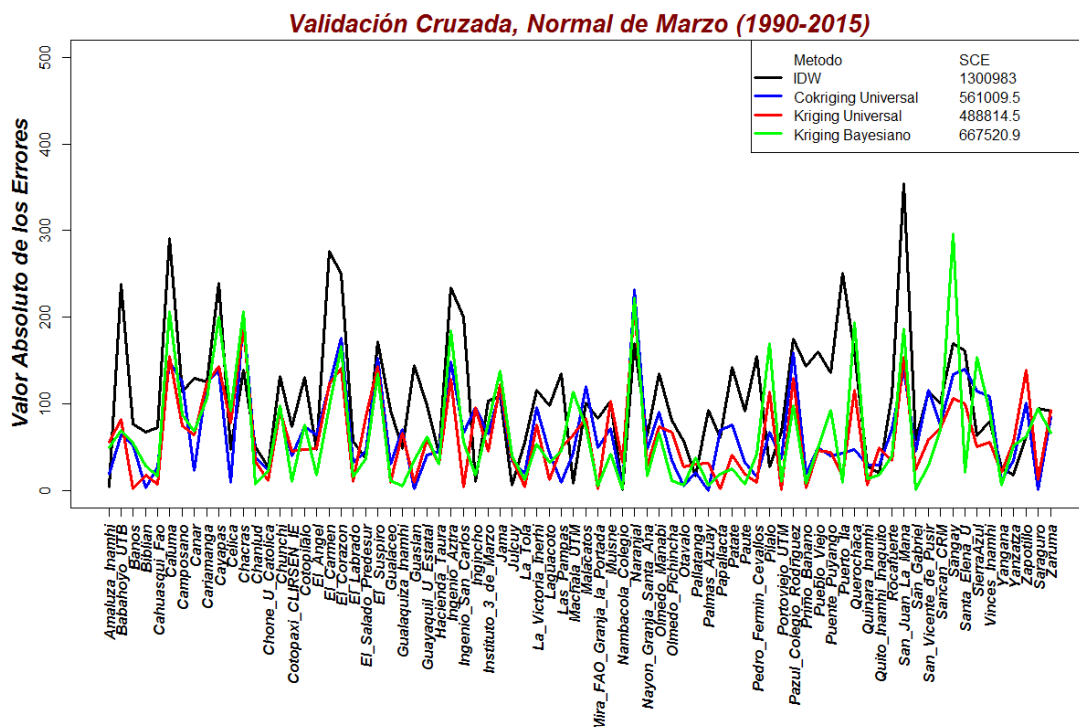


Gráfico 4.12. Validación cruzada Marzo.  
Elaborado por: Gabriel Mantilla.

El modelo que mejor se desempeñó fue el Pentaesférico para el método de cokriging con un SCE de 28.1, siendo este el menor, le sigue el K-Bessel del método de kriging universal. De igual manera en la validación cruzada el kriging y el cokriging presentan menor SCE con respecto a los otros dos métodos.

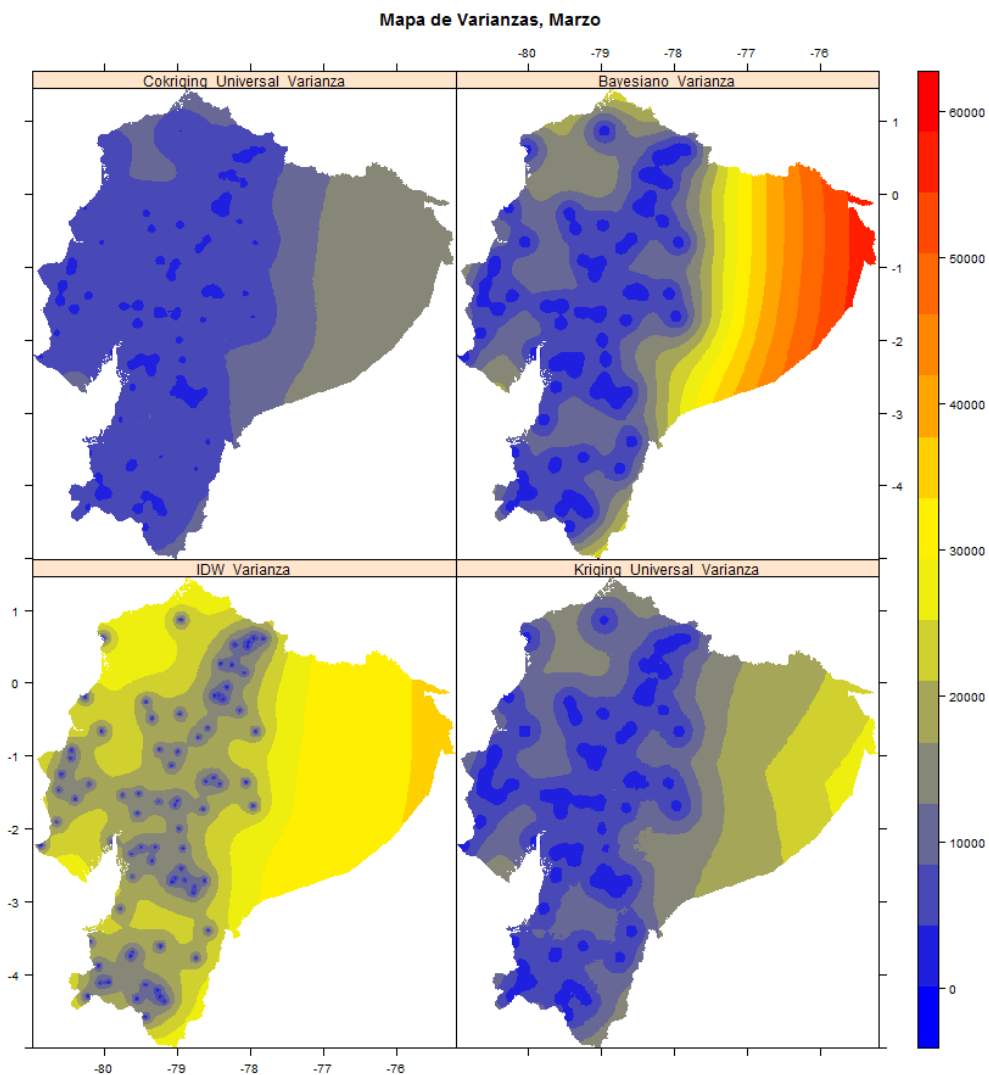


Ilustración 4.5. Varianzas de los interpoladores para Marzo.  
Elaborado por: Gabriel Mantilla.

El método de kriging universal presenta una varianza un poco más alta con respecto al mes anterior debido a que la variabilidad de lluvias aumenta en el mes de Marzo. El método del kriging bayesiano estimó una varianza más alta en la zona oriental, en donde no hay estaciones.



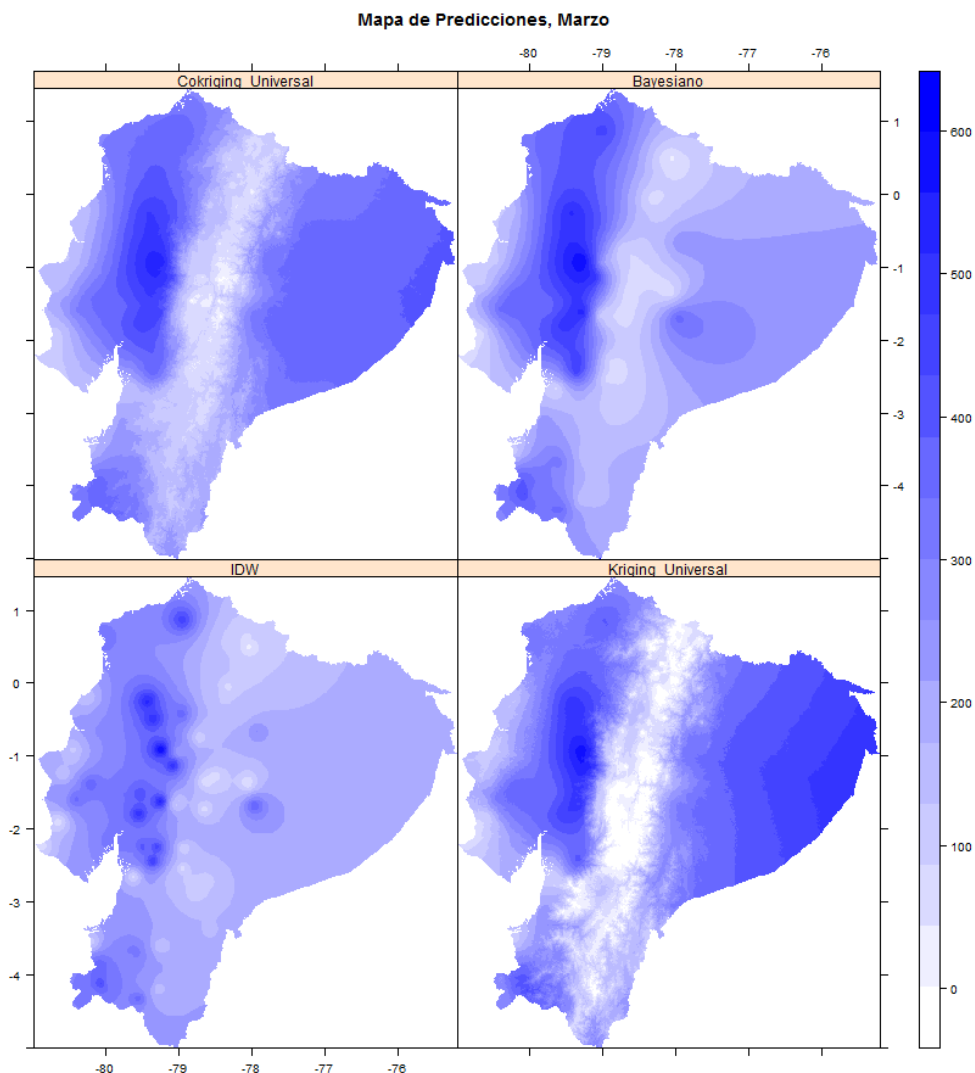


Ilustración 4.6. Interpolación de los métodos para Marzo.  
Elaborado por: Gabriel Mantilla.

La ilustración 4.6 muestra la interpolación de los métodos para el mes de marzo, marzo es un mes diferente dado que en las estadísticas descriptivas representa el mes más lluvioso, aprovechando eso podemos ver la interacción entre la variable elevación con las distancias al mar, dado que presenta mayores precipitaciones en la zona interandina, siendo el caso opuesto para el kriging universal en esa zona, de igual manera pronostica el kriging bayesiano.

Resultados del mes de Abril

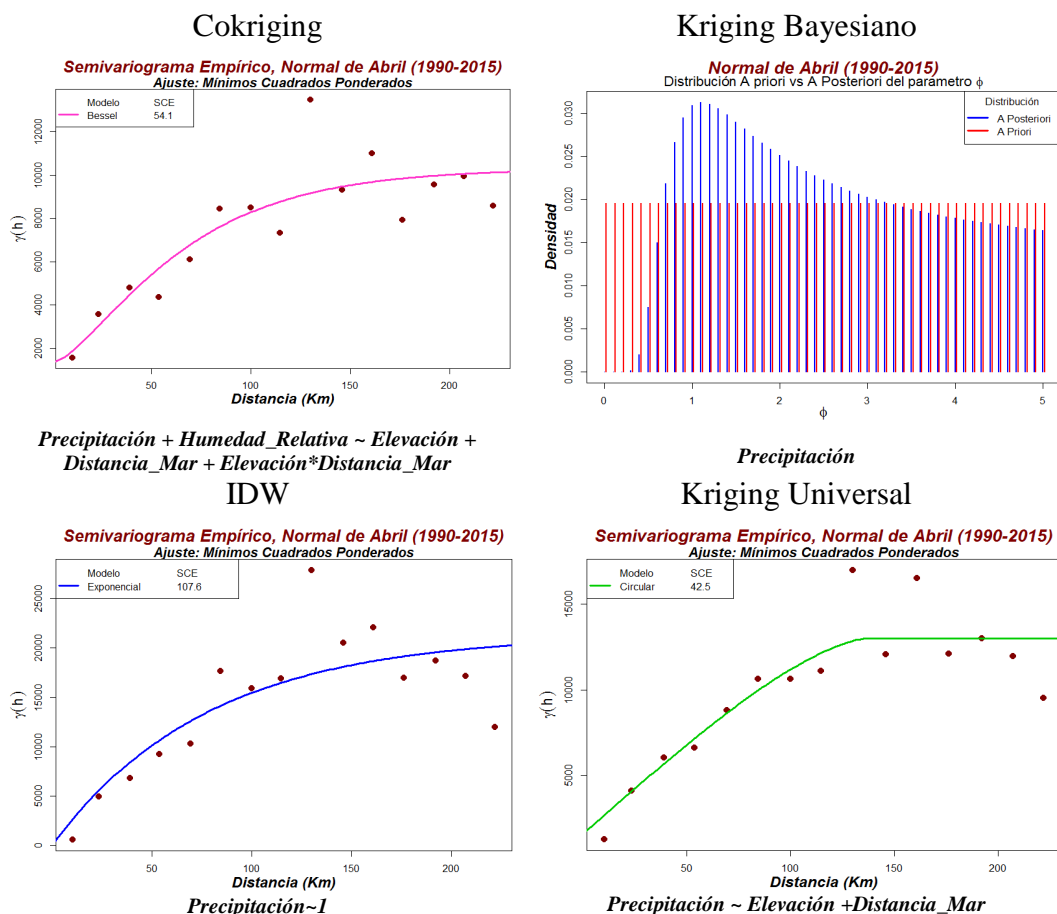


Gráfico 4.13. Medidas de Ajuste, Normal de Abril.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	21390.44	0	77.927	Exponencial	107.6
Kriging Universal	11431.785	1557.59	136.29	Circular	42.5
Cokriging	8935.41	1363.16	44.21	K-Bessel	54.1

Tabla 4.6. Parámetros de los Modelos para Abril.

Elaborado por: Gabriel Mantilla.

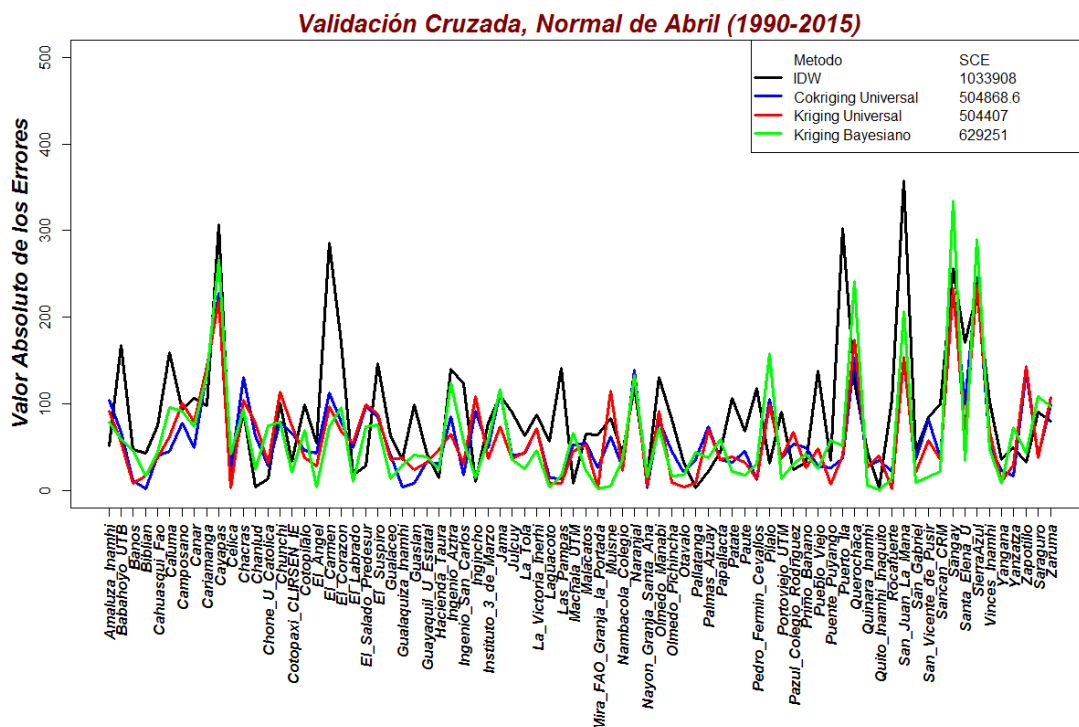


Gráfico 4.14. Validación cruzada Abril.  
Elaborado por: Gabriel Mantilla.

La validación cruzada en el gráfico 4.14 vemos que el pronóstico del cokriging se asemeja mucho al del kriging, teniendo SCE muy similares, a diferencia del IDW que tiene el doble de Suma cuadrática del Error.

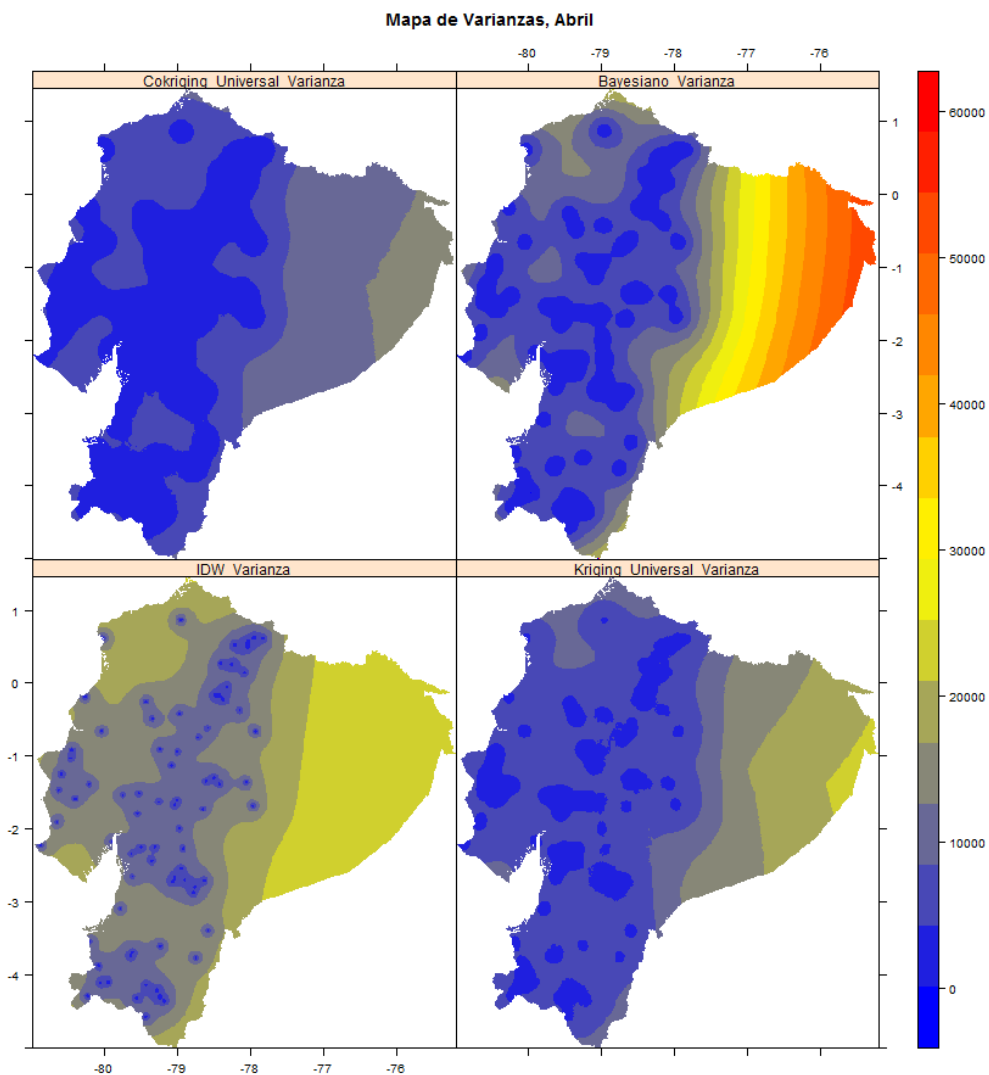


Ilustración 4.7. Varianzas de los interpoladores para Abril.  
Elaborado por: Gabriel Mantilla.

El cokriging para el mes de abril presenta de igual manera una varianza más pequeña en las zonas donde hay observaciones, que en donde no las hay. El kriging bayesiano dispara la varianza en la zona oriental donde no hay valores observados. Esto es importante al momento de ver un pronóstico donde la varianza es pequeña, podríamos sentirnos un poco más confiados de que es más ajustado a la realidad.

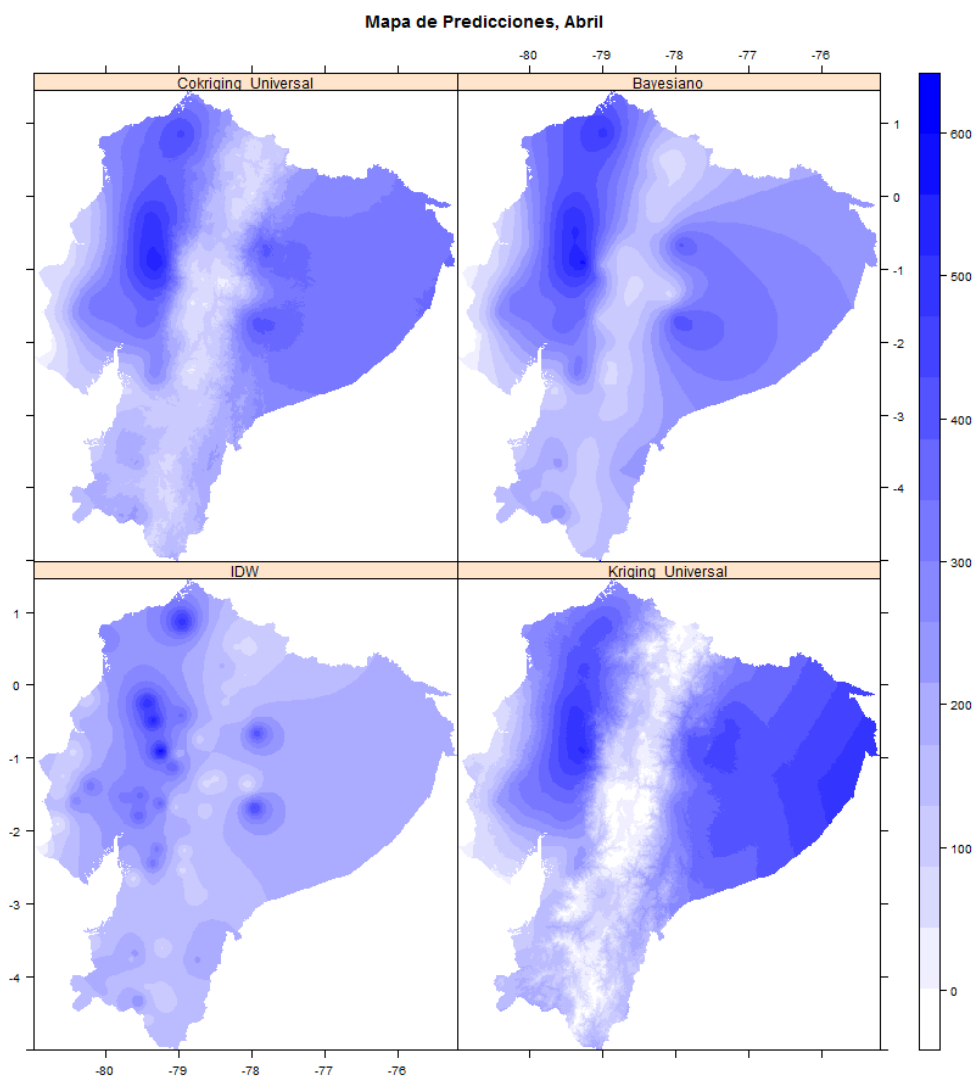
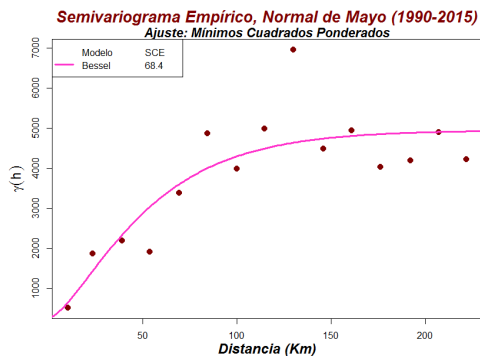


Ilustración 4.8. Interpolación de los métodos para Abril.  
Elaborado por: Gabriel Mantilla.

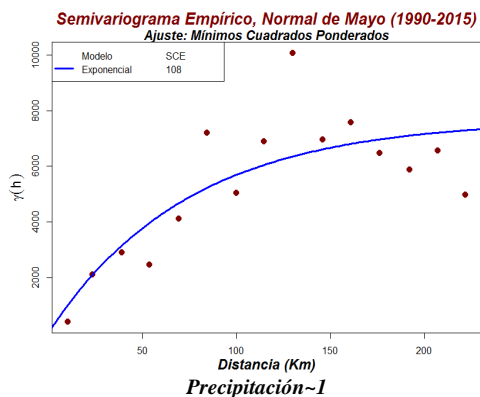
La ilustración 4.8 nos muestra que el único interpolador que está pronosticando lluvias de intensidad fuerte en la parte oriental es el método del kriging, en donde los otros interpoladores no lo hacen con aproximadamente 500 mm.

Resultados del mes de Mayo

Cokriging

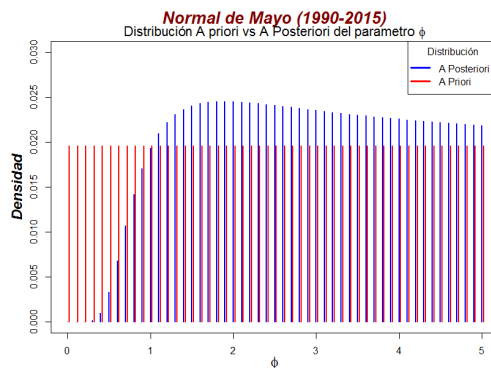


*Precipitación + Humedad\_Relativa ~ Elevación + Distancia\_Mar + Elevación\*Distancia\_Mar*  
IDW

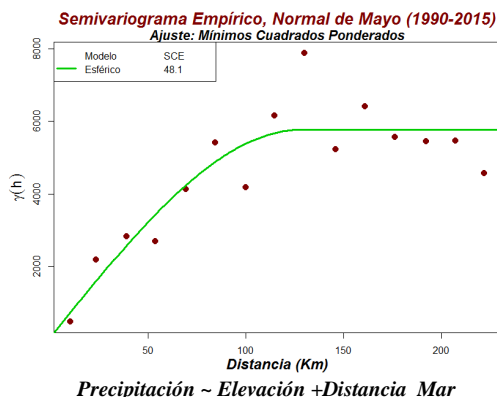


*Precipitación ~ I*

Kriging Bayesiano



*Precipitación*  
Kriging Universal



*Precipitación ~ Elevación + Distancia\_Mar*

Gráfico 4.15. Medidas de Ajuste, Normal de Mayo.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	7677.945	0	73.997	Exponencial	108
Kriging Universal	34.92	34.91	128.317	Esférico	48.1
Cokriging	4673.06	273.38	35.34	K-Bessel	68.4

Tabla 4.7. Parámetros de los Modelos para Mayo.

Elaborado por: Gabriel Mantilla.

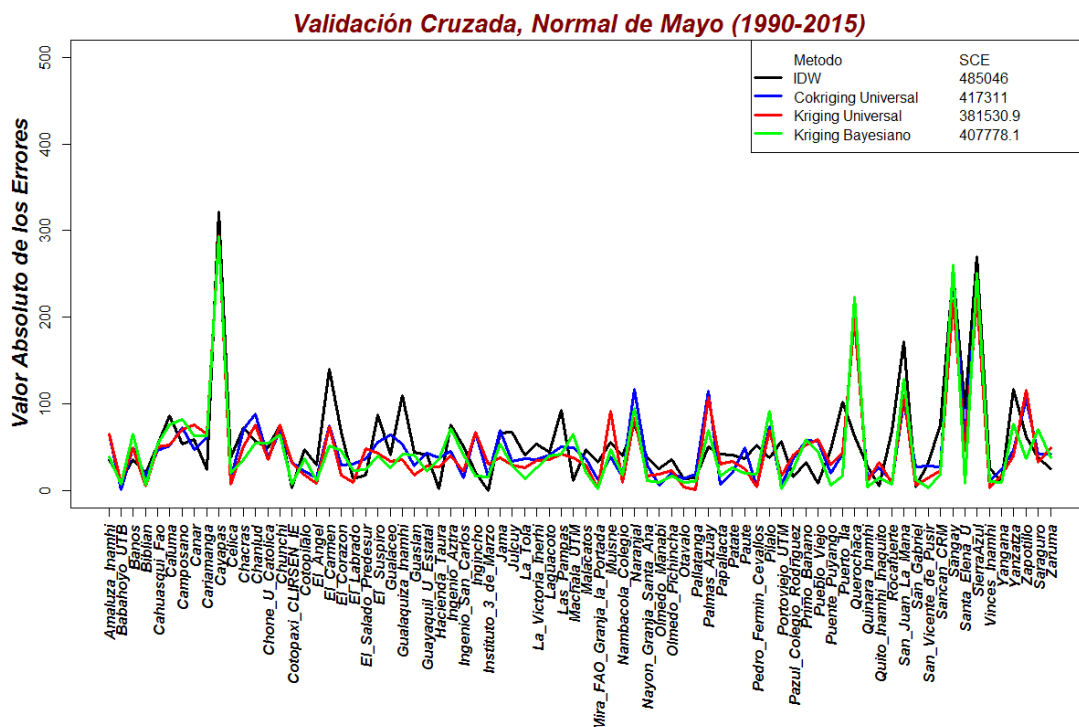


Gráfico 4.16. Validación cruzada Mayo.  
Elaborado por: Gabriel Mantilla.

De el gráfico 4.15 podemos notar que el modelo Esférico del kriging universal, mediante mínimos cuadrados, minimizo el SCE. La validación cruzada para el mes de mayo, nos muestra que es más riesgoso pronosticar en las estaciones de Cayapas, Sierra azul y Sangay por sus errores altos.

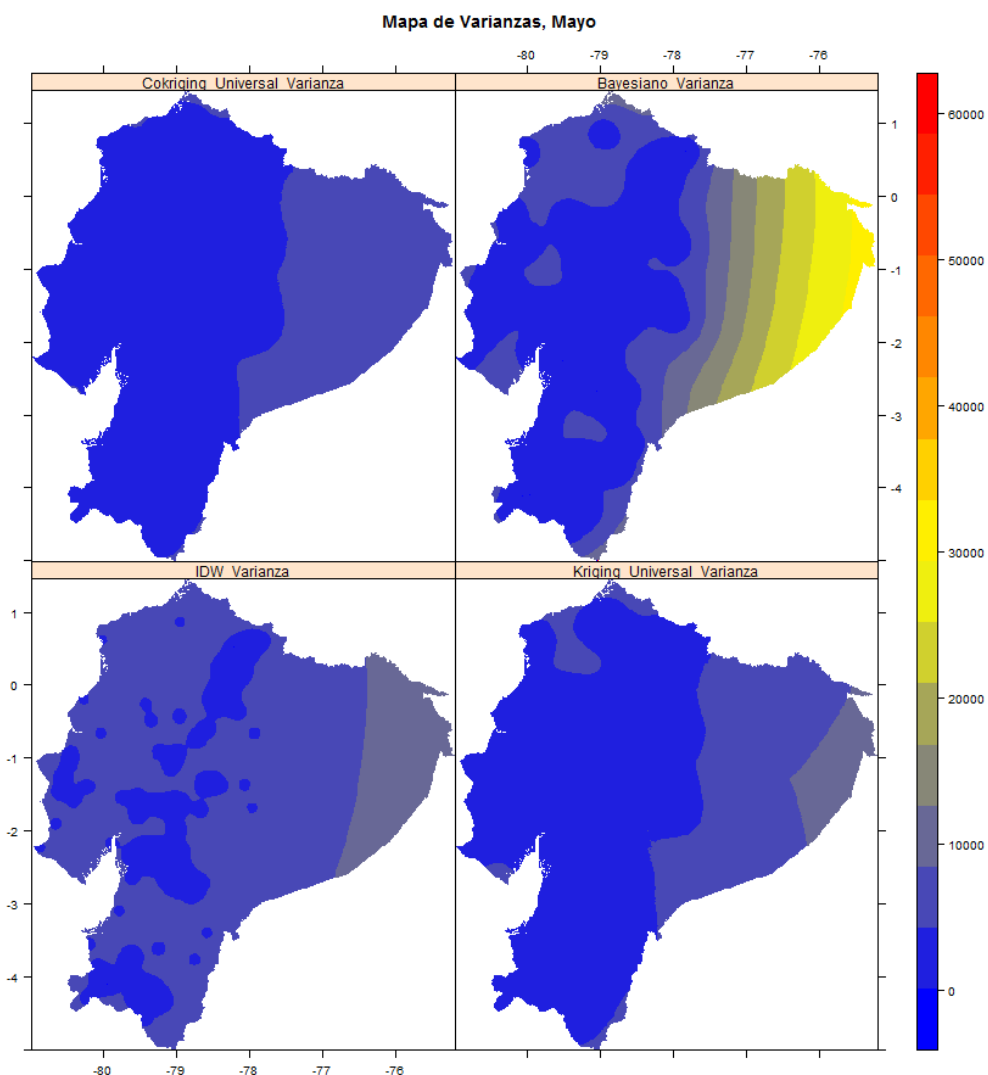


Ilustración 4.9. Varianzas de los interpoladores para Mayo.  
Elaborado por: Gabriel Mantilla.

De la ilustración 4.9, podemos concluir que las varianzas del cokriging y del kriging son casi similares, y las varianzas de los otros dos métodos son más bajas, esto es debido a medida que nos acercamos a la época seca en donde la variabilidad de la precipitación disminuye., las varianzas también lo hacen.



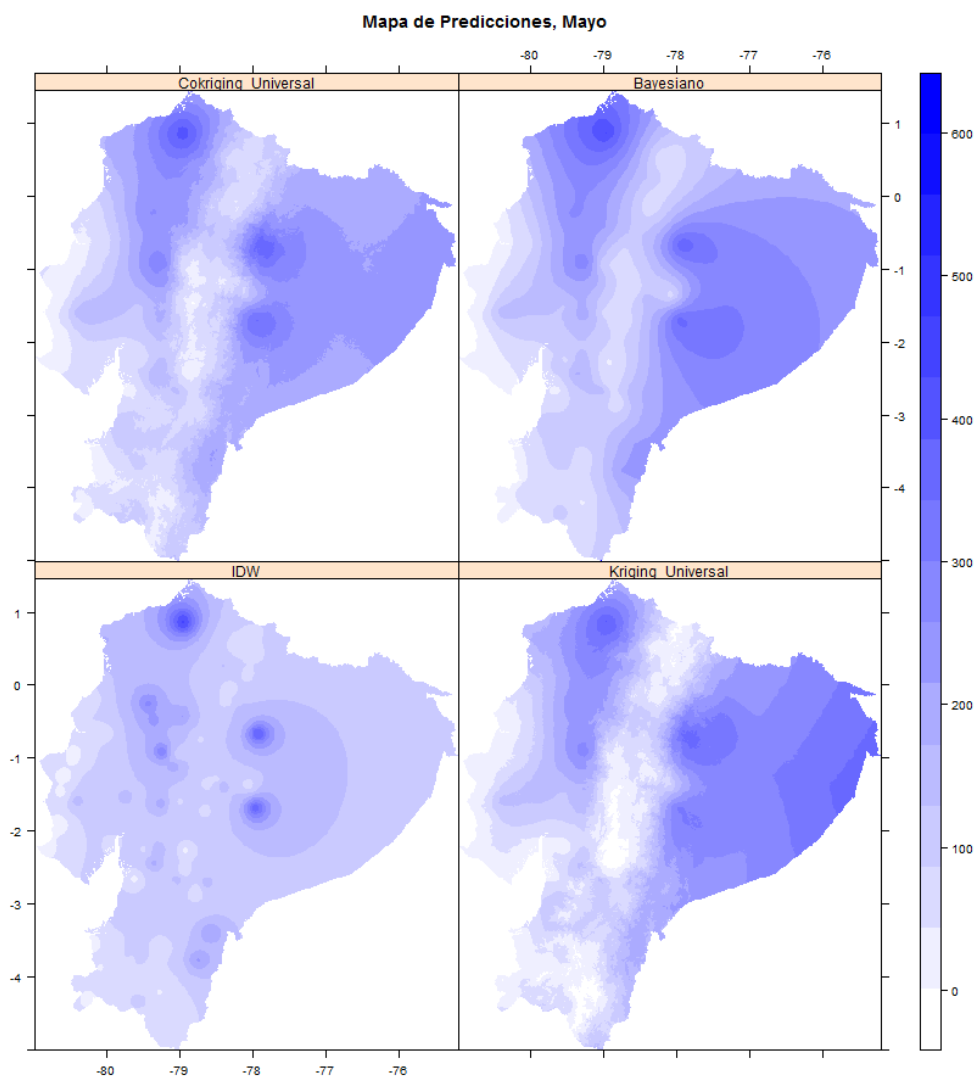
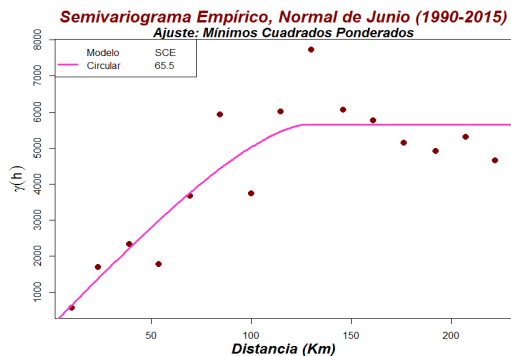


Ilustración 4.10. Interpolación de los métodos para Mayo.  
Elaborado por: Gabriel Mantilla.

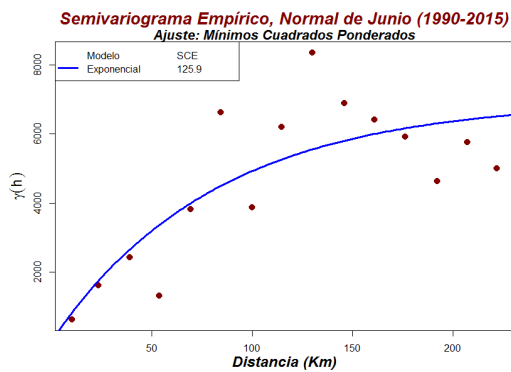
La ilustración 4.10 nos muestra precipitaciones más moderadas en la cuenca del Río Guayas entre los 50 y 100 mm, a diferencia de las estaciones de Cayapas, Sierra Azul y Sangay que registraron lluvias entre los 300 a 400 mm, siendo estos los puntos más azules.

Resultados del mes de Junio

Cokriging

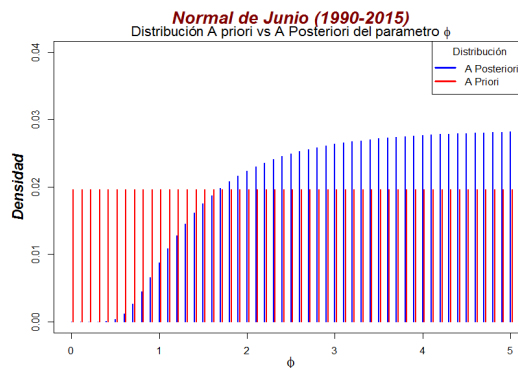


Precipitación + Humedad\_Relativa ~ Elevación + Distancia\_Mar + Elevación\*Distancia\_Mar  
IDW

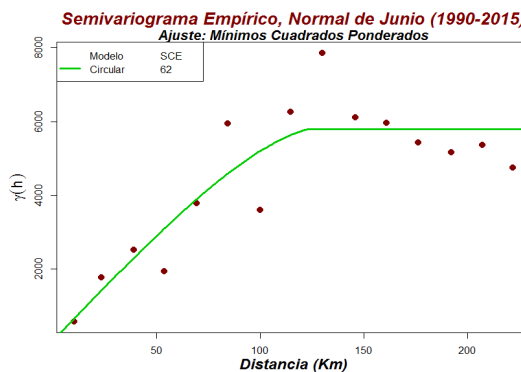


Precipitación~I

Kriging Bayesiano



Precipitación  
Kriging Universal



Precipitación ~ Elevación +Distancia\_Mar

Gráfico 4.17. Medidas de Ajuste, Normal de Junio.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	6964.06	0	81.42	Exponencial	125.9
Kriging Universal	5744.31	57.32	125.18	Circular	62
Cokriging	5586.38	63.45	126.58	Circular	65.5

Tabla 4.8. Parámetros de los Modelos para Junio.

Elaborado por: Gabriel Mantilla.

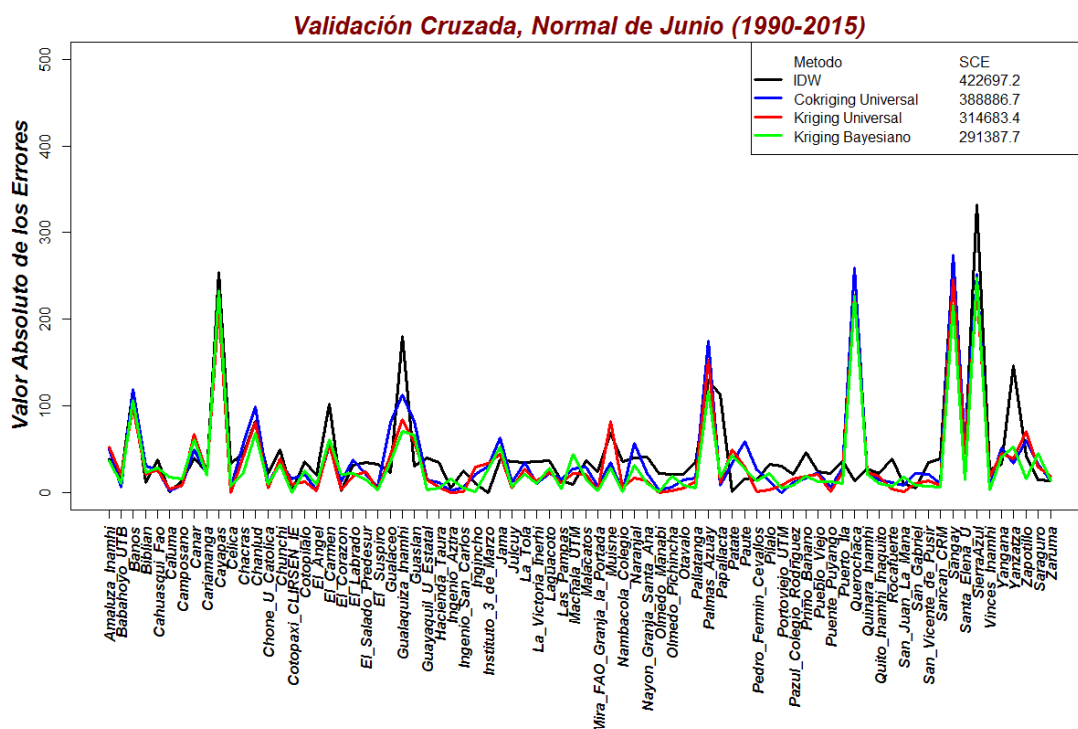


Gráfico 4.18. Validación cruzada Junio.  
Elaborado por: Gabriel Mantilla.

La validación cruzada para el mes de junio, nos muestra que el kriging bayesiano tiene un SCE menor con respecto a los otros interpoladores, algo que no se ha dado en los meses anteriores. Esto nos quiere decir que en meses donde la variabilidad de las lluvias no es tan alta, el método del kriging bayesiano funciona mucho mejor que métodos multivariados como los kriging simples.

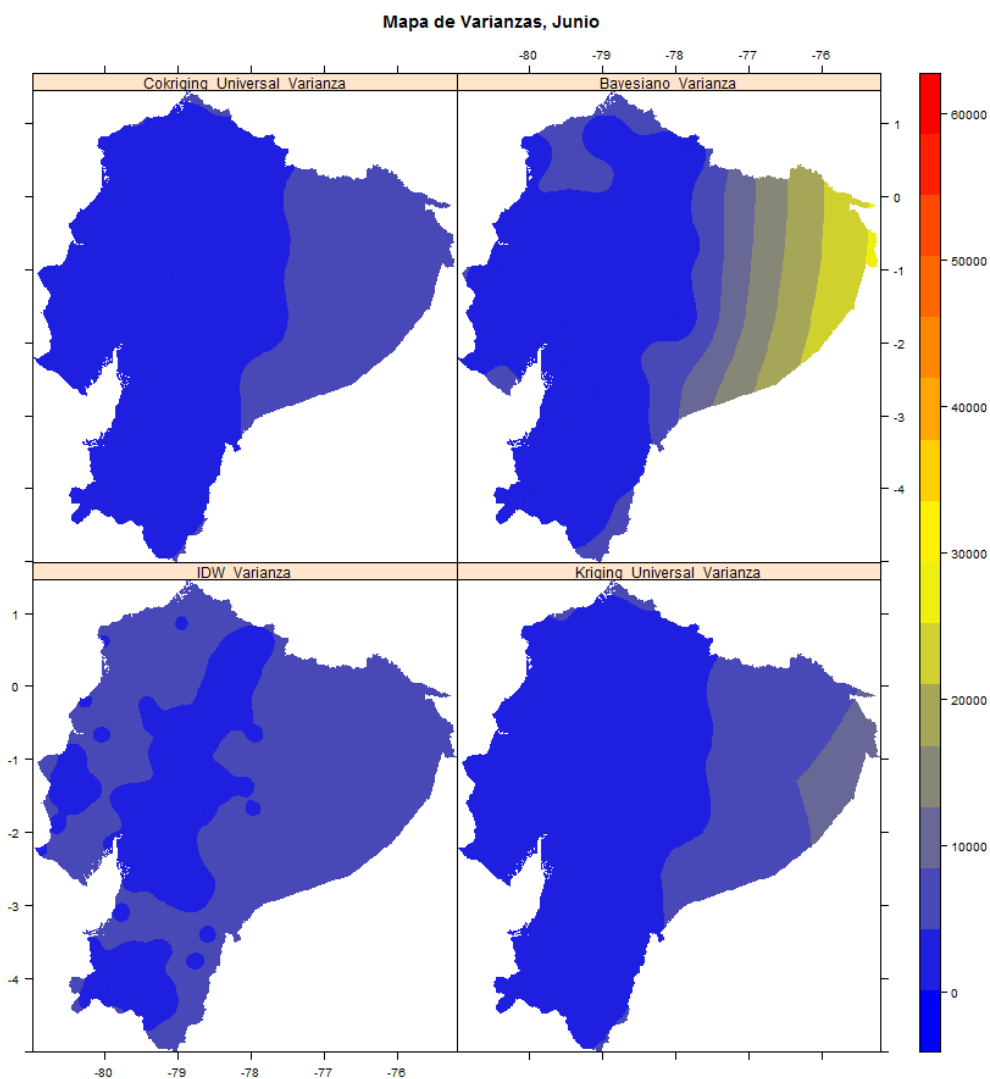


Ilustración 4.11. Varianzas de los interpoladores para Junio.  
Elaborado por: Gabriel Mantilla.

Los mapas de varianzas muestran que para el mes de junio, la varianza de los métodos es cercana a 0, el método del kriging bayesiano tiene mayor varianza en donde no ha sido posible obtener observaciones.

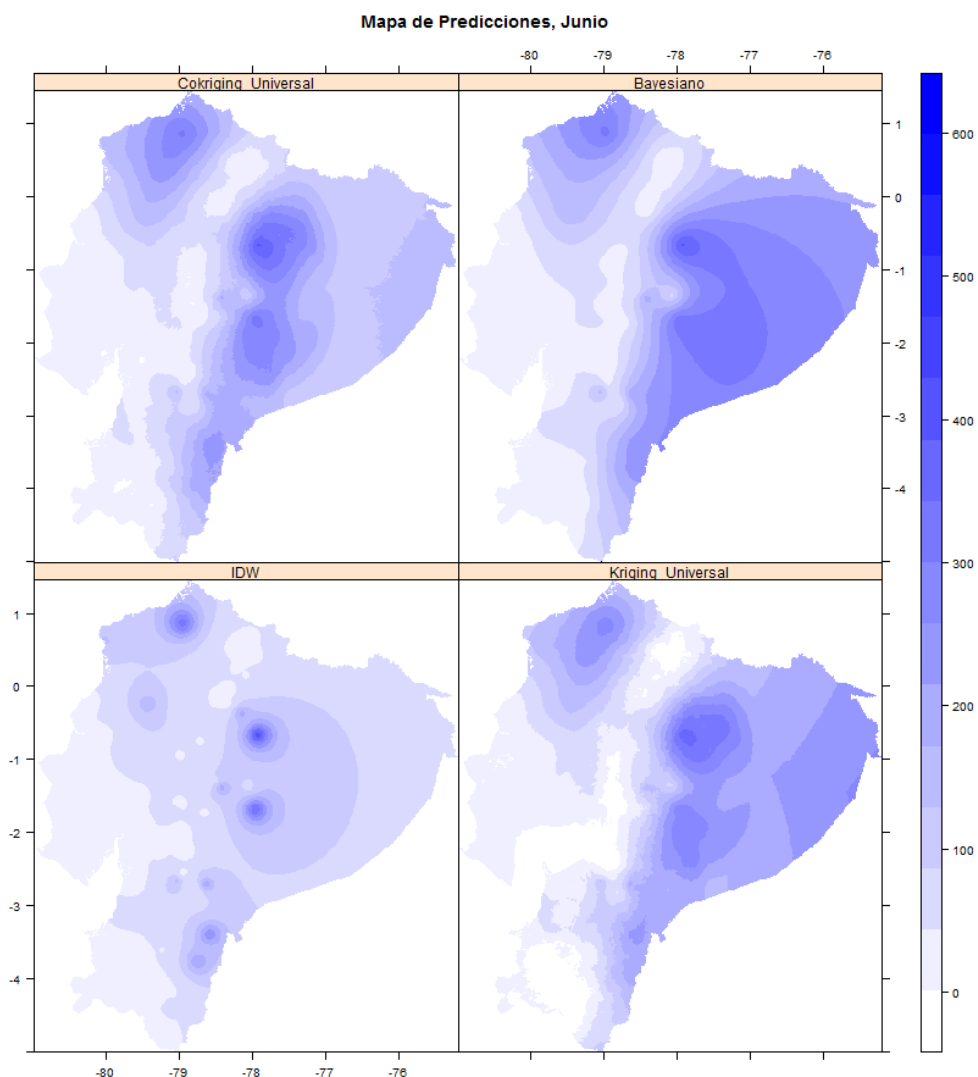
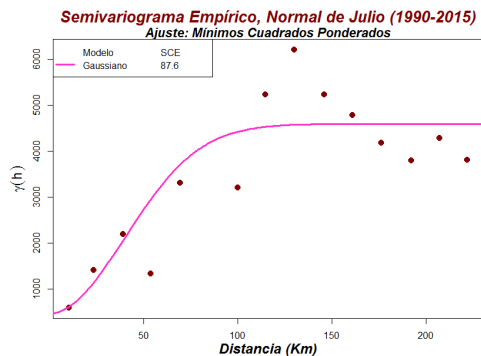


Ilustración 4.12. Interpolación de los métodos para Junio.  
Elaborado por: Gabriel Mantilla.

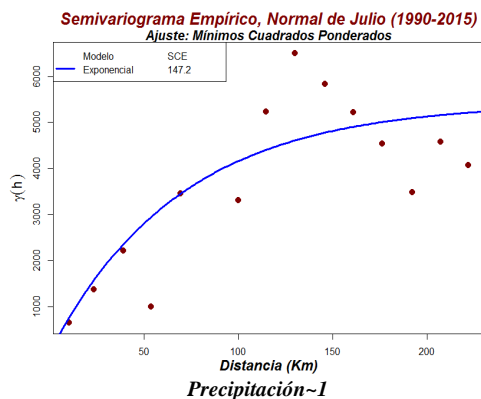
El método del IDW no sobrestima los valores observados, los puntos más azules muestran que las estaciones en esas zonas registraron precipitaciones alrededor de los 300 a 400 mm, solo en 3 estaciones, a pesar de esto los demás interpoladores sobre estiman los valores observados, suavizándolos a tal punto de que la zona coloridas tuvieran concordancia con los puntos cercanos a ella. .

Resultados del mes de Julio

Cokriging

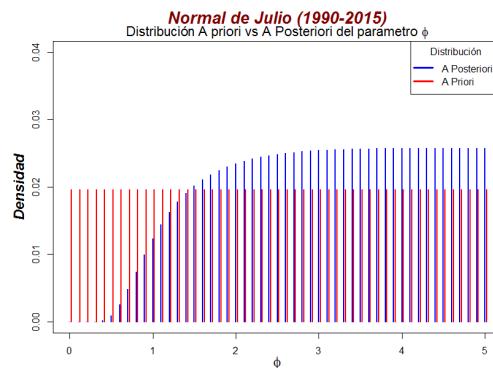


*Precipitación + Humedad\_Relativa ~ Elevación + Distancia\_Mar + Elevación\*Distancia\_Mar*  
IDW

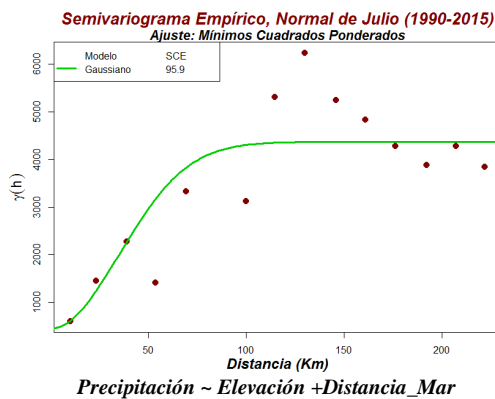


*Precipitación ~ I*

Kriging Bayesiano



*Precipitación*  
Kriging Universal



*Precipitación ~ Elevación + Distancia\_Mar*

Gráfico 4.19. Medidas de Ajuste, Normal de Julio.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	5427.52	0	68.71	Exponencial	147.2
Kriging Universal	3921.34	449.07	49.34	Gaussiano	95.9
Cokriging	4118.62	477.73	55.998	Gaussiano	87.6

Tabla 4.9. Parámetros de los Modelos para Julio.

Elaborado por: Gabriel Mantilla.

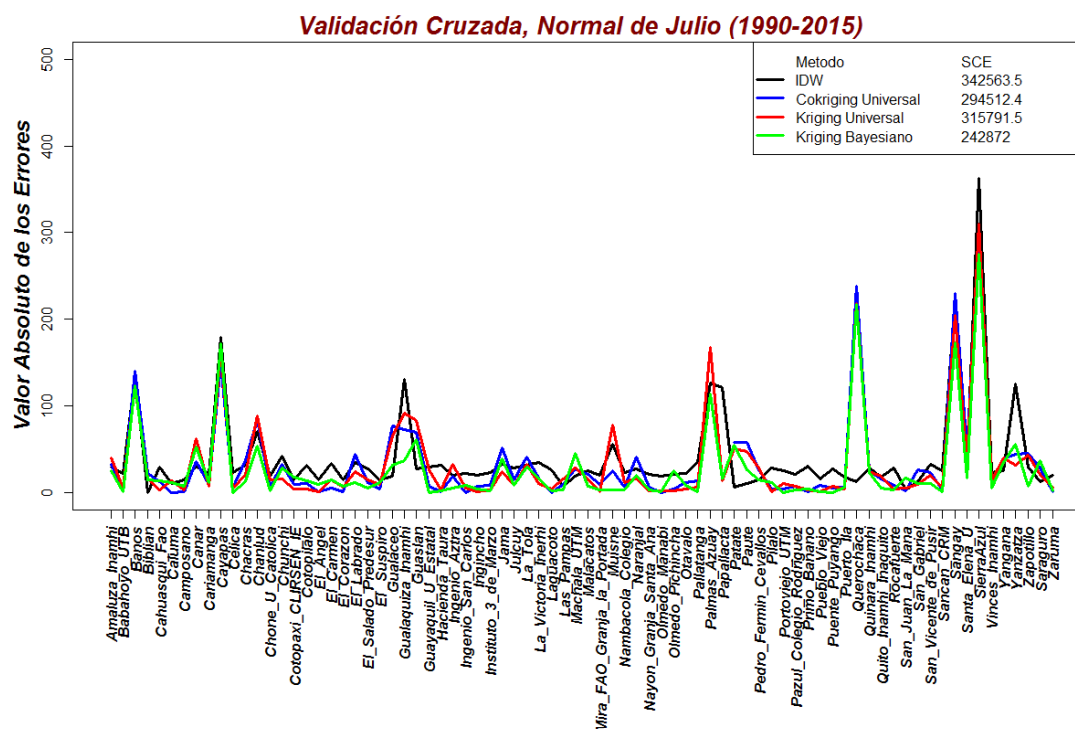


Gráfico 4.20. Validación cruzada Julio.  
Elaborado por: Gabriel Mantilla.

La validación cruzada del mes de Julio, muestra que el método del kriging bayesiano tiene menor SCE con respecto a los otros métodos con 242872. Si observamos los errores del gráfico 4.20, todos los métodos tienen inconvenientes para realizar un buen pronóstico en la estación de Sangay y Sierra Azul, siendo los picos más altos del gráfico.

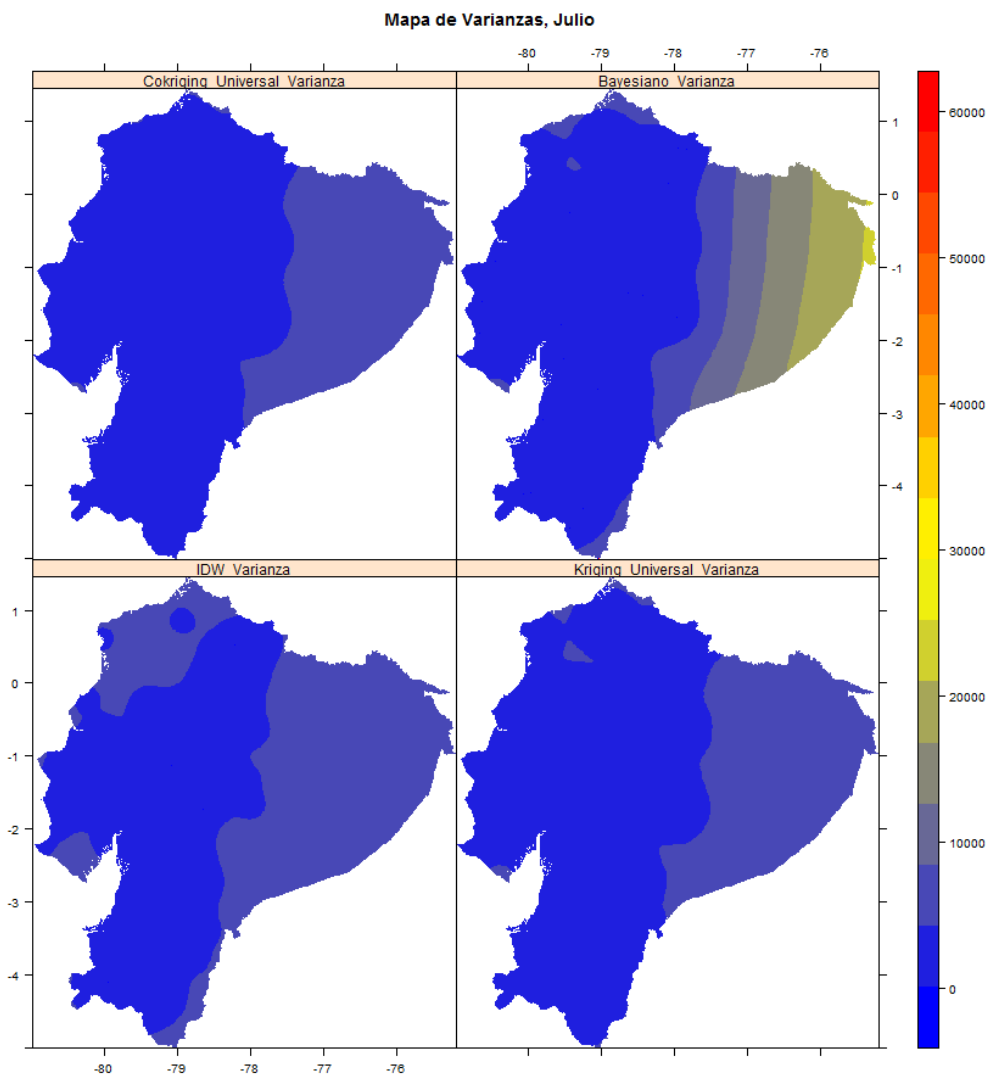


Ilustración 4.13. Varianzas de los interpoladores para Julio.  
Elaborado por: Gabriel Mantilla.



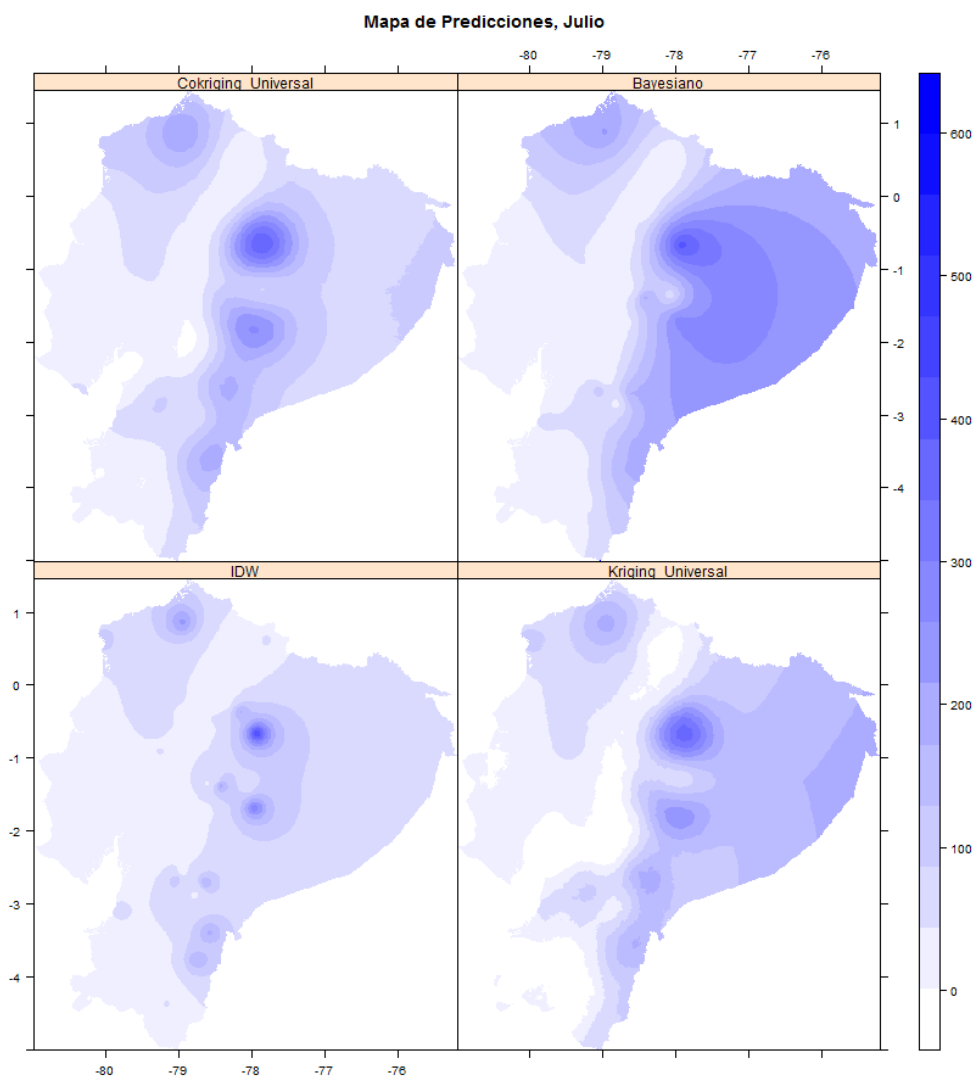


Ilustración 4.14. Interpolación de los métodos para Julio.  
Elaborado por: Gabriel Mantilla.

La validación cruzada del gráfico 4.20, muestra los errores más altos para las estaciones de Sangay y Sierra azul presentan mayor error de pronóstico, siendo los dos círculos azules que vemos en los mapas con precipitaciones de 300 a 350 mm, presentando mayor variabilidad con respecto a los otros valores observados.

Resultados del mes de Agosto.

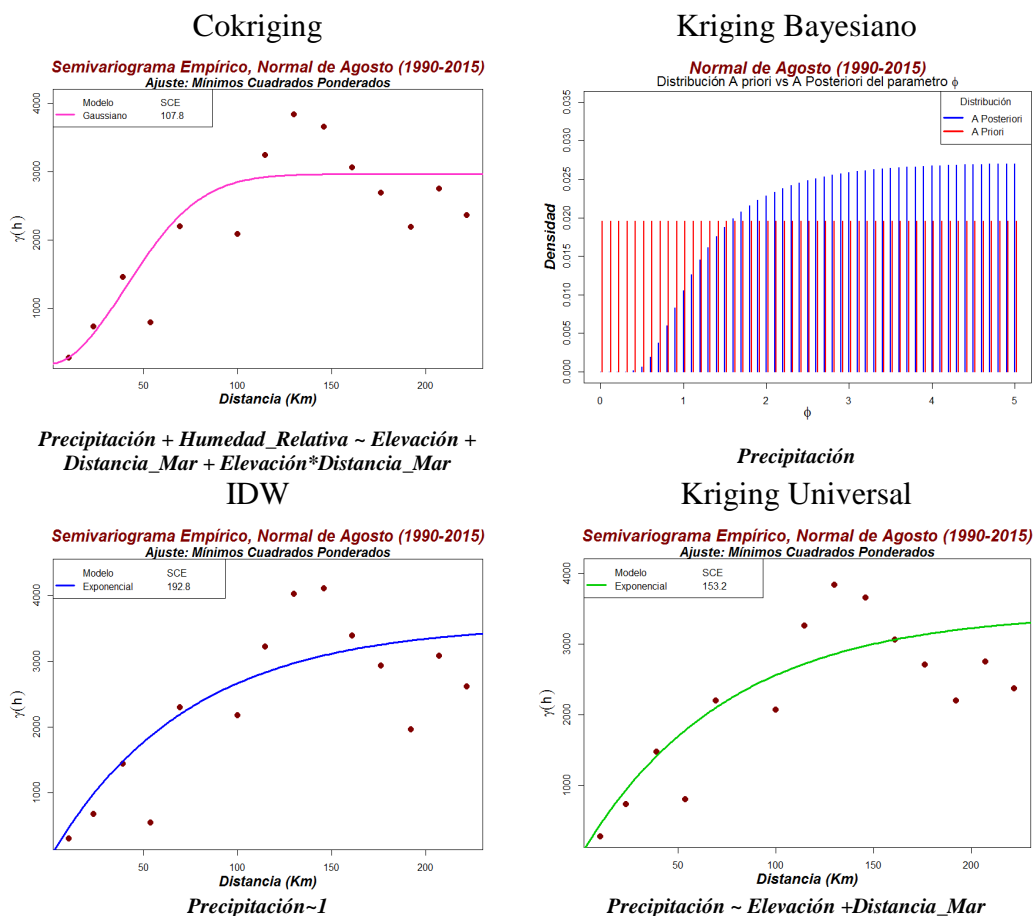


Gráfico 4.21. Medidas de Ajuste, Normal de Agosto.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	3570.06	0	72.85	Exponencial	192.8
Kriging Universal	3455.833	0	74.14	Exponencial	153.2
Cokriging	2774.984	191.32	56.22	Gaussiano	107.8

Tabla 4.10. Parámetros de los Modelos para Agosto.

Elaborado por: Gabriel Mantilla.

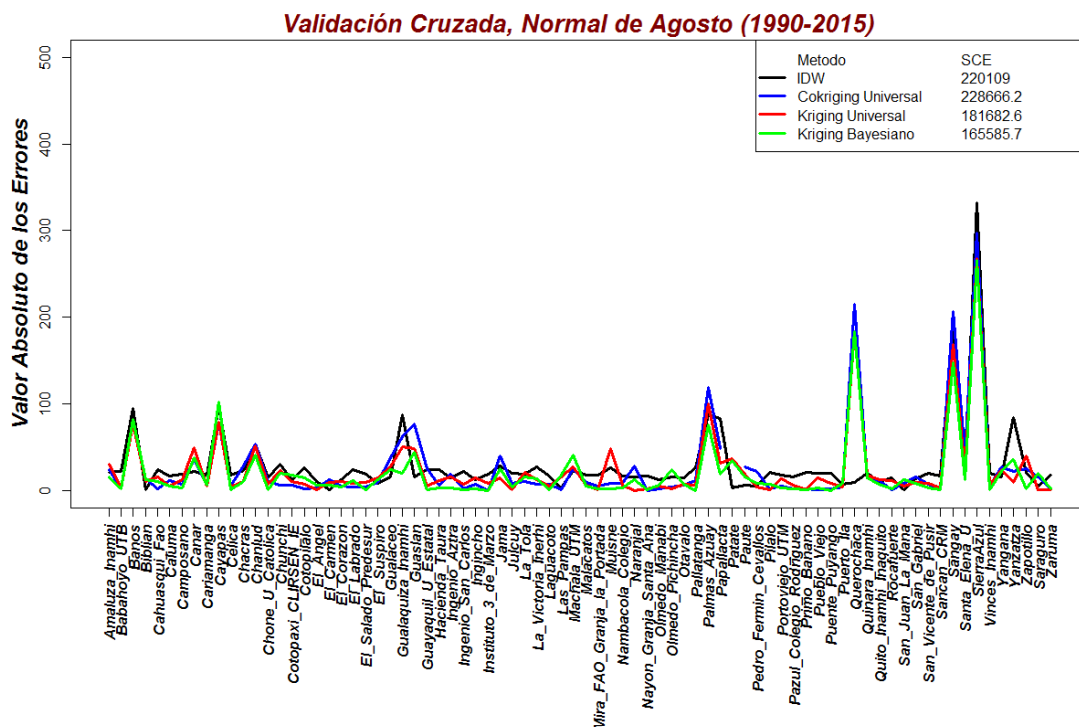


Gráfico 4.22. Validación cruzada Agosto.  
Elaborado por: Gabriel Mantilla.

De la tabla 4.10 podemos observar que los SCE de los métodos para los semivariogramas empíricos son altos, en la validación cruzada presenta errores altos para las estaciones de Querochaca y Sangay, siendo estos los picos más altos del gráfico. La Validación Cruzada del gráfico 4.22 para el mes de Agosto muestra que el kriging bayesiano tiene menor SCE con respecto a los otros métodos.

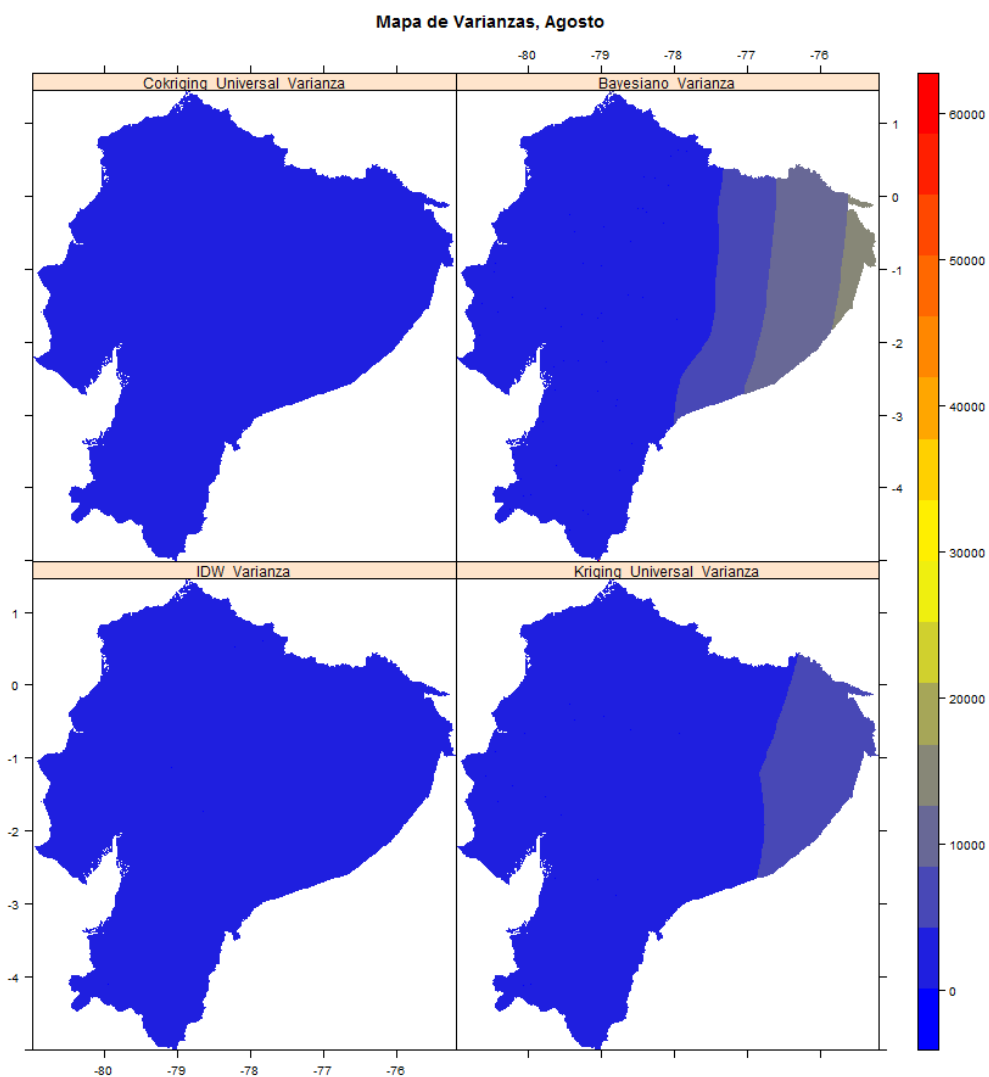


Ilustración 4.15. Varianzas de los interpoladores para Agosto.  
Elaborado por: Gabriel Mantilla.

La ilustración 4.15 nos muestra que las varianzas del mes de agosto para los 4 métodos son bajas, alrededor de 0, categorizando a las zonas del mapa de color azul. La varianza del kriging bayesiano es baja en el oriente, en donde no hay valores observados.

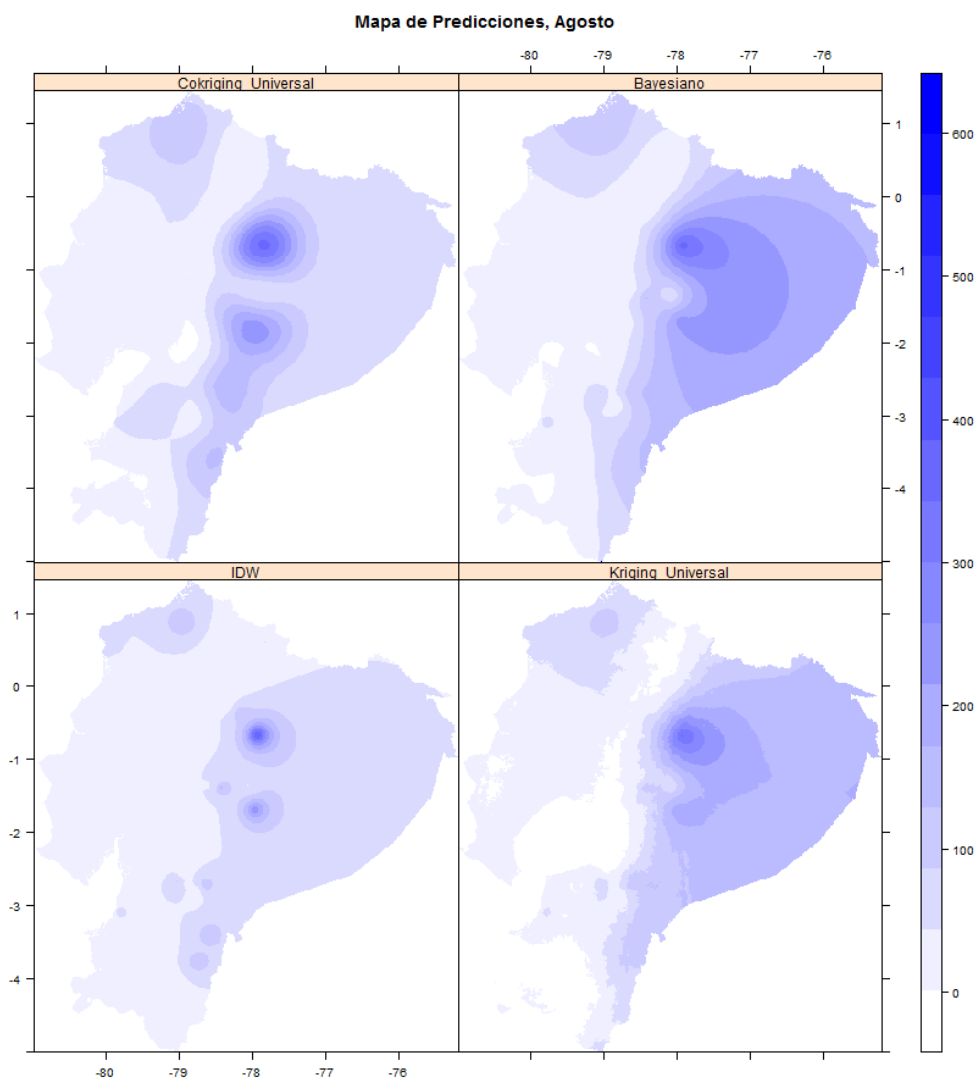
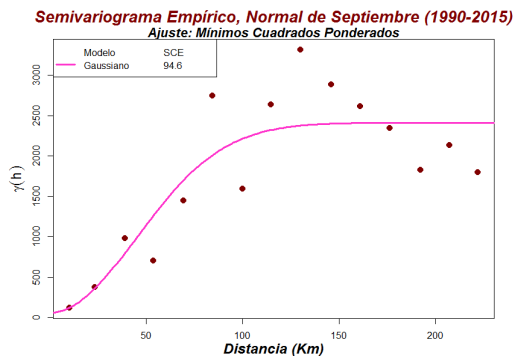


Ilustración 4.16. Interpolación de los métodos para Agosto.  
Elaborado por: Gabriel Mantilla.

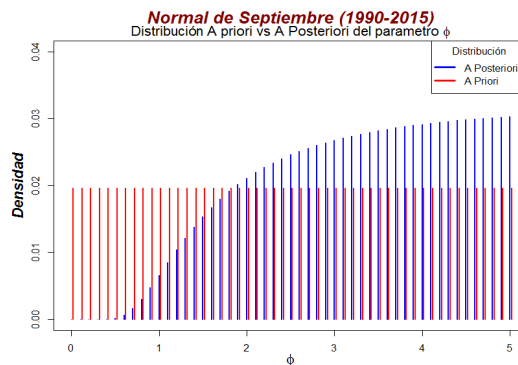
La ilustración 4.16 nos muestra la interpolación de los 4 métodos, en la zona más cercana al borde costero, se ven precipitaciones alrededor de los 75 a 150 mm, y un pico de 368 mm en la estación de Sangay.

Resultados del mes de Septiembre

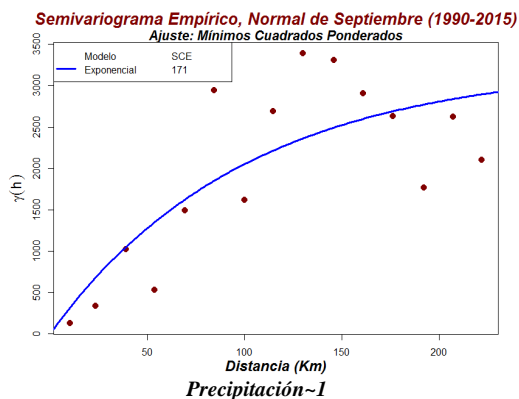
Cokriging



Kriging Bayesiano



Precipitación + Humedad\_Relativa ~ Elevación + Distancia\_Mar + Elevación\*Distancia\_Mar  
IDW



Precipitación  
Kriging Universal

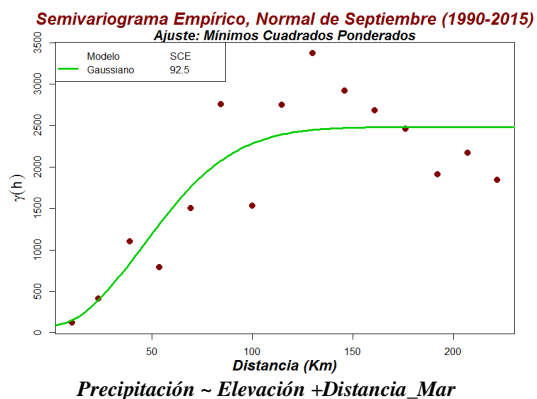


Gráfico 4.23. Medidas de Ajuste, Normal de Septiembre.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	3249.121	0	100.14	Exponencial	171
Kriging Universal	2396.28	86.97	63.49	Gaussiano	92.5
Cokriging	2352.065	59.589	63.544	Gaussiano	94.6

Tabla 4.11. Parámetros de los Modelos para Septiembre.

Elaborado por: Gabriel Mantilla.

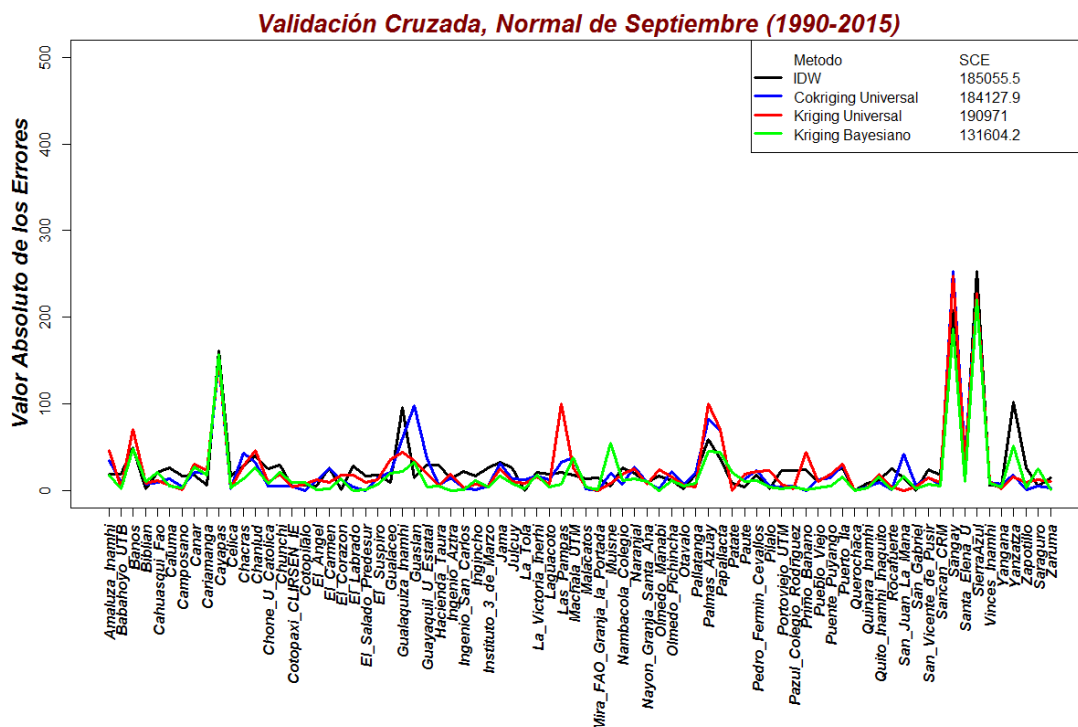


Gráfico 4.24. Validación cruzada Septiembre.  
Elaborado por: Gabriel Mantilla.

La validación cruzada del mes de Septiembre muestra que el kriging bayesiano tiene menor SCE con respecto a los otros métodos con 131604. Los errores más significativos de la VC del mes de septiembre son de la estación de Sangay y Sierra Azul.

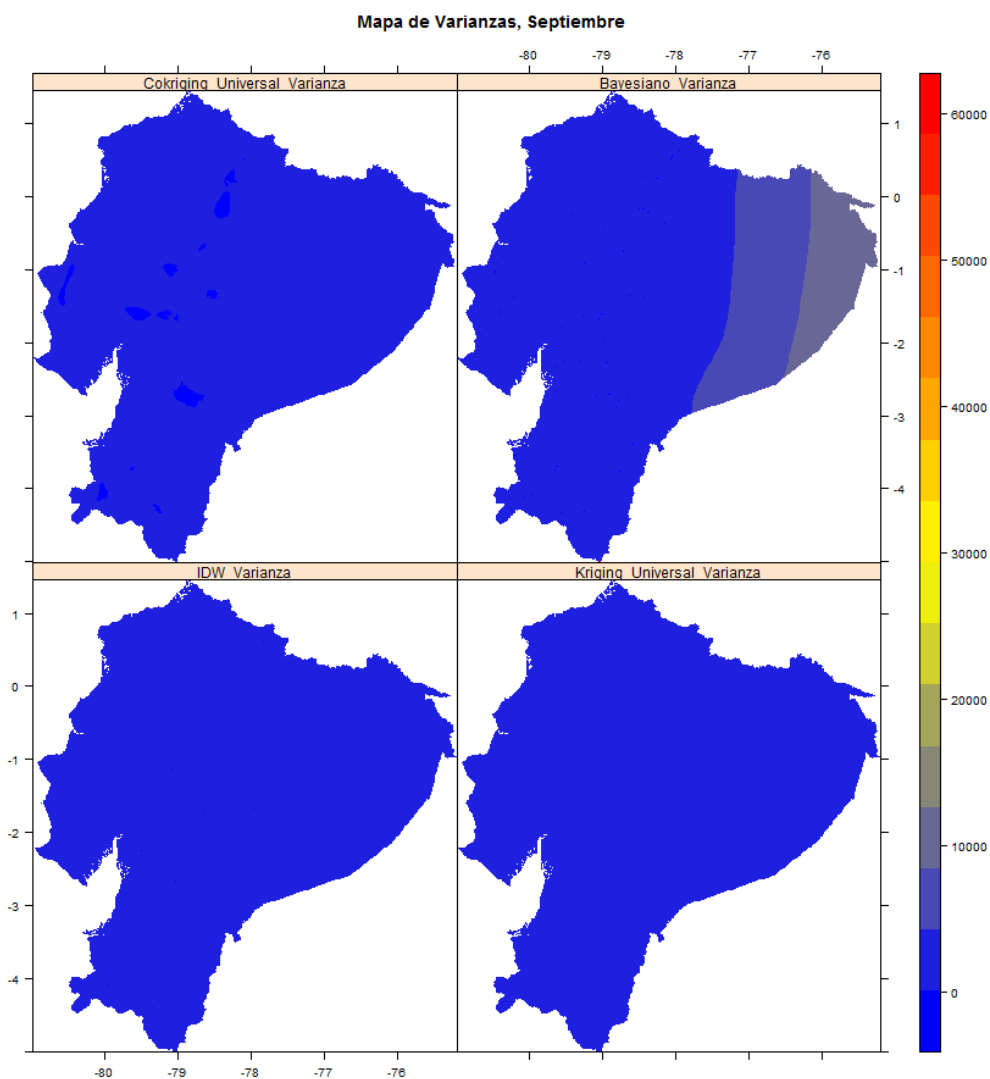


Ilustración 4.17. Varianzas de los interpoladores para Septiembre.  
Elaborado por: Gabriel Mantilla.

La ilustración 4.17 nos muestra las varianzas de los 4 métodos para el mes de Septiembre, siendo todas bajas alrededor de 0, de igual manera el kriging bayesiano muestra varianzas más altas en donde no se encuentran valores observados.



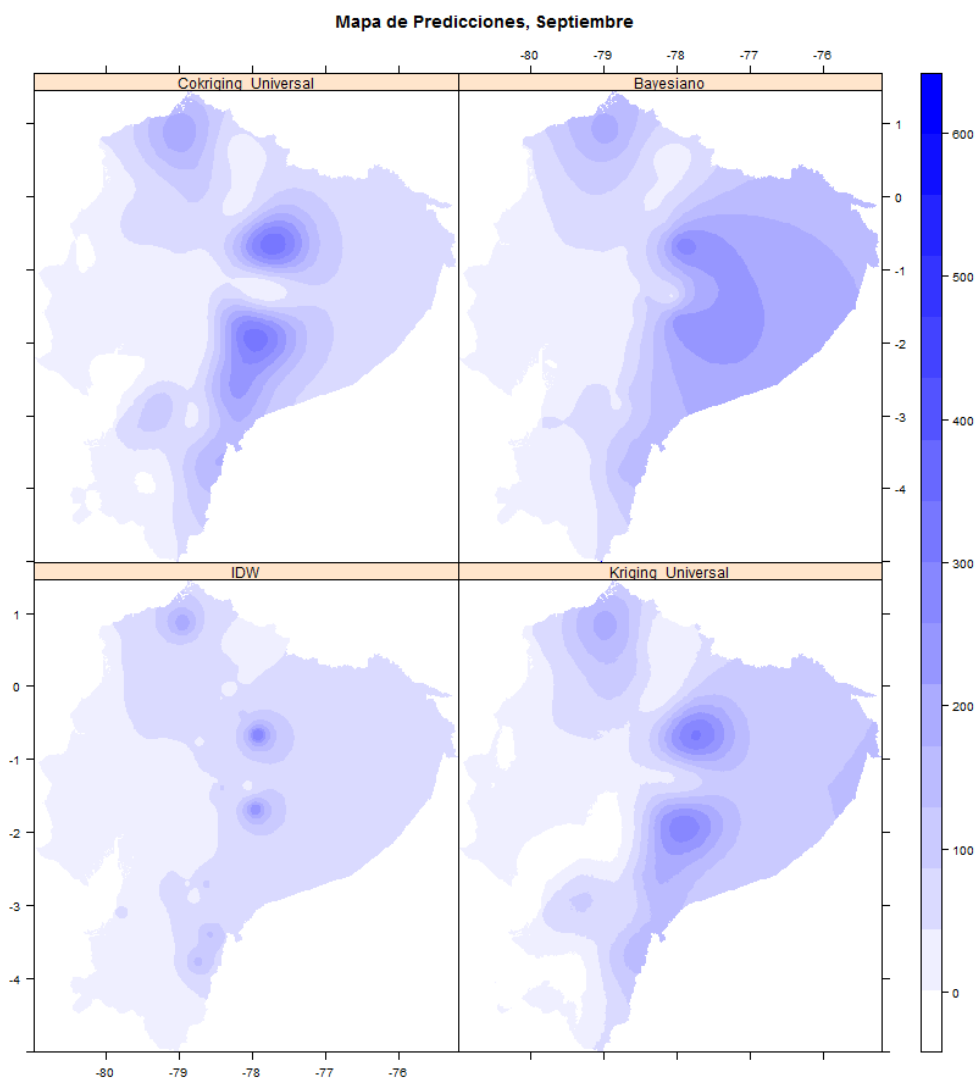


Ilustración 4.18. Interpolación de los métodos para Septiembre.  
Elaborado por: Gabriel Mantilla.

Resultados del mes de Octubre.

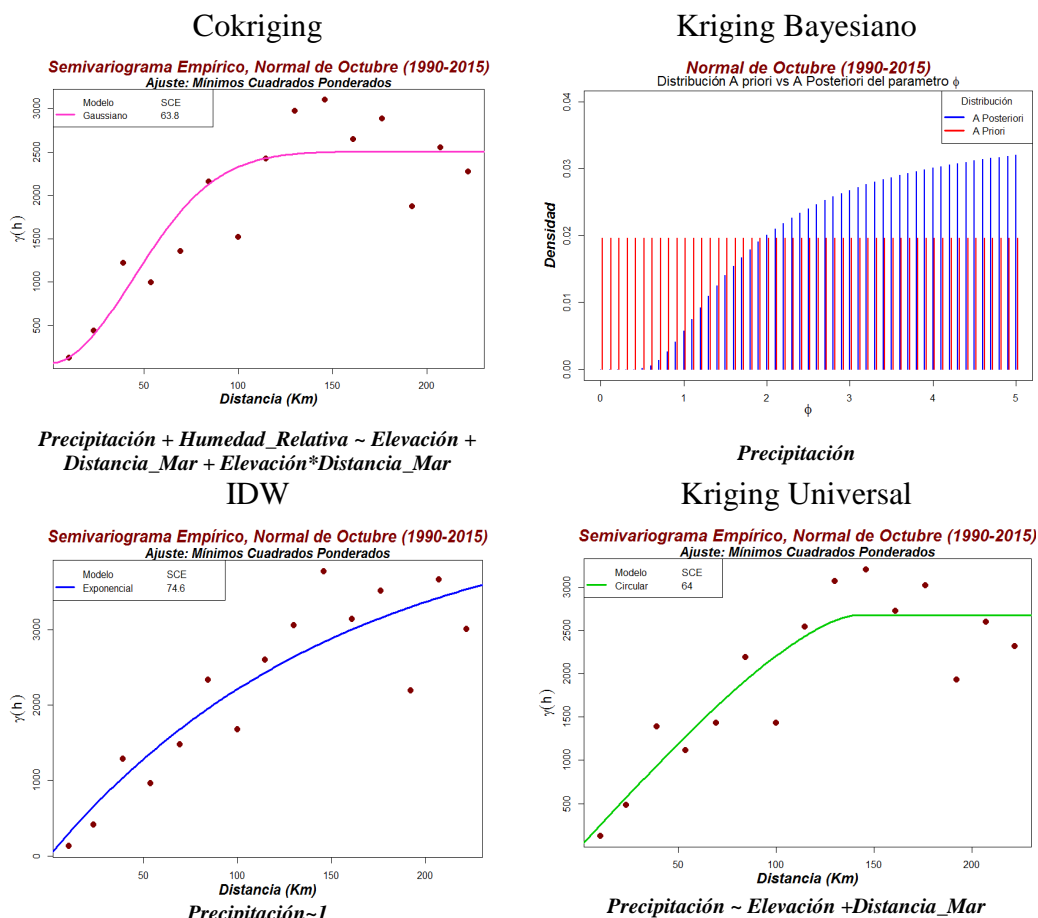


Grafico 4.25. Medidas de Ajuste, Normal de Octubre.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	4630.174	0	154.15	Exponencial	74.6
Kriging Universal	2673.769	0	140.069	Circular	64
Cokriging	2440.745	65.836	61.92	Gaussiano	63.8

Tabla 4.12. Parámetros de los Modelos para Octubre.

Elaborado por: Gabriel Mantilla.

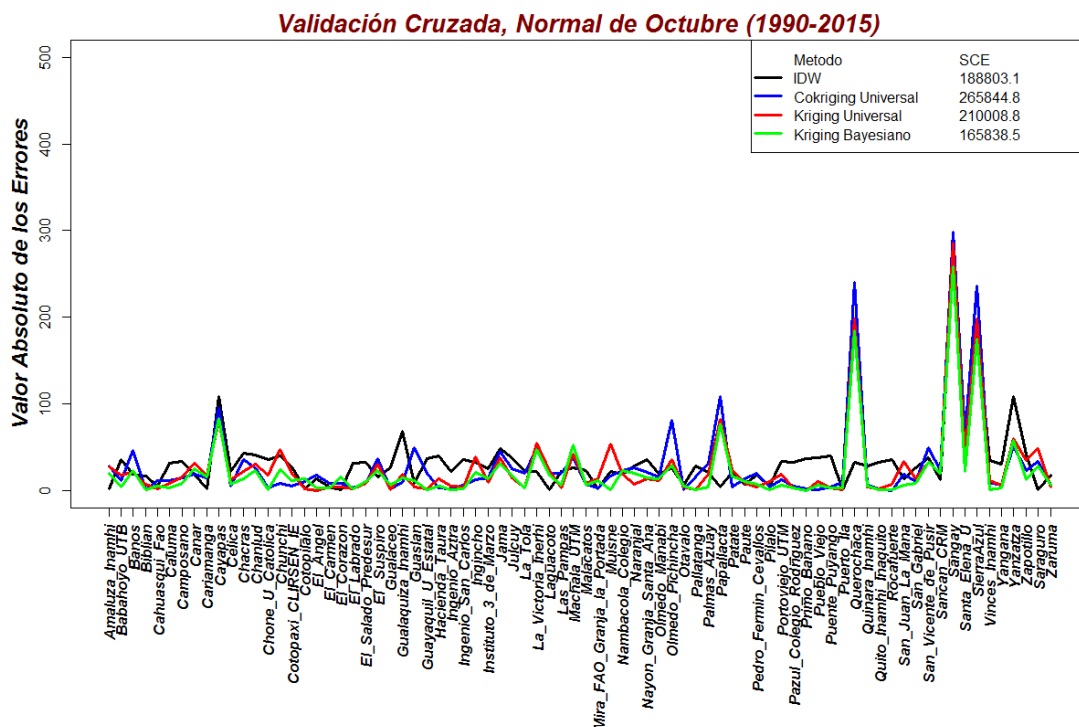


Gráfico 4.26. Validación cruzada Octubre.  
Elaborado por: Gabriel Mantilla.

La VC del mes de Octubre muestra que la estación de Sangay tiene el pico más alto en la gráfica de errores. El método del kriging bayesiano presenta menor SCE con respecto a los otros métodos con 165838.

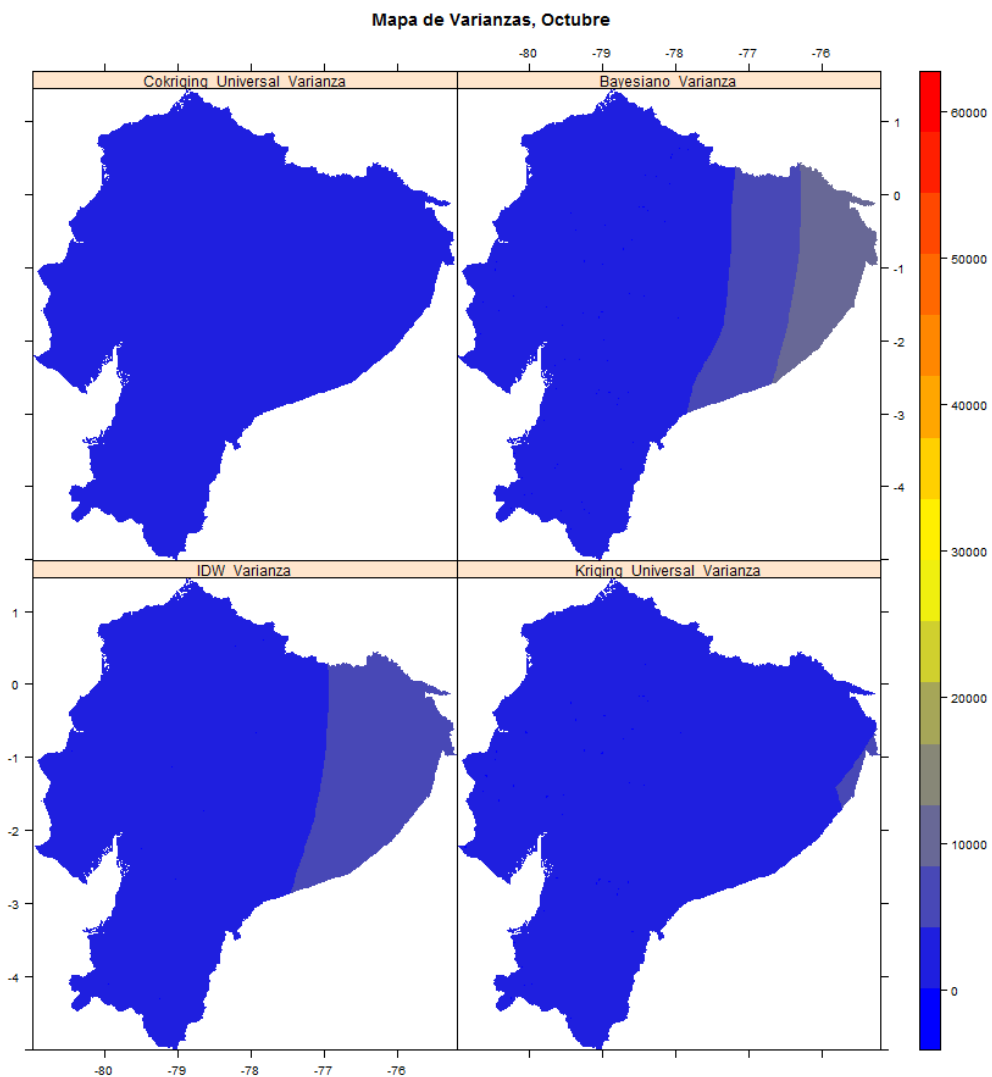


Ilustración 4.19. Varianzas de los interpoladores para Octubre.  
Elaborado por: Gabriel Mantilla.

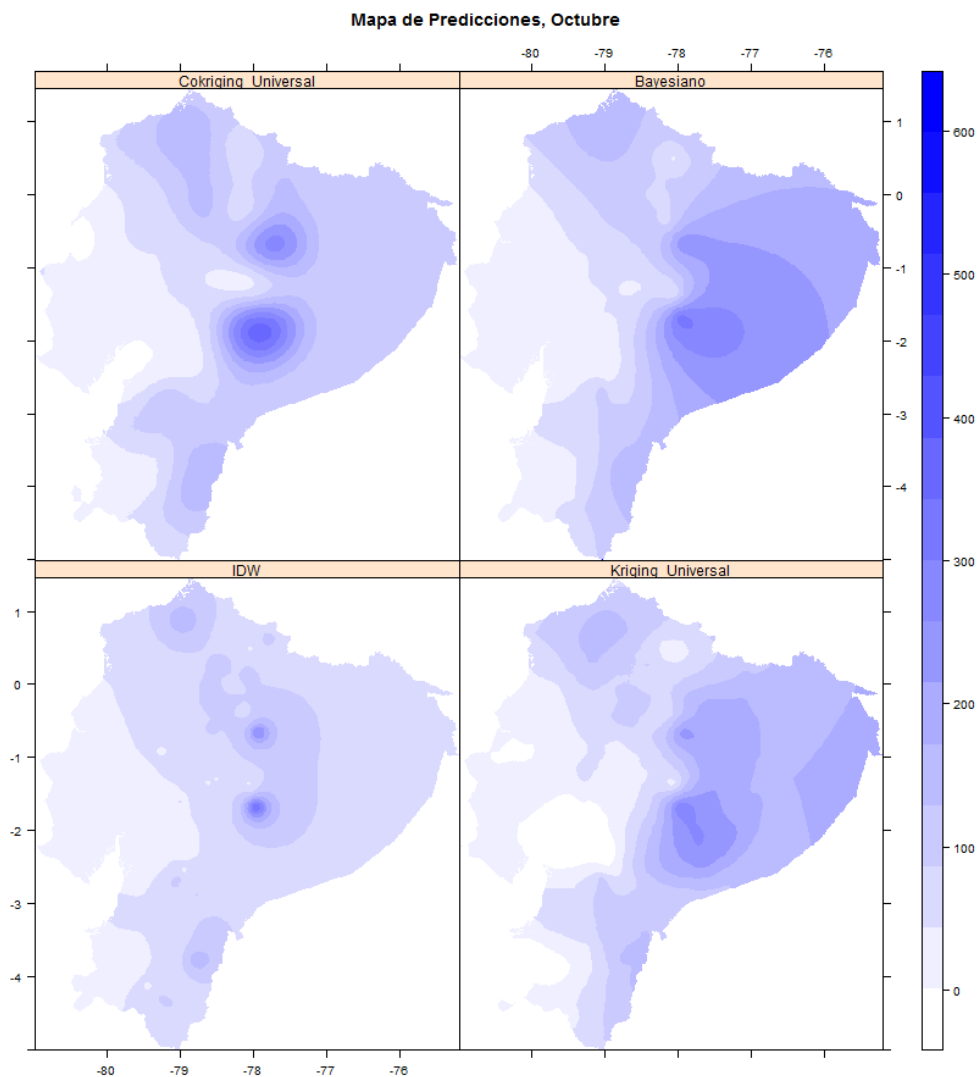


Ilustración 4.20. Interpolación de los métodos para Octubre.  
Elaborado por: Gabriel Mantilla.

El color más azul pintado en los mapas, corresponde a la estación de Sangay, con 324 mm de precipitación, siendo el valor más alto con respecto a las demás estaciones. La estación con el valor más bajo es la de Pazul Colegio Rodriguez con 7.4 mm pintado alrededor del color blanco.

Resultados del mes de Noviembre.

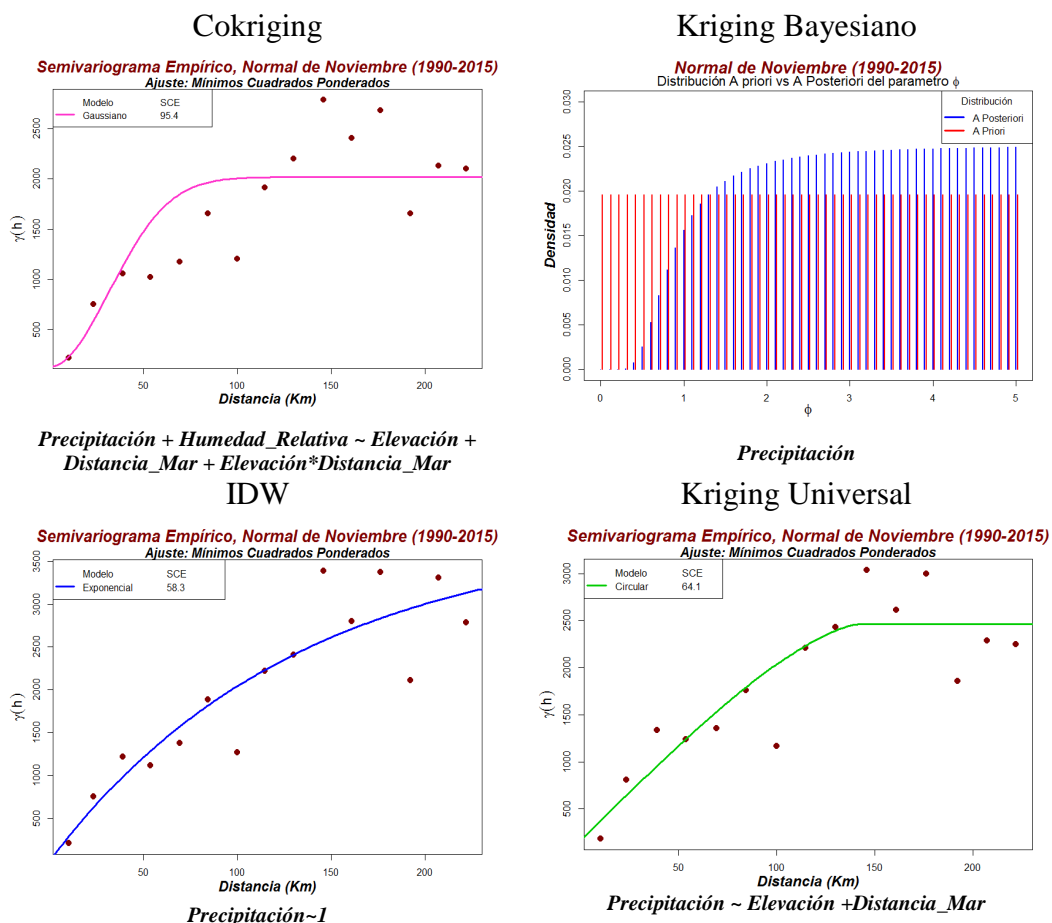


Grafico 4.27. Medidas de Ajuste, Normal de Noviembre.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	3850.58	0	132.264	Exponencial	58.3
Kriging Universal	2304.09	158.57	142.42	Circular	64.1
Cokriging	1887.11	132.18	44.8	Gaussiano	95.4

Tabla 4.13. Parámetros de los Modelos para Noviembre.

Elaborado por: Gabriel Mantilla.

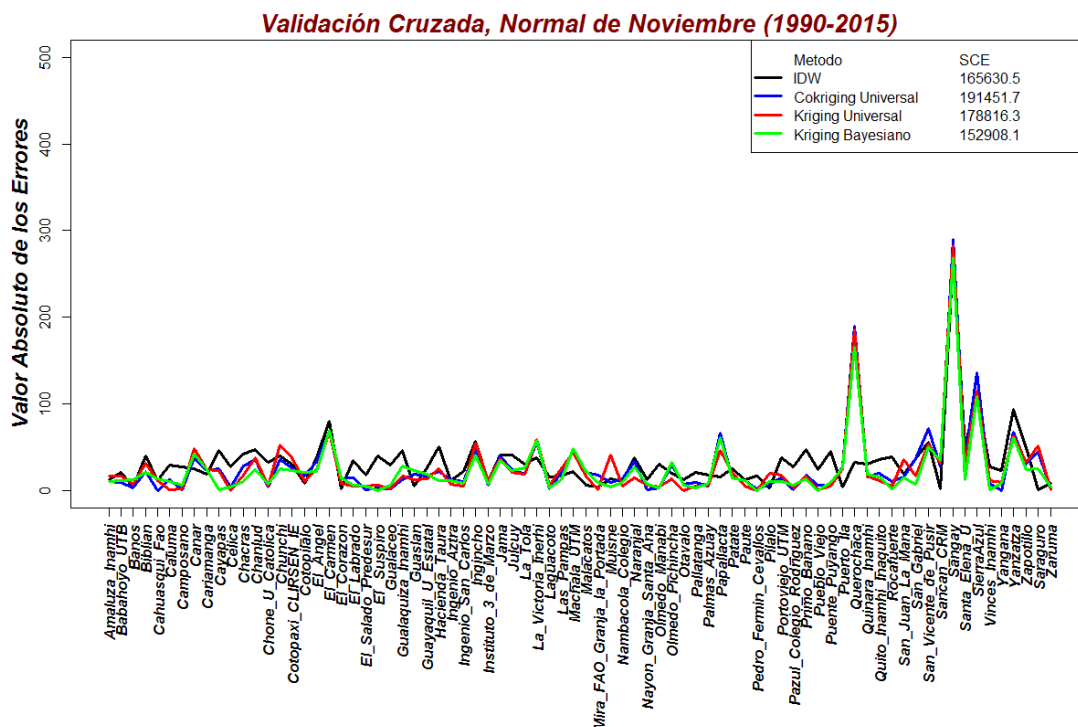


Gráfico 4.28. Validación cruzada Noviembre.  
Elaborado por: Gabriel Mantilla.

La tabla 4.13 nos muestra el modelo que mejor se ajustó al semivariograma empírico, en el caso de la ecuación del kriging universal, el modelo circular minimizo de mejor manera la suma cuadrática de los errores, con respecto a los otros modelos teóricos con 64.1. El gráfico 4.28 nos muestra que el kriging bayesiano presenta menor SCE con 152908, con respecto a los otros métodos.

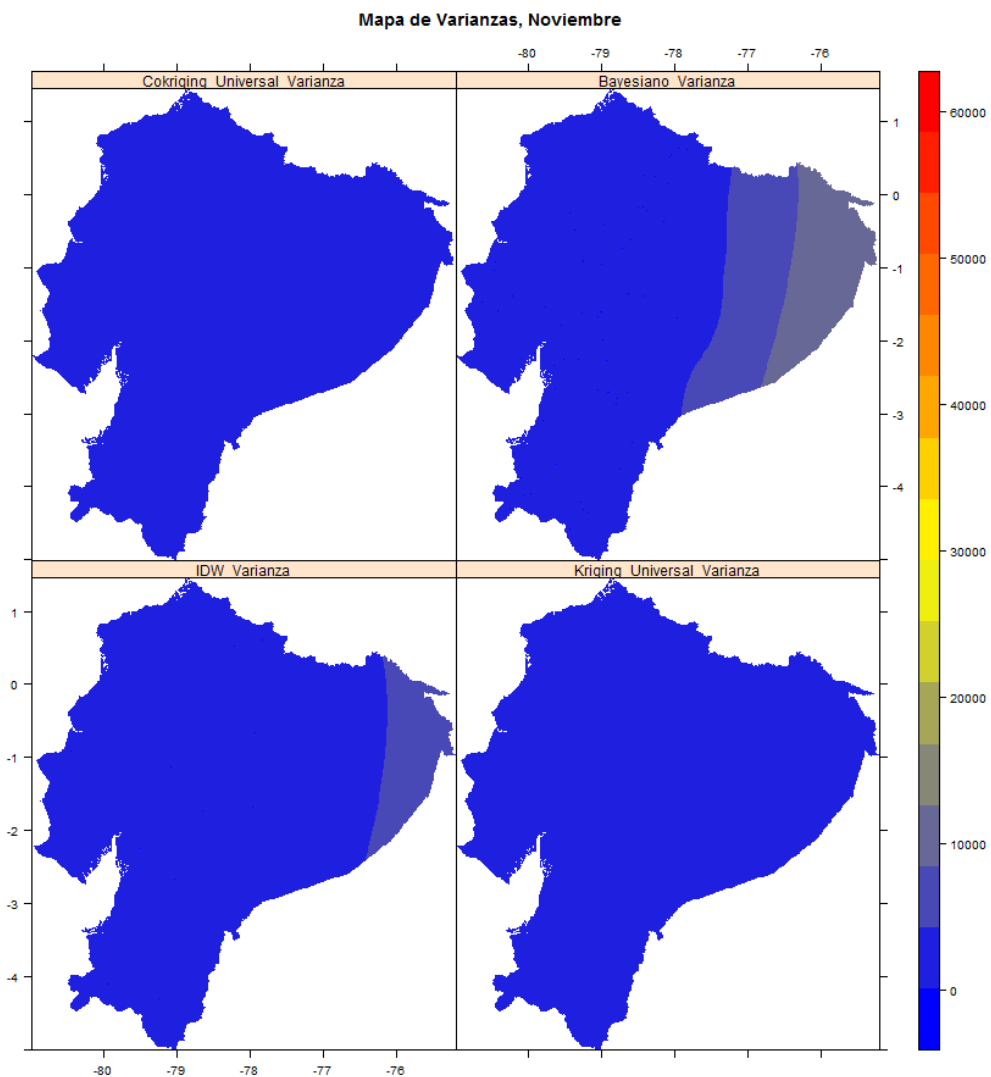


Ilustración 4.21. Varianzas de los interpoladores para Noviembre.  
Elaborado por: Gabriel Mantilla.



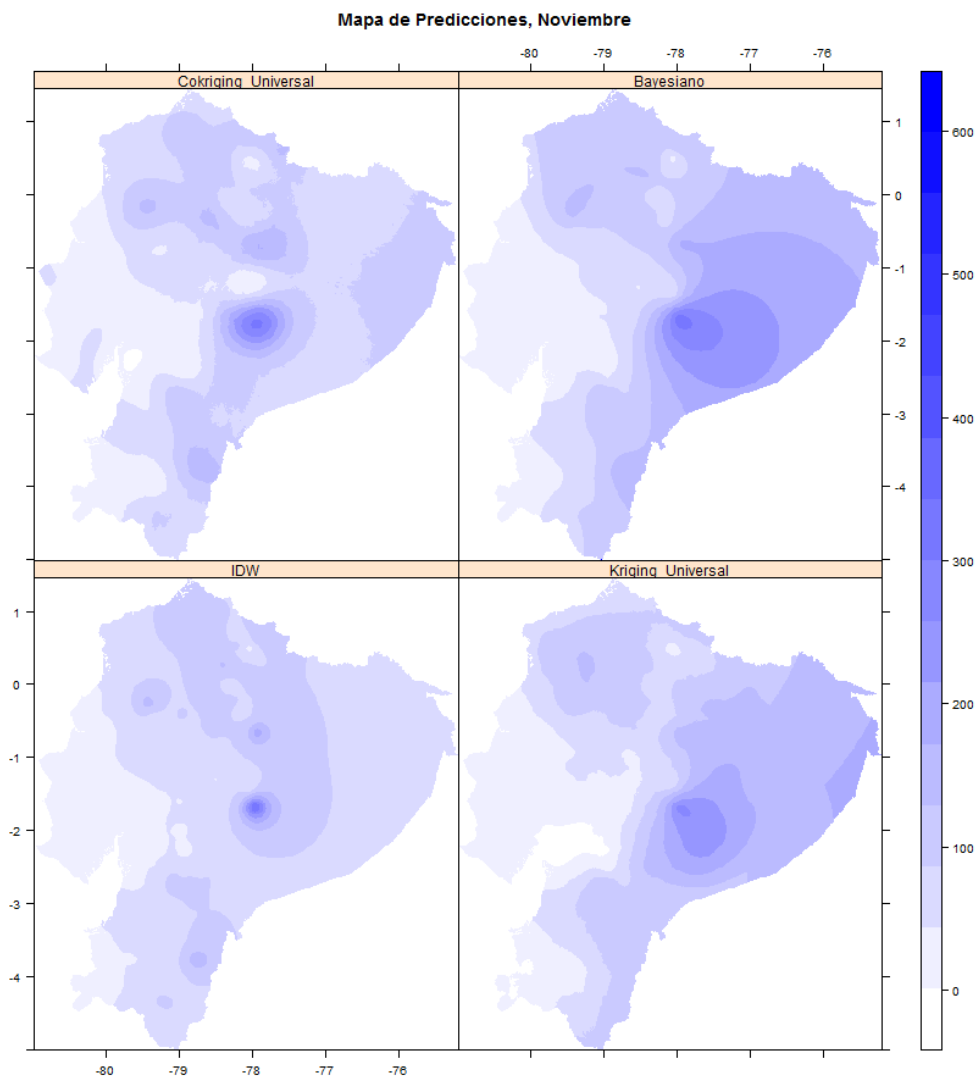


Ilustración 4.22. Interpolación de los métodos para Noviembre.  
Elaborado por: Gabriel Mantilla.

El color más azul del mapa corresponde a la estación de Sangay con 332 mm de precipitación, en el mes de noviembre se termina la época seca, los 4 métodos presentan equivalencia entre ellos en sus pronósticos, dado que los valores observados no son muy variables.

Resultados del mes de Diciembre.

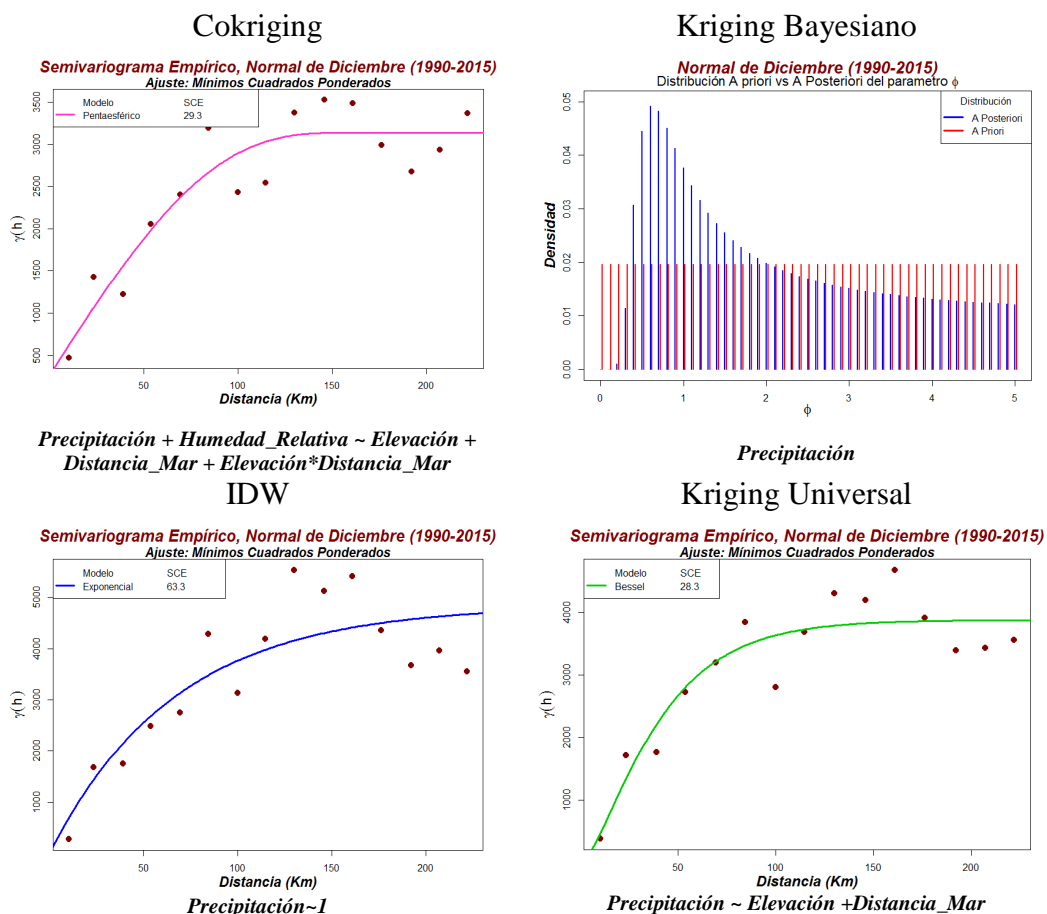


Gráfico 4.29. Medidas de Ajuste, Normal de Diciembre.

Elaborado por: Gabriel Mantilla.

Parámetro	Silo	Nugget	Rango	Modelo	SCE
Metodología					
IDW	4844.95	0	66.71	Exponencial	63.3
Kriging Universal	3876.68	0	26.77	K-Bessel	28.3
Cokriging	2869	266.20	155.18	Penta-esférico	29.3

Tabla 4.14. Parámetros de los Modelos para Diciembre.

Elaborado por: Gabriel Mantilla.

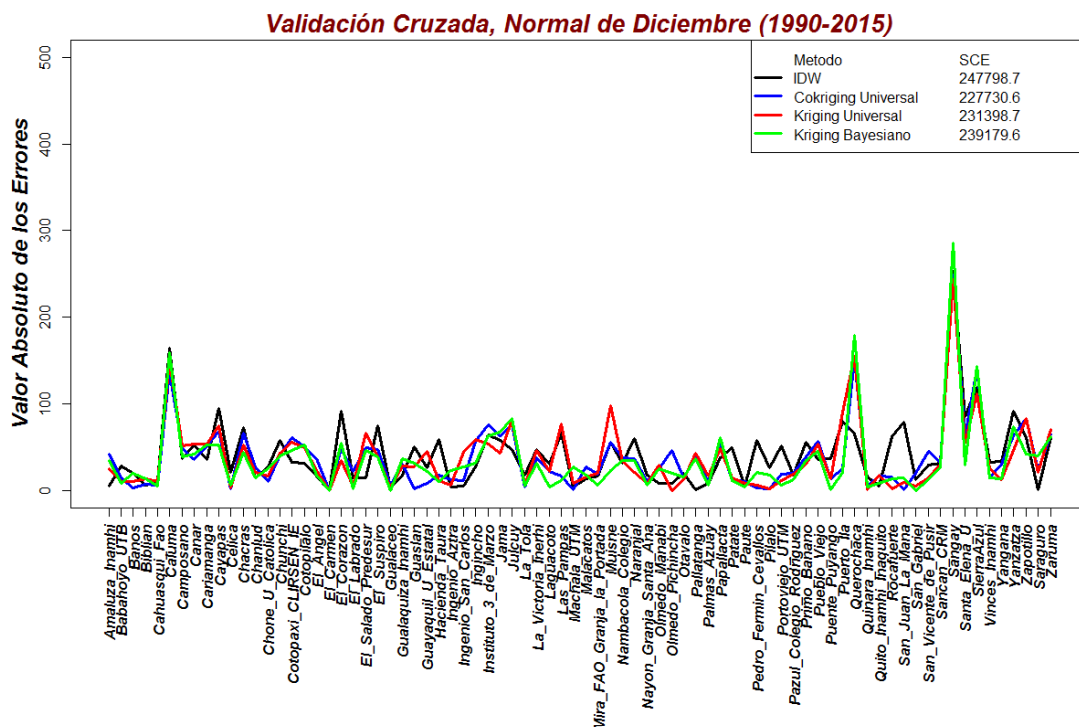


Gráfico 4.30 Validación cruzada Diciembre.  
Elaborado por: Gabriel Mantilla.

La tabla 4.14 nos muestra que el modelo que minimizo mejor el SCE fue el K-Bessel para el método de kriging universal con 28.3, haciendo que su pronóstico sea mejor ajustado. En la validación cruzada del mes de diciembre muestra que el método con menor SCE es el cokriging con 227730. El pico más alto de los errores lo presenta el pronóstico de la estación de Sangay.

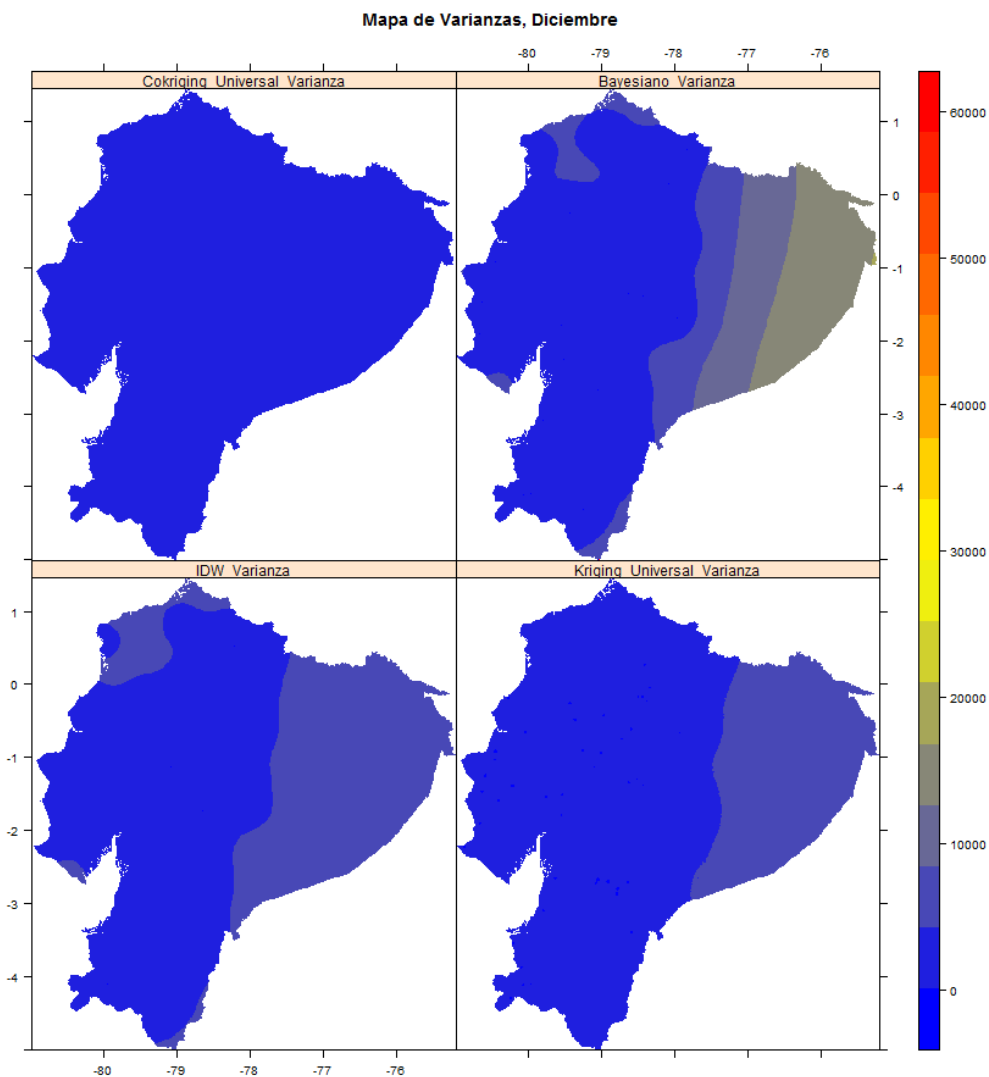


Ilustración 4.23. Varianzas de los interpoladores para Diciembre.  
Elaborado por: Gabriel Mantilla.

Las varianzas para el mes de Diciembre se mantienen bajas, en las zonas donde no se encuentran valores observados, las varianzas se incrementan ligeramente a un color azul claro, incluso el método del kriging bayesiano tiene las varianzas más altas en la zona del oriente.

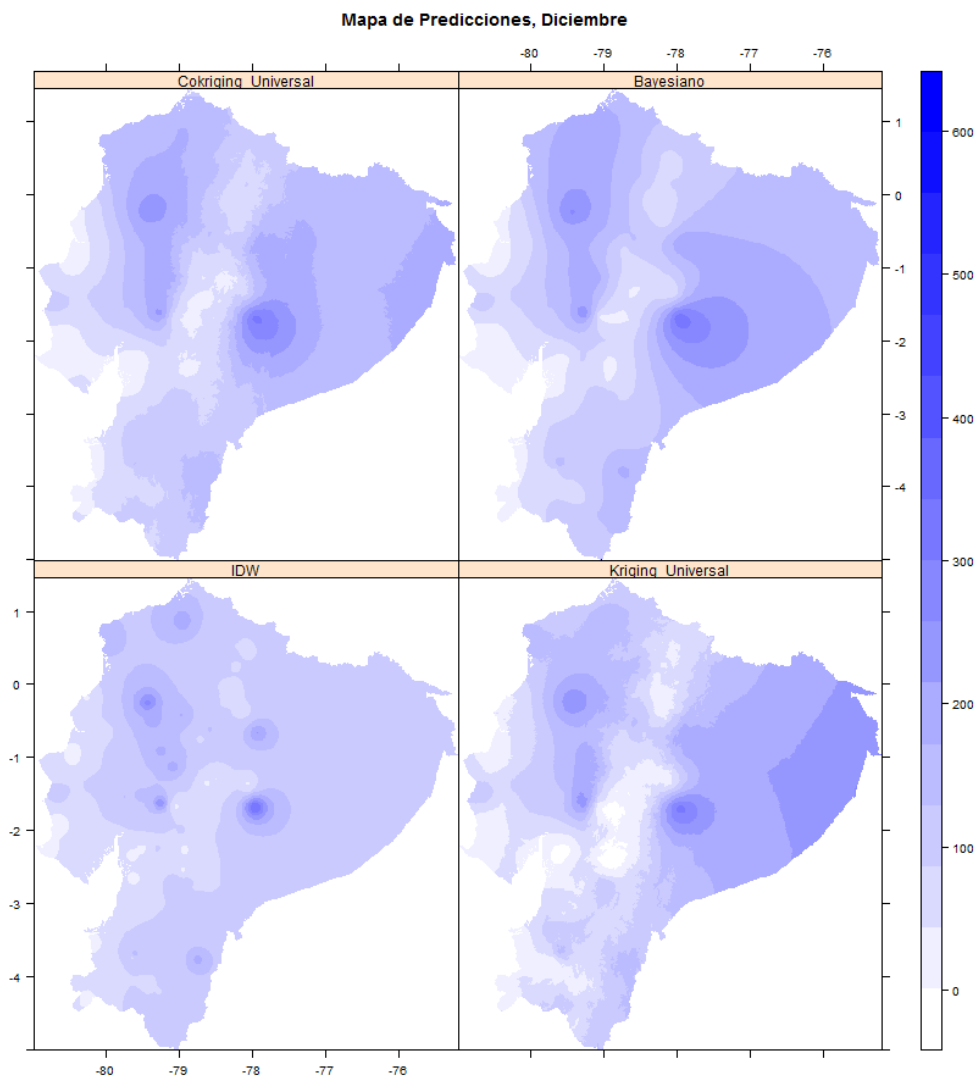


Ilustración 4.24. Interpolación de los métodos para Diciembre.  
Elaborado por: Gabriel Mantilla.

La ilustración 4.24 nos muestra el pronóstico de los 4 métodos, los valores observados se incrementan a medida que entran a la época lluviosa, El mes de diciembre es conocido como el mes de transición de la época seca a la época lluviosa. El color más azulado

## CONCLUSIONES Y RECOMENDACIONES

La variabilidad en los meses más lluviosos de diciembre a mayo se refleja en los mapas de varianza, por sus colores más intensos en la zona oriental, debido a que en este análisis no se contaba con estaciones meteorológicas en la zona Oriental de Ecuador.

El método que presentó mayor varianza es el kriging bayesiano, debido a que la distribución espacial aprende de las observaciones, por tanto al no haber estaciones la varianza se eleva considerablemente, por consiguiente está el método IDW que muestra varianzas elevadas incluso en todo el mapa donde si hay observaciones.

El método que menor varianzas presentó fue el cokriging, esto se debe al aporte de las variables de explicación que condicionan a la precipitación a medida que estamos más cerca de un cerro, más lejos de la costa y la interacción entre ambas. La humedad que esta sumada a la variable de estudio, hace posible la implementación de este método.

El método del kriging universal también presentó varianzas pequeñas, gracias al aporte de información de la altitud y la distancia con respecto al mar que se encuentra una estación meteorológica, siendo la precipitación la única variable de estudio.

La validación cruzada dio buenos resultados, nos enseñó que la suma cuadrática de los errores (SCE) es menor para unos métodos en ciertos meses que para otros métodos en los meses restantes. Para los meses lluviosos de diciembre, enero y febrero la SCE es menor utilizando el cokriging. También para los meses de marzo, abril y mayo la SCE es menor utilizando el kriging universal. Mientras que para los meses secos de junio a noviembre la SCE es menor para el kriging bayesiano.

Podemos concluir que en los meses con mayor variabilidad siendo los meses lluviosos de diciembre a mayo es recomendable utilizar un método multivariado (Kriging Universal y Cokriging) apoyándonos en variables de explicación. Mientras en los meses menos lluviosos de junio a noviembre, es recomendable usar un método estadístico basado en simulaciones como lo es el kriging bayesiano, esto tiene concordancia con los mapas de varianza.

Para los métodos es recomendable tener un número considerable de estaciones, a medida que uno se aleja más de la estación, la incertidumbre aumenta, esto lo podemos ver en los mapas de varianza.

Es útil comparar metodologías complejas con metodologías sencillas, el IDW es un método sencillo que no sobrestima los valores observados, para de esta forma poder observar en donde el kriging está sobrestimando y modificando los valores observados, esto lo podemos comprobar en los mapas de precipitación.

## BIBLIOGRAFÍA

- [1] [http://issuu.com/inamhi/docs/inamhi\\_revista\\_institucional\\_2013](http://issuu.com/inamhi/docs/inamhi_revista_institucional_2013)
- [2] <http://www.buenvivir.gob.ec/versiones-plan-nacional;jsessionid=E8A69C59F3AEA50B28E649E05C6F2E69>
- [3] <http://www.serviciometeorologico.gob.ec/la-institucion/>
- [4] Organización Meteorológica mundial, 2011. Guía de prácticas climatológicas
- [5] Romina Mezher, Pablo Mercuri, 2008. Análisis espacial y temporal de la ocurrencia de eventos de granizo sobre argentina. XV Congreso Brasileiro de Meteorología.
- [6] Donald Shepard, 1968. A two-dimensional interpolation for irregularly-spaced data.  
  
Peter A. Burrough. Principles of Geographical Information Systems for Land Resources Assessment. Monographs on Soil and Resources Survey. Clarendon Press, Oxford, U.K., 1986.  
  
Peter A. Burrough and R. McDonnell. Principles of Geographical Information Systems. Oxford University Press, Oxford, U.K., 1998.
- [7] Noel Cressie, 1990, The Origins of Kriging  
  
Lance A Waller, Carol A Gotway, 2004, Applied spatial statistics for public health data



- [8] Ramón Giraldo H, 2001. Propuesta de un indicador como variable auxiliar en el análisis cokriging.
- [9] <http://desktop.arcgis.com/es/desktop/latest/guide-books/extensions/geostatistical-analyst/what-is-empirical-bayesian-kriging-.htm>
- [10] Paulo J. Ribeiro Jr, Peter J. Diggle, Bayesian inference in Gaussian model-based geostatistics.  
<http://www.leg.ufpr.br/geoR/geoRdoc/bayeskrige.pdf>
- [11] Dennis D. Wackerly, William Mendenhall, Richard L. Scheaffer, 2008. Mathematical Statistics with Applications
- [12] Noel Cressie, 1985. Fitting Variogram Models by Weighted Least Squares.
- [13] Peter Craven, Grace Wahba, 1979. Smoothing Noisy Data with Spline Functions.
- [14] K.W. Lau, Q.H. Whu, 2002, leave one support vector out cross validation for fast estimation of generalization errors

**Bibliografía Teoría del Arte**

- ✓ Jaime Gómez, Eduardo Cassiraga, Carolina Guardiola, Javier Alvarez, Incorporation information from a digital Elevation Model for improving the Areal Estimation of Rainfal.
- ✓ A. Sarani, C. A. Cox. C.A. Madramootoo, Geostatical Methods for Prediction of Spatial Variability of Rainfall in a Mountainous Region.
- ✓ Gabriel Díaz, Ignacio Sánchez, Roberto Quiroz, Jaime Garatuza, Christopher Watts, Roberto Cruz. Interpolación Espacial de la Precipitación Pluvial en la zona de Barlovento y Sotavento del Golfo de México.
- ✓ Larry Niño, Uso de la función semivariograma y estimación kriging en el análisis espacial de un indicador entomológico de *Aedes aegypti* (Diptera: Culicidae).

## ANEXOS

### A. Glosario de Términos.

**Atípico.-** Información que es numéricamente distinta del resto de los datos.

**Altitud.-** Distancia vertical con respecto al nivel 0 del mar.

**Bimodal.-** Frecuencia de distribución de datos numéricos que presenta dos modos o picos diferentes.

**Cuartiles.-** Son los tres valores que dividen al conjunto de datos ordenados en cuatro partes porcentualmente iguales.

**Distribución.-** Función que determina a cada acontecimiento definido sobre la variable aleatoria, la probabilidad de que dicho suceso ocurra.

**Estación meteorológica.-** instalación propuesta a medir y registrar regularmente diversas variables meteorológicas.

**Hidrometeorología.-** Es una rama de la ciencias de la atmósfera y de la Hidrografía que estudia la transferencia de agua y energía entre la superficie y la atmósfera.

**Humedad Relativa.-** Relación entre la cantidad de vapor de agua que contiene y la que tendría si estuviera completamente saturada; cuanto más se aproxima a 100 representa mayor humedad.

**Interpolación.-** Obtención de nuevos puntos partiendo del conocimiento de un conjunto conocido de puntos.

**Meteorología.-** Ciencia que estudia la física de la atmósfera y su estado del tiempo.

**Milímetros.-** Unidad de medida usada para cuantificar las precipitaciones, un milímetro representa un litro de agua acumulada por cada metro cuadrado.

**Precipitación.-** Caída de agua sólida o líquida sobre la superficie terrestre debido a la condensación del vapor.

**Predecir.-** Acción de anunciar un suceso en el futuro sin ninguna validación formal del mismo.

**Pronóstico.-** Conocimiento anticipado de algún suceso de interés.

**Variograma.-** Valor o cálculo que permite analizar el comportamiento espacial de una variable sobre un área definida.

**Unimodal.-** Frecuencia de distribución de datos numéricos que presenta un solo modo o pico.

**Validación cruzada.-** Técnica utilizada para evaluar los resultados de análisis estadístico, analizando la diferencia entre los valores observados y los predichos.

## B. Código Fuente.

### Código Fuente del método IDW

```

setwd("C:/Users/gabriel/Dropbox/Materia Integradora/Metodologia/IDW")

BASE=data.frame(read.delim("BASE.txt",header=TRUE, dec=","))

coordinates(BASE)=c("long", "lat")

proj4string(BASE)=CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0")

VARIABLES=data.frame(read.delim("VARIABLES.txt",header=TRUE, dec=","))

coordinates(VARIABLES)=~x+y

proj4string(VARIABLES) <- CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0") ## for example

#varcolor=c("#FF0000", "#FF2600", "#FF4D00", "#FF7300",
"#FF9900", "#FFBF00", "#FFE600", "#FFFF00", "#F5F70D", "#EBF01A", "#E0E826",
"#D6E033", "#CCD940", "#C2D14C", "#B8C959", "#ADC266", "#A3BA73", "#99B28
0", "#8FAB8C", "#85A399", "#7A9CA6", "#7094B2", "#668CBF", "#5C85CC", "#527D
D9", "#4775E6", "#3D6EF2", "#3366FF")

#precicolor=c("#FFFFFF", "#00FFFF", "#00CCFF", "#0099FF", "#0066FF",
"#0033FF", "#0000FF", "#6600CC")

Svg=function(Co,Cs,h,a){if(h>0){Smv=Co+Cs*(1-exp(-h/a))}else{Smv=0}}

Dist=function(A,B){sapply(1:length(A$x), function(i)
sapply(1:length(B$x),function(j) if(A$x[i]==B$x[j] &
A$y[i]==B$y[j]){Smv=0}else{Smv=6374.7*acos(sin(A$y[i]*pi/180)*sin(B$y[j]*pi/

```

```

180)+cos(A$y[i]*pi/180)*cos(B$y[j]*pi/180)*cos(B$x[j]*pi/180-
A$x[i]*pi/180)))))}

Prediccion=function(Predictora,x1,y1,x2,y2,Observada,Estimada,Sill,Nuget,Rango)
{
  IDW=idw(Predictora~1, Observada,Estimada)

  AL=data.frame(x=x1,y=y1)
  BL=data.frame(x=x2,y=y2)

  D=Dist(BL,AL)
  Dij=Dist(AL,AL)

  L=sapply(1:length(x2), function(j) sapply(1:length(x1), function(i)
(1/D[i,j])/sum(1/D[1:length(x1),j])))

  Vgij=sapply(1:length(x1), function(i) sapply(1:length(x1), function(j)
Svg(Nuget,Sill,Dij[i,j],Rango)))

  Vg0=sapply(1:length(x2), function(j) sapply(1:length(x1), function(i)
Svg(Nuget,Sill,D[i,j],Rango)))

  data.frame(x=VARIABLES$x,y=VARIABLES$y,IDW=IDW$var1.pred,IDW_Vari
anza=sapply(1:length(x2), function(i) 2*sum(Vg0[,i]*L[,i]-
t(L[,i])%*%Vgij%*%L[,i]))
}

#ENERO

v1=variogram(ENE_RR~1,BASE)

```

```

Modelo1=fit.variogram(v1, vgm(psill = 18810,range = 73.8,model = "Exp",nugget
=0 ), fit.method=2)

MEXP=variogramLine(Modelo1,dist_vector=v1$dist)

SCE=round(sum(v1$np*(v1$gamma/MEXP$gamma-1)^2),1)

plot(gamma~dist,v1, xlab="Distancia (Km)",ylab=expression(gamma(h)), pch=19,
col="#800000",cex=1.4, cex.lab=1.5,font.lab=4,mgp = c(2.3, 1, 0))

title("Semivariograma Empírico, Normal de Enero (1990-
2015)",cex.main=1.7,font.main=4,col.main="#800000")

mtext("Ajuste: Mínimos Cuadrados Ponderados",cex = 1.4,font=4)

lines(variogramLine(Modelo1, 250), col="blue", lwd = 3)

legend("topleft",c("Modelo", "Exponencial", "SCE",format(SCE)),pch=c(NA,"-
",NA,NA),ncol=2, lwd=c(NA,3,NA,NA),col="blue")

Prediccion1=Prediccion(BASE$ENE_RR,BASE$long,BASE$lat,VARIABLES$x,V
ARIABLES$y,BASE,VARIABLES,18795.35,0,73.67228)

```

### **Código Fuente Kriging Universal**

```

setwd("C:/Users/gabriel/Dropbox/Materia
Integradora/Metodologia/KRIGING_UNIVERSAL")

BASE=data.frame(read.delim("BASE.txt",header=TRUE, dec=","))

coordinates(BASE)=c("long", "lat")

```

```

proj4string(BASE)=CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0")

VARIABLES=data.frame(read.delim("VARIABLES.txt",header=TRUE, dec=","))

coordinates(VARIABLES)=~x+y

proj4string(VARIABLES) <- CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0")

#KRIGING UNIVERSAL MULTIVARIADO

Variograma=function(Predictando,Predictora1,Predictora2,Sill,Rango,Nuget,BASE)
{

V=variogram(Predictando~Predictora1+Predictora2,BASE)

M11e=vgm(Sill, "Exp", Rango, Nuget)#Exponencial

M12e=vgm(Sill, "Gau", Rango, Nuget)#Gausiano

M13e=vgm(Sill, "Sph", Rango, Nuget)#Esferical

M14e=vgm(Sill, "Bes", Rango, Nuget)#Besses

M15e=vgm(Sill, "Cir", Rango, Nuget)#Circular

M16e=vgm(Sill, "Lin", Rango, Nuget)#Lineal

M17e=vgm(Sill, "Pen", Rango, Nuget)#Pentaspherical

M21e=fit.variogram(V, M11e)

```



M22e=fit.variogram(V, M12e)

M23e=fit.variogram(V, M13e)

M24e=fit.variogram(V, M14e)

M25e=fit.variogram(V, M15e)

M26e=fit.variogram(V, M16e)

M27e=fit.variogram(V, M17e)

M31e=fit.variogram(V, M21e, fit.method=2)

M32e=fit.variogram(V, M22e, fit.method=2)

M33e=fit.variogram(V, M23e, fit.method=2)

M34e=fit.variogram(V, M24e, fit.method=2)

M35e=fit.variogram(V, M25e, fit.method=2)

M36e=fit.variogram(V, M26e, fit.method=2)

M37e=fit.variogram(V, M27e, fit.method=2)

#Segun Distancia Cuadratica ajuste normal

ExpDCe=variogramLine(M21e, dist\_vector=V\$dist)

GauDCe=variogramLine(M22e, dist\_vector=V\$dist)

ShpDCe=variogramLine(M23e, dist\_vector=V\$dist)

BesDCe=variogramLine(M24e, dist\_vector=V\$dist)

CirDCe=variogramLine(M25e, dist\_vector=V\$dist)

LinDCe=variogramLine(M26e, dist\_vector=V\$dist)

PenDCe=variogramLine(M27e, dist\_vector=V\$dist)

#Segun Ponderacion

ExpPone=variogramLine(M31e, dist\_vector=V\$dist)

GauPone=variogramLine(M32e, dist\_vector=V\$dist)

ShpPone=variogramLine(M33e, dist\_vector=V\$dist)

BesPone=variogramLine(M34e, dist\_vector=V\$dist)

CirPone=variogramLine(M35e, dist\_vector=V\$dist)

LinPone=variogramLine(M36e, dist\_vector=V\$dist)

PenPone=variogramLine(M37e, dist\_vector=V\$dist)

#suma cuadratica ponderada del error Ponderacion

S31e=sum(V\$np\*(V\$gamma/ExpPone\$gamma-1)^2)

S32e=sum(V\$np\*(V\$gamma/GauPone\$gamma-1)^2)

S33e=sum(V\$np\*(V\$gamma/ShpPone\$gamma-1)^2)

S34e=sum(V\$np\*(V\$gamma/BesPone\$gamma-1)^2)

```

S35e=sum(V$np*(V$gamma/CirPone$gamma-1)^2)

S36e=sum(V$np*(V$gamma/LinPone$gamma-1)^2)

S37e=sum(V$np*(V$gamma/PenPone$gamma-1)^2)

#El metodo de minimos cuadrados ponderados

dev.off()

plot(gamma~dist,V, xlab="Distancia",
ylab=expression(gamma(h)),main=expression(frac(N(h),(gamma[theta](h))^2)),
pch=19, col="#800080" )

lines(variogramLine(M31e, 250), col="#F62217")

lines(variogramLine(M32e, 250), col="#ff00ff")

lines(variogramLine(M33e, 250), col="#808000")

lines(variogramLine(M34e, 250), col="#ffbf00")

lines(variogramLine(M35e, 250), col="#306EFF")

lines(variogramLine(M36e, 250), col="#F6358A")

lines(variogramLine(M37e, 250), col="#FFA62F")

op <- par(bg = "antiquewhite1")

legend("topleft", c("Exponencial","Gausiano","Esferical","Bessel","Circular",
"Lineal", "Pentaesferical",format(round(S31e, 1)),format(round(S32e,

```

```

1)),format(round(S33e, 1)),format(round(S34e, 1)),format(round(S35e,
1)),format(round(S36e, 1)),format(round(S37e, 1)))

,ncol = 2,col

=c("#F62217", "#ff00ff", "#808000", "#ffbf00", "#306EFF", "#F6358A", "#FFA62F"),
lty = 0,pch = "-", cex = 0.6)

min(S31e,S32e,S33e,S34e,S35e,S36e,S37e)

}

#ENERO

V1=variogram(ENE_RR~ELEVACION+DISTANCIA,BASE)

#Variograma(BASE$ENE_RR,BASE$ELEVACION,BASE$DISTANCIA,5000,100
,1000,BASE)

Modelo1=fit.variogram(V1,vgm(7896.849,"Cir",120,1960.510), fit.method=2)

MEXP1=variogramLine(Modelo1,dist_vector=V1$dist)

SCE1=round(sum(V1$np*(V1$gamma/MEXP1$gamma-1)^2),1)

plot(gamma~dist,V1, xlab="Distancia (Km)",ylab=expression(gamma(h)), pch=19,
col="#800000",cex=1.4, cex.lab=1.5,font.lab=4,mgp = c(2.3, 1, 0))

title("Semivariograma Empírico, Normal de Enero (1990-
2015)",cex.main=1.7,font.main=4,col.main="#800000")

mtext("Ajuste: Mínimos Cuadrados Ponderados",cex = 1.4,font=4)

```

```

lines(variogramLine(Modelo1, 250), col="#00cc00", lwd = 3)

legend("topleft",c("Modelo", "Circular", "SCE", format(SCE1)),pch=c(NA, "-
", NA, NA),ncol=2, lwd=c(NA, 3, NA, NA),col="#00cc00")

KU1=krige(ENE_RR~ELEVACION+DISTANCIA,BASE,VARIABLES,model=Modelo1)

KU1$var1.pred[which(KU1$var1.pred< 0)]=0

```

### **Código Fuente Cokriging.**

```

setwd("C:/Users/gabriel/Dropbox/Materia
Integradora/Metodologia/KRIGING_UNIVERSAL")

BASE=data.frame(read.delim("BASE.txt",header=TRUE, dec=","))

coordinates(BASE)=c("long", "lat")

proj4string(BASE)=CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0")

VARIABLES=data.frame(read.delim("VARIABLES.txt",header=TRUE, dec=","))

coordinates(VARIABLES)=~x+y

proj4string(VARIABLES) <- CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0") ## for example

```

```
#Variograma=function(Predictando,Predictora1,Predictora2,Predictora3,Sill,Rango,
Nuget,BASE){

  V=variogram(Predictando~Predictora1+Predictora2+Predictora3,BASE)

  M11e=vgm(Sill, "Exp", Rango, Nuget)#Exponencial

  M12e=vgm(Sill, "Gau", Rango, Nuget)#Gausiano

  M13e=vgm(Sill, "Sph", Rango, Nuget)#Esferical

  M14e=vgm(Sill, "Bes", Rango, Nuget)#Besses

  M15e=vgm(Sill, "Cir", Rango, Nuget)#Circular

  M16e=vgm(Sill, "Lin", Rango, Nuget)#Lineal

  M17e=vgm(Sill, "Pen", Rango, Nuget)#Pentaspherical

  M21e=fit.variogram(V, M11e)

  M22e=fit.variogram(V, M12e)

  M23e=fit.variogram(V, M13e)

  M24e=fit.variogram(V, M14e)

  M25e=fit.variogram(V, M15e)

  M26e=fit.variogram(V, M16e)

  M27e=fit.variogram(V, M17e)

  M31e=fit.variogram(V, M21e, fit.method=2)
```

```
M32e=fit.variogram(V, M22e, fit.method=2)
```

```
M33e=fit.variogram(V, M23e, fit.method=2)
```

```
M34e=fit.variogram(V, M24e, fit.method=2)
```

```
M35e=fit.variogram(V, M25e, fit.method=2)
```

```
M36e=fit.variogram(V, M26e, fit.method=2)
```

```
M37e=fit.variogram(V, M27e, fit.method=2)
```

```
#Segun Distancia Cuadratica ajuste normal
```

```
ExpDCe=variogramLine(M21e, dist_vector=V$dist)
```

```
GauDCe=variogramLine(M22e, dist_vector=V$dist)
```

```
ShpDCe=variogramLine(M23e, dist_vector=V$dist)
```

```
BesDCe=variogramLine(M24e, dist_vector=V$dist)
```

```
CirDCe=variogramLine(M25e, dist_vector=V$dist)
```

```
LinDCe=variogramLine(M26e, dist_vector=V$dist)
```

```
PenDCe=variogramLine(M27e, dist_vector=V$dist)
```

```
#Segun Ponderacion
```

```
ExpPone=variogramLine(M31e, dist_vector=V$dist)
```

```
GauPone=variogramLine(M32e, dist_vector=V$dist)
```

```

ShpPone=variogramLine(M33e, dist_vector=V$dist)

BesPone=variogramLine(M34e, dist_vector=V$dist)

CirPone=variogramLine(M35e, dist_vector=V$dist)

LinPone=variogramLine(M36e, dist_vector=V$dist)

PenPone=variogramLine(M37e, dist_vector=V$dist)

#suma cuadratica ponderada del error Ponderacion

S31e=sum(V$np*(V$gamma/ExpPone$gamma-1)^2)

S32e=sum(V$np*(V$gamma/GauPone$gamma-1)^2)

S33e=sum(V$np*(V$gamma/ShpPone$gamma-1)^2)

S34e=sum(V$np*(V$gamma/BesPone$gamma-1)^2)

S35e=sum(V$np*(V$gamma/CirPone$gamma-1)^2)

S36e=sum(V$np*(V$gamma/LinPone$gamma-1)^2)

S37e=sum(V$np*(V$gamma/PenPone$gamma-1)^2)

#El metodo de minimos cuadrados ponderados

dev.off()

plot(gamma~dist,V, xlab="Distancia",
ylab=expression(gamma(h)),main=expression(frac(N(h),(gamma[theta](h))^2)),
pch=19, col="#800080" )

```



```

lines(variogramLine(M31e, 250), col="#F62217")

lines(variogramLine(M32e, 250), col="#ff00ff")

lines(variogramLine(M33e, 250), col="#808000")

lines(variogramLine(M34e, 250), col="#ffbf00")

lines(variogramLine(M35e, 250), col="#306EFF")

lines(variogramLine(M36e, 250), col="#F6358A")

lines(variogramLine(M37e, 250), col="#FFA62F")

op <- par(bg = "antiquewhite1")

legend("topleft", c("Exponencial","Gausiano","Esferical","Bessel","Circular",
"Lineal", "Pentaesferical",format(round(S31e, 1)),format(round(S32e,
1)),format(round(S33e, 1)),format(round(S34e, 1)),format(round(S35e,
1)),format(round(S36e, 1)),format(round(S37e, 1)))

      ,ncol = 2,col
=c("#F62217", "#ff00ff", "#808000", "#ffbf00", "#306EFF", "#F6358A", "#FFA62F") ,
lty = 0,pch = "-", cex = 0.6)

      min(S31e,S32e,S33e,S34e,S35e,S36e,S37e)

}

```

```

V1=variogram(ENE_RR+ENE_HUME~ELEVACION+DISTANCIA+ELEVACION*DISTANCIA,BASE)

#Variograma(BASE$ENE_RR+BASE$ENE_HUME,BASE$ELEVACION,BASE$DISTANCIA,(BASE$ELEVACION)*(BASE$DISTANCIA),4000,70,2000,BASE)

Modelo1=fit.variogram(V1,vgm(4000,"Bes",70,1000), fit.method=2)

MEXP1=variogramLine(Modelo1,dist_vector=V1$dist)

SCE1=round(sum(V1$np*(V1$gamma/MEXP1$gamma-1)^2),1)

plot(gamma~dist,V1, xlab="Distancia (Km)",ylab=expression(gamma(h)), pch=19,
col="#800000",cex=1.4, cex.lab=1.5,font.lab=4,mgp = c(2.3, 1, 0))

title("Semivariograma Empírico, Normal de Enero (1990-2015)",cex.main=1.7,font.main=4,col.main="#800000")

mtext("Ajuste: Mínimos Cuadrados Ponderados",cex = 1.4,font=4)

lines(variogramLine(Modelo1, 250), col="#ff33cc", lwd = 3)

legend("topleft",c("Modelo","Bessel","SCE",format(SCE1)),pch=c(NA,"-",NA,NA),ncol=2, lwd=c(NA,3,NA,NA),col="#ff33cc")

rm(g,v)

g = gstat(NULL, "ELEVACION", ENE_RR+ENE_HUME~ELEVACION,BASE)

g = gstat(g, "DISTANCIA", ENE_RR+ENE_HUME~DISTANCIA,BASE)

```

```

g = gstat(g, "MODELO",
ENE_RR+ENE_HUME~ELEVACION+DISTANCIA+ELEVACION*DISTANCIA
,BASE)

v = variogram(g)

g = gstat(g, model = Modelo1, fill.all = TRUE)

g.fit = fit.lmc(v, g)

CO=predict(g.fit,VARIABLES)

CO$MODELO.pred[which(CO$MODELO.pred< 0)]=0

CO$CORRELACION=CO$cov.ELEVACION.DISTANCIA/(sqrt(CO$ELEVACIO
N.var)*sqrt(CO$DISTANCIA.var))

```

### **Código Fuente Kriging Bayesiano.**

```

setwd("C:/Users/gabriel/Dropbox/Materia
Integradora/Metodologia/KRIGING_UNIVERSAL")

BASE=data.frame(read.delim("BASE.txt",header=TRUE, dec=","))

coordinates(BASE)=c("long", "lat")

proj4string(BASE)=CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0")

VARIABLES=data.frame(read.delim("VARIABLES.txt",header=TRUE, dec=","))

```

```

coordinates(VARIABLES)=~x+y

proj4string(VARIABLES) <- CRS("+proj=longlat +datum=WGS84 +ellps=WGS84
+towgs84=0,0,0")

#ENERO

DATA=data.frame(read.delim("BASE.txt",header=TRUE, dec=","))

BASEE=data.frame(x=DATA$long,y=DATA$lat,data=DATA$ENE_RR)

BASEE=as.geodata(BASEE)

DATOSNUEVOS=data.frame(VARIABLES$x,VARIABLES$y)

bsp=krige.bayes(BASEE,loc=DATOSNUEVOS,prior=prior.control(phi.discrete=seq
(0,5,l=51)),output=output.control(n.predictive=0))

A=data.frame(A1=bsp$posterior$phi$distribution,row.names = seq(0,5,l=51))

B=data.frame(B1=bsp$prior$phi$probs,row.names=seq(0,5,l=51)+0.02)

plot(row.names(A),A$A1,type="h",col="blue",lwd=2,cex.lab=1.5,font.lab=4,mgp =
c(2.3, 1, 0),xlab=expression(phi),ylab="Densidad")

lines(row.names(B),B$B1,type="h",col="red",lwd=2)

title("Normal de Enero (1990-
2015)",cex.main=1.7,font.main=4,col.main="#800000")

mtext(expression(paste("Distribución A priori vs A Posteriori del parametro
",phi)),cex = 1.4,font=4)

```

```
legend("topright",title="Distribución",c("A Posteriori","A Priori"),pch=c("-", "-"),ncol=1, lwd=c(3,3),col=c("blue","red"))
```