

UNIVERSIDAD DE CUENCA



FACULTAD DE CIENCIAS QUÍMICAS  
ESCUELA DE INGENIERÍA AMBIENTAL

“UTILIZACIÓN DE HERRAMIENTAS GEOESTADÍSTICAS PARA LA GENERACIÓN  
DE INFORMACIÓN ESPACIAL METEOROLÓGICA EN LA PROVINCIA DEL AZUAY”

Trabajo de titulación previo a la obtención  
del título de Ingeniero Ambiental

AUTORA:

STEFANIA TANIA CAGUANA REYES

C.I. 0107424103

DIRECTOR:

BLGO. JULIO DANILO MEJÍA CORONEL, MSc

C.I. 0103638581

CUENCA – ECUADOR

2017



## RESUMEN

La información meteorológica se ha convertido, en la actualidad, en un factor de suma importancia para el desarrollo de varias actividades humanas, pues permite tener conocimiento de las características climáticas de un lugar, facilitando llevarlas a cabo con éxito. Y una de las herramientas que permite obtener este tipo de datos de lugares en los que no existen estaciones meteorológicas, son los métodos geostatísticos. El presente trabajo de titulación tuvo como objetivo generar información de precipitación y temperatura de la Provincia del Azuay (zona noreste), a través de un análisis geoestadístico en base a datos a escala horaria recolectados en 16 estaciones meteorológicas pertenecientes a la Corporación Eléctrica del Ecuador (CELEC) y a la Universidad Politécnica Salesiana (UPS). Los métodos de análisis empleados fueron Kriging y Cokriging, utilizando como herramienta de modelado el software ArcGis. Los resultados indican una temperatura promedio-hora en un rango de 5.86 a 20.23 °C y una precipitación promedio-hora en un rango de 0.03 a 0.19mm para las distintas zonas del área de estudio durante el período de análisis. Además el proceso demostró que el método de las proporcionalidades y el de la media aritmética son adecuados para completar la información meteorológica faltante y que el método Kriging Ordinario Exponencial y Gaussiano son los modelos que mejor se ajustan a las variables de estudio a escala temporal horaria. Los resultados obtenidos serán compartidos e implementados en un geoportal, el cual permite el acceso a esta información depurada a diferentes usuarios.

**Palabras clave:** geostatística, Kriging, Cokriging, precipitación, temperatura, método de las proporcionalidades, media aritmética.



## ABSTRACT

Meteorological information has become, at present, in a factor of great importance for the development of several human activities, since it allows to know the climatic characteristics of a place, facilitating to carry out this activities with success. One of the tools that allow to obtain this type of data of places in which there are no meteorological stations, are the geostatistic methods. The present research work had the objective of generate precipitation and temperature information from the Azuay Province (noreast zone), through a geostatistical analysis based on hourly data collected in 16 meteorological stations belonging to the Corporación Eléctrica del Ecuador (CELEC) and to the Universidad Politécnica Salesiana (UPS). The methods of analysis used were the Kriging and Cokriging, using as modeling tool the ArcGis software. The results indicate an average-hour temperature in a range of 5.09 to 20.23 °C and an average-hour rainfall in a range of 0.03 a 0.19mm for the different zones of the province. In addition, the process showed that the proportionality method and the arithmetic mean method are adequate to complete the missing meteorological information and that the Exponential and Gaussian Ordinary Kriging method are the models that best fit the study variables at the time scale. The results obtained will be shared and implemented in a geoportal, which allows the access to this information purified to different users.

Keywords: geostatistics, Kriging, Cokriging, precipitation, temperature, proportionality method, arithmetic mean.



## ÍNDICE

RESUMEN.....	2
ABSTRACT .....	3
1. INTRODUCCIÓN.....	10
1.1 Antecedentes .....	10
1.2 Justificación .....	10
1.3 Objetivos .....	11
2. MARCO TEÓRICO.....	13
2.1 Estudios antecedentes .....	13
2.1.1 Relleno de información climática .....	13
2.1.2 Geostatística .....	15
2.1.3 Validación del método de interpolación .....	16
2.1.4 Sistemas de información geográfica (SIG).....	16
2.2 Conceptos y Teorías.....	17
3. METODOLOGÍA.....	23
3.1 Tipo de estudio.....	23
3.2 Área de estudio.....	23
3.3 Recopilación de información.....	24
3.3.1 Variables de estudio.....	24
3.3.2 Tipo de información: Secundaria .....	24
3.3.3 Número de estaciones meteorológicas.....	24
3.3.4 Período de análisis.....	25
3.3.5 Análisis estadístico.....	25
3.3.6 Programa de análisis .....	26
3.4 Métodos.....	26
3.4.1 Relleno de información meteorológica.....	26
3.4.2 Aplicación de Kriging y Cokriging .....	31
3.4.3 Validación del modelo geoestadístico.....	34
3.4.4 Publicación de resultados .....	35
4. RESULTADOS Y DISCUSIÓN.....	36
4.1 Relleno de información meteorológica.....	36
4.1.1 Análisis de datos.....	36
4.1.2 Determinación de la calidad de los datos.....	37
4.1.3 Identificación del patrón de datos faltantes .....	38



4.1.4 Estimación de valores faltantes .....	39
4.2 Aplicación de Kriging y Cokriging .....	41
4.2.1 Análisis exploratorio de los datos .....	42
4.2.2 Análisis estructural de los datos.....	43
4.2.3 Validación del modelo Kriging y Cokriging .....	56
4.2.4 Publicación de resultados .....	59
5. CONCLUSIONES.....	61
6. RECOMENDACIONES Y LIMITACIONES .....	62
7. BIBLIOGRAFÍA.....	63
8. ANEXOS .....	69
Anexo 1: Zona de estudio .....	69
Anexo 2: Correlaciones entre las estaciones.....	72
Anexo 3: Estaciones con datos faltantes vs Estaciones con datos completos .....	73
Anexo 4: Temperatura y Precipitación mensual.....	77

## ÍNDICE DE ILUSTRACIONES

<i>Ilustración 1: Estudio Geostadístico .....</i>	21
<i>Ilustración 2: Mapa de ubicación del área de estudio .....</i>	23
<i>Ilustración 3: Proceso de relleno de datos meteorológicos.....</i>	26
<i>Ilustración 4: Tabla de interpretación de valores de correlación .....</i>	29
Ilustración 5 Resumen global datos meteorológicos perdidos.....	37
Ilustración 6: Semivariograma empírico .....	44
Ilustración 7: Anisotropía.....	45
Ilustración 8: Análisis exploratorio: tendencia.....	46
Ilustración 9: Modelos de semivariograma Kriging-Precipitación.....	51
Ilustración 10: Modelos de Semivariograma Kriging-Temperatura .....	53
Ilustración 11 Modelos de Semivariograma Cokriging-Precipitación.....	54
Ilustración 12: Modelos de semivariograma Cokriging-Temperatura .....	55
Ilustración 13: Mapas de error.....	58
Ilustración 14: Geoportal-Provincia del Azuay.....	60
Ilustración 15 Parroquias pertenecientes al área de estudio .....	71

## ÍNDICE DE TABLAS

Tabla 1: Variables meteorológicas de estudio.....	24
Tabla 2: Estaciones Meteorológicas UPS-CELEC.....	24
Tabla 3: Información precipitación y temperatura de estaciones meteorológicas .....	27
Tabla 4 Información estaciones: Kriging-Cokriging .....	31
Tabla 5: Estaciones meteorológicas (% datos faltantes) .....	36
Tabla 6: Estaciones con outliers para variable precipitación .....	38
Tabla 7: Tabla de Medias marginales- Precipitación.....	38



## Universidad de Cuenca

Tabla 8: Tabla de Medias marginales- Temperatura.....	38
Tabla 9: Grupos de estaciones para el método Media Aritmética.....	40
Tabla 10: Estadística de relleno: ME, MEA, RMSE .....	40
Tabla 11: Análisis exploratorio: Histograma.....	42
Tabla 12: Prueba de normalidad-transformación logarítmica .....	43
Tabla 13: Prueba de normalidad- transformación 1/x. Precipitación .....	43
Tabla 14: Semivariograma teórico Precipitación.....	47
Tabla 15: Semivariograma teórico Temperatura.....	48
Tabla 16: Prueba normalidad-altura.....	51
Tabla 17: Parámetros estadísticos Validación.....	56



Universidad de Cuenca

### Cláusula de Propiedad Intelectual

---

Yo Stefania Tania Caguana Reyes, autora del trabajo de titulación "UTILIZACIÓN DE HERRAMIENTAS GEOESTADÍSTICAS PARA LA GENERACIÓN DE INFORMACIÓN ESPACIAL METEOROLÓGICA EN LA PROVINCIA DEL AZUAY" certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autora.

Cuenca, 13 de noviembre del 2017

A handwritten signature in blue ink, reading "Stefania Tania Caguana Reyes", written over a horizontal line.

Stefania Tania Caguana Reyes

C.I: 0107424103



Universidad de Cuenca

## Cláusula de licencia y autorización para publicación en el Repositorio Institucional

---

Stefania Tania Caguana Reyes en calidad de autora y titular de los derechos morales y patrimoniales del trabajo de titulación "UTILIZACIÓN DE HERRAMIENTAS GEOESTADÍSTICAS PARA LA GENERACIÓN DE INFORMACIÓN ESPACIAL METEOROLÓGICA EN LA PROVINCIA DEL AZUAY", de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el repositorio institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, 13 de noviembre del 2017

A handwritten signature in blue ink, reading "Stefania Tania Caguana Reyes", positioned above a horizontal line.

Stefania Tania Caguana Reyes

C.I: 0107424103





## AGRADECIMIENTO

*A Dios, por haberme guiado, por estar junto a mí y darme la fortaleza para seguir adelante y nunca darme por vencida.*

*A mis padres, Juan y Mariana, los pilares de mi vida, por todo el amor, apoyo, confianza y comprensión que me han dado a lo largo de estos años, sin el cual no hubiese podido lograr todas las metas que me he propuesto.*

*A mis hermanos, Esther, Magaly y Cristian por siempre estar a mi lado y motivarme a seguir adelante y dar lo mejor de mí. Gracias por todos los consejos y comprensión.*

*A mi mejor amigo y novio, Jhonny gracias por estar conmigo en los momentos difíciles, por tu apoyo, por escucharme, por todo el cariño que me das y por toda la ayuda que me has brindado para culminar esta meta.*

*A mis sobrinos, Karen, Jeancarlos, Linda, Nahomi, Andrea y Aarón por estar junto a mí, compartir tantos lindos momentos y siempre sacarme una sonrisa.*

*A mis amigas, Claudia y Lisseth, por todas las experiencias que hemos vivido juntas, por siempre estar pendiente de mí y por todos estos años en los que nuestra amistad nos ha motivado y ayudado a lograr una meta más.*

*A mi director, Blgo. Danilo Mejía por la confianza puesta en mí durante la realización de este trabajo y por todo el tiempo y paciencia invertidos durante la realización del mismo.*

*Al Ing. Alex Áviles, por todas las recomendaciones y sugerencias dadas para mejorar la investigación realizada.*

*A las instituciones, Corporación Eléctrica del Ecuador y a la Universidad Politécnica Salesiana por facilitarme el acceso a los datos necesarios para llevar a cabo este trabajo de investigación.*



## 1. INTRODUCCIÓN

### 1.1 Antecedentes

La falta de información meteorológica de diversos lugares constituye un problema en la actualidad, pues esta información influye significativamente en la población y en todos los sectores de la sociedad.

El conocer el estado de la atmósfera posibilita las labores de vigilancia necesarias en distintas actividades; situaciones de emergencia por accidentes químicos o físicos, monitoreo de incendios o cualquier tipo de actividad que se desarrolle al aire libre, son actividades en las que el comportamiento de la meteorología de un lugar pueda causar problemas. Para ello es necesario conocer cuál ha sido, es o será el comportamiento del medio atmosférico. Tener esta información ayuda a las poblaciones a estar preparadas para mitigar o disminuir los impactos que pueden ocasionar los desastres naturales y permite a su vez que las sociedades se desarrollen en ámbitos tanto económicos como productivos de manera satisfactoria.

Actualmente varias instituciones instalan y operan estaciones meteorológicas automáticas a lo largo del país. Sin embargo, no existen los recursos económicos necesarios para implementarlas en puntos remotos y zonas sin información, por lo que se hace necesario buscar métodos de transmisión, procesamiento y cálculo o métodos estadísticos que permitan extrapolar datos para generar la información meteorológica desconocida de estos lugares.

Por ello, se hace necesario utilizar métodos de investigación más prácticos, robustos y económicos, que permitan obtener estos datos sin la necesidad de instalar estaciones. Y una de las herramientas más empleadas es la geoestadística, la cual posibilita la aplicación del conocimiento meteorológico para diversos fines.

### 1.2 Justificación

La información sobre meteorología es indispensable para todos los países que utilizan esta información en la toma de decisiones. Predecir cuál será el comportamiento del clima o tiempo en una región constituye la clave para proyectos tales como la



## Universidad de Cuenca

implementación de energía solar, térmica o eólica; así como en actividades productivas como la agricultura, las industrias, el turismo, etc.; pues de las variables meteorológicas dependerá su éxito o fracaso. Por ello la información meteorológica constituye un elemento clave para el desarrollo de una nación.

El conocimiento de esta información permite a los gobiernos y a las demás partes interesadas minimizar los costos en caso de desastres naturales, mediante medidas preventivas y de mitigación ante sus efectos. Por todo ello, se ve la necesidad de buscar métodos y tecnologías que permita obtener y difundir la información de forma rápida y accesible para usuarios públicos y privados.

Una forma de compartir esta información son los geoportales o Portales Geoespaciales, en los cuales se puede acceder a datos vía Internet. La información que existe en los geoportales es muy variada y se puede encontrar datos informativos, turísticos, etc. Esta herramienta permite el intercambio de información entre las diferentes instituciones, ciudadanos y agentes sociales de toda índole.

Por lo tanto, generar este tipo de información de manera clara y precisa, así como compartirla para que pueda ser utilizada por diversos usuarios de manera gratuita y oportuna, constituyen factores importantes para el desarrollo de un país.

En la presente investigación se pretendía generar información de temperatura y precipitación de toda la provincia del Azuay, sin embargo al no contar con datos en la zona norte, central y sur de la provincia, se procedió a generar la información de la zona noreste de la misma, área en la cual se contaba con información.

### 1.3 Objetivos

#### *General*

-Generar información de precipitación y temperatura de la zona noreste de la Provincia del Azuay a través un análisis geoestadístico.

#### *Específicos*

-Rellenar información de temperatura y precipitación en estaciones con vacíos de información.



## Universidad de Cuenca

-Implementar un método de interpolación para obtener datos de precipitación y temperatura en lugares no existentes de esta información a través datos de las estaciones adyacentes.

-Validar el método de interpolación con los datos y publicar la información obtenida en un geoportal a nivel web para su Acceso e Inferencia Geoestadística.



## 2. MARCO TEÓRICO

### 2.1 Estudios antecedentes

Para varios investigadores, como Membreño & Zavala (2015), una de las grandes necesidades que se presentan en los países en los últimos años, es la realización de estudios donde intervengan variables climáticas como la precipitación y la temperatura, pues permiten el análisis del comportamiento continuo del territorio, facilitando las diferentes actividades que en él se practican

Así también, para Guerra, Gómez, González, & Zambrano (2009) la precipitación y la temperatura son procesos atmosféricos muy importantes en la caracterización de las condiciones ambientales y climáticas de cualquier área terrestre, por lo que el estudio de su distribución y variabilidad necesita ser revisada y evaluada permanentemente, para así ser utilizada en diversos fines de investigación, planificación, ordenación del territorio, prevención de riesgos, entre otros.

Estudios que recalcan la importancia de conocer y evaluar las variables climáticas en cualquier proyecto a seguir. Sin embargo, en mucho de los casos, esta información es inexistente o cuenta con muchos vacíos por lo que se requiere rellenar o compensar esta información de alguna manera.

#### 2.1.1 Relleno de información climática

Uno de los problemas que se puede presentar durante el análisis de información climática es la ausencia de datos en las estaciones meteorológicas existentes en las áreas de estudio. Si los datos son recolectados y almacenados por algún sistema automático, pueden ocurrir fallos en el automatismo, ocasionados por la falta de energía, o por algún tipo de mala configuración del mecanismo (Ferreira y Ocerín, 2005).

Generalmente los datos de precipitación y temperatura obtenidos mediante instrumentos de medición, deben recibir tratamientos específicos con el objetivo de hacerlos más confiables y que puedan ser utilizados para procesos más complejos como la geoespacialización. Por ello diversas técnicas se han desarrollado a través del tiempo para procesar estas variables climáticas tales como la verificación de outliers o de datos faltantes (Guerra, Gómez, González, & Zambrano, 2006a).



## Universidad de Cuenca

Todo aquel que trabaja con series de tiempo climáticas se encuentra con el problema que en muchos casos las series están incompletas. Algunos métodos de análisis pueden adecuarse a esta situación pero otros requieren que las series estén completas. Por ello es necesario aplicar ciertos métodos que permitan completar series de datos, ya que los vacíos en las series temporales pueden comprometer severamente la utilidad de la información que se requiera utilizar para distintos fines. Campozano, Sánchez, Avilés, y Samaniego (2015), en una investigación realizada en los Andes ecuatorianos manifiestan que para el relleno de datos de precipitación a escala diaria la mejor técnica es la regresión lineal múltiple y para los datos de temperatura a escala diaria la mejor opción es la media climatológica del día.

Para otros autores como Urrutia, Palomino y Salazar (2010) el uso de los modelos de relleno de información, deben ser aplicados cuando se tienen datos faltantes menor o igual al 20%, mientras que si este porcentaje aumenta recomiendan hacer uso de otros modelos de imputación más robustos.

Otros de los aportes importantes en este sentido, es el realizado por Caldera, Piyathisse, & Nandalal (2016), quienes en su investigación concluyen que para las estaciones meteorológicas objetivo que tienen sólo una estación vecina con un alto coeficiente de correlación, el método de Regresión Lineal da buenas predicciones.

Por lo tanto, y en base a la información descrita, el método de regresión lineal utilizando la información de estaciones vecinas con alta correlación, puede ser considerado como un método óptimo que ofrece en el relleno de información faltante con menor error para este tipo de variables.

Para el caso de estaciones vecinas que tienen baja correlación, el método de proporcionalidades que utiliza la información de la misma estación de relleno resulta ser adecuado (Lituma & Andreina, 2011).

Monte Alegre (1990), manifiesta que el método de proporcionalidades es uno de los más utilizados en aquellos casos en que no existen datos de comparación y por tanto, la serie tiene que servir de referencia para el relleno de datos faltantes de sí misma.

El método de proporcionalidades a pesar de en ciertas ocasiones producir mayores errores, también produce estimaciones insesgadas y minimiza varianzas (Congreso, 2004).

Una vez que se han aplicado los métodos de relleno, se tiene una base de datos completa y confiable con la cual se puede trabajar y aplicar los métodos geostadísticos más apropiados.



### 2.1.2 Geostatística

Con el pasar del tiempo y el avance de la ciencia se generó una nueva tendencia en los estudios estadísticos para variables que se encuentran distribuidas espacialmente, esta tendencia es conocida como geoestadística.

Se conoce como geoestadística a la parte la estadística descriptiva que permite estudiar fenómenos tomando en cuenta su ubicación en el tiempo y en el espacio. Generalmente, el tratamiento de esta información se realiza a través de medidas descriptivas de tendencia central, de posición y dispersión o variabilidad, de análisis de correlación y regresión. Mientras que para el análisis de variables la geoestadística utiliza métodos de interpolación, encontrándose metodologías específicas para cada zona dependiendo de la variable de interés que se esté estudiando y permitiendo además mejorar la representación espacial de variables climatológicas (González, Guerra, & Gómez, 2007).

La geoestadística se ha convertido, en una herramienta útil para explicar la variación de propiedades en el espacio y su uso se ha generalizado recientemente y de manera particular para estimar valores de variables en sitios donde no existe información (Quevedo y Sánchez, 2009).

Autores como Chávez y Moreano (2013) manifiestan que los métodos geostatísticos constituyen una herramienta útil para atender de manera ágil consultas de la comunidad dirigidas a los servicios meteorológicos nacionales, así como también para evaluaciones climáticas y en especial para determinar las características del clima en áreas con insuficientes registros históricos.

Como parte de la Geoestadística, la Geoespacialización de las variables climáticas, sobre todo para modelos que requieren el uso de datos diarios, es un problema complejo cuya solución comprende varios aspectos básicos entre los que destacan la necesidad de una mejor densidad y distribución espacial de las estaciones climatológicas, contar con datos de largo registro y garantizar la continuidad y calidad de las mediciones (Machado, Puche, Vilorio, Parra, & Primavera, 2007).



Por ello, para obtener mejores resultados es necesario que se cuente con una red meteorológica amplia, distribuida a lo largo de la zona de estudio y que cuente con gran cantidad de datos.

Aplicados los métodos geostatísticos de interpolación más apropiados para los datos que se disponen, es necesario que se realice una validación de los resultados obtenidos, para ello se han desarrollado varias técnicas como la validación cruzada.

### 2.1.3 Validación del método de interpolación

Existen varias técnicas para determinar qué tan confiable es el método utilizado y que tan certeras fueron las decisiones tomadas. El método más utilizado es la validación cruzada (técnica de Jackknife), la cual consiste en eliminar uno de los valores de la variable de estudio, calcular el semivariograma de estos datos y estimar el valor que se eliminó anteriormente.

En una investigación realizada por Adisoma (1993) se determina que la utilización de la técnica Jackknife complementa mejor el uso del Kriging Ordinario, pues muestra superioridad en comparación con otras técnicas de validación y además maximiza la utilización de los datos estimados.

Para realizar una validación cruzada es recomendable que la base de datos con la que se haya realizado la interpolación sea amplia, pues de ello dependerá que la eficacia en la determinación de la funcionalidad del método utilizado (Zhang & Wang, 2010).

Para optimizar los resultados de la interpolación geostatística Pesquer, Masó y Pons (2007) sugieren realizar analizar las variables meteorológicas conjuntamente con técnicas de interpolación espacial y procedimientos de análisis SIG.

### 2.1.4 Sistemas de información geográfica (SIG)

Burrough y McDonnell, 1998 citados en García-Martí et al (2004) expresan que el desarrollo de los sistemas de información geográfica (SIG) y la aplicación de la geoestadística han supuesto un nuevo impulso al análisis de la distribución espacial aplicada a la ecología y a las ciencias ambientales.





## Universidad de Cuenca

El Instituto Geográfico militar define un SIG como Sistemas de Información Geográfica que son el resultado de la aplicación de las llamadas Tecnologías de la Información (TI) a la gestión de la Información Geográfica (IG).

Para Abad (2014) las tecnologías de información geográfica, como los SIG y la geoestadística son herramientas robustas que permiten apoyar a los expertos en la obtención de mapas continuos de precipitación y temperatura a partir de datos puntuales como son los registros de las estaciones de medición.

La utilización de los Sistemas de Información geográficas resultan de este modo una herramienta útil pues permite representar la información de una manera más didáctica mejorando el análisis de información resultante.

### 2.2 Conceptos y Teorías

En el presente trabajo de titulación se manejan varios términos estadísticos cuyos conceptos se resumen en el siguiente apartado:

#### Estadística

“Es el arte de aprender a partir de los datos. Está relacionada con la recopilación de datos, su descripción subsiguiente y su análisis, lo que no lleva a extraer conclusiones” (Ross, 2007).

#### Método de las proporcionalidades

Es uno de los métodos más utilizados en casos en que no existen datos de comparación y por lo tanto, la serie de datos tiene que servir de referencia para el relleno de datos faltantes de sí misma.

#### Método de la media aritmética

Es el método en el cual se utiliza la media (medida estadística) de la serie de datos registrados en un intervalo de tiempo y con este valor medio se rellena los espacios sin información, constituyendo de este modo el método más simple de relleno de información faltante (Guzman & León, 2012).



## Estadística espacial

La estadística espacial es la parte de la estadística encargada de realizar un análisis descriptivo e inferencial de datos geográficos, mediante un estudio de la relación de dependencia o auto correlación espacial de las observaciones de las variables de estudio (Martori, 2007).

La estadística espacial se subdivide en tres grandes áreas, una de ellas es la geoestadística (“Geoestadística - EcuRed,” 2010).

### Geoestadística

“La geoestadística se refiere al conjunto de técnicas estadísticas utilizadas para el estudio de variables numéricas distribuidas en el espacio” (González, Guerra, & Gómez, 2007).

Para Díaz (2002) la geoestadística tiene como objeto analizar y modelar la variabilidad espacial y predecir los fenómenos naturales en el espacio y en el tiempo. Mientras que para Carrasco (2014) el objetivo es caracterizar e interpretar el comportamiento de los datos que están distribuidos espacialmente.

La geoestadística es aplicada en varias áreas para estimar y simular variables regionalizadas. Comprende múltiples técnicas que se pueden clasificar 5 grupos: geoestadística lineal, no estacionaria, multivariada, no lineal y las simulaciones (Martinez-Vargas & Ramírez-Garcia, 2005).

### Variable regionalizada

“Una variable regionalizada  $z(x)$  es aquella que está geo-referenciadas en el tiempo o en el espacio presentando una estructura espacial de correlación” (Díaz, 2002).

### Función aleatoria

Se define como función aleatoria al conjunto de variables aleatorias en las que a cada punto  $x$  de nuestra área de muestreo le corresponde una variable aleatoria  $z(x)$ , siendo dependiente de esta manera (Díaz, 2002).

### Semivariograma



## Universidad de Cuenca

El variograma o semivariograma  $\gamma(h)$  es una herramienta esencial de la geostatística que expresa información de forma resumida. Es una función creciente que alcanza un valor límite llamado meseta que es igual a la varianza que tenga la muestra. La meseta se alcanza para un valor de  $h$  conocido como rango o alcance. El cual determina la zona de influencia en torno a un punto, más allá del cual la auto correlación es nula (García, 2004). Durante el análisis geostatístico primero se construye un semivariograma experimental y en base a este un semivariograma teórico (modelo ajustado) para obtener una función matemática precisa.

### Efecto Nugget

“Indica la discontinuidad del semivariograma para distancias que sean menores, que la menor distancia que se da entre los puntos muestrales” (Carrasco, 2014).

### Rango

“El rango es la distancia a partir de la cual dos observaciones son independientes. El rango se interpreta como la zona de influencia” (Abad, 2014).

### Modelos de semivariograma

Giraldo (2002) describe los modelos de semivariograma teórico más empleados:

- Modelo esférico

En este modelo el crecimiento es rápido al aproximarse al origen, pero entre mayor sea la distancia este va disminuyendo y a distancias superiores al rango los incrementos son nulos.

- Modelo gaussiano

En el modelo gaussiano la dependencia espacial disminuye cuando las distancias tienden al infinito. Tiene forma parabólica cerca al origen.

- Modelo exponencial

La dependencia espacial incrementa exponencialmente respecto a la distancia entre los puntos conocidos.



## Isotropía y Anisotropía

Cuando en los semivariogramas se considera que la variación del valor de la variable con el espacio es igual en todas las direcciones de éste (semivariograma omnidireccional), se dice que estas tienen un comportamiento isotrópico. Pero no siempre es así, y puede ser que la variación espacial sea diferente en las distintas direcciones del espacio teniendo un comportamiento anisotrópico (Gallardo, 2006).

## Estacionariedad

En el análisis geoestadístico la estacionariedad es un supuesto en el que se asume que el fenómeno es estacionario, es decir que el nivel promedio de la variable de estudio debe ser constante en todos los puntos del área. Ello confirma que las propiedades de la variable analizada no dependen de su posición en el espacio. Esta hipótesis se puede confirmar a través de la elaboración de un diagrama coordenada geográfica vs variable cuantitativa de estudio (Giraldo, 2002).

## Estudio geoestadístico

Según Vanderlinden, Giráldez, & Van Meirvenne (2003), los pasos a seguir para hacer un estudio geoestadístico (Ilustración 1) se pueden resumir de la manera siguiente:

1. Análisis exploratorio de datos: consiste en hacer un estudio de los datos para verificar su exactitud y representatividad. Se realiza un análisis descriptivo de los datos de manera que se cumplan las hipótesis de la geoestadística, esto es, verificar que los datos se ajusten a una distribución normal.
2. Análisis estructural o semivariográfico: cumplidas las condiciones anteriores se procede a calcular el semivariograma experimental y posteriormente se lo ajusta a un modelo teórico. Este semivariograma experimental se obtiene con los datos de muestra e indica el modelo de distribución espacial de la variable a partir de la correlación espacial entre los datos medidos en el área o región.
3. Interpolación o estimación espacial: consiste en estimar el valor que asume la variable de estudio para diferentes puntos dentro de la región, en base al semivariograma que define la correlación espacial existente entre los valores muestrales.

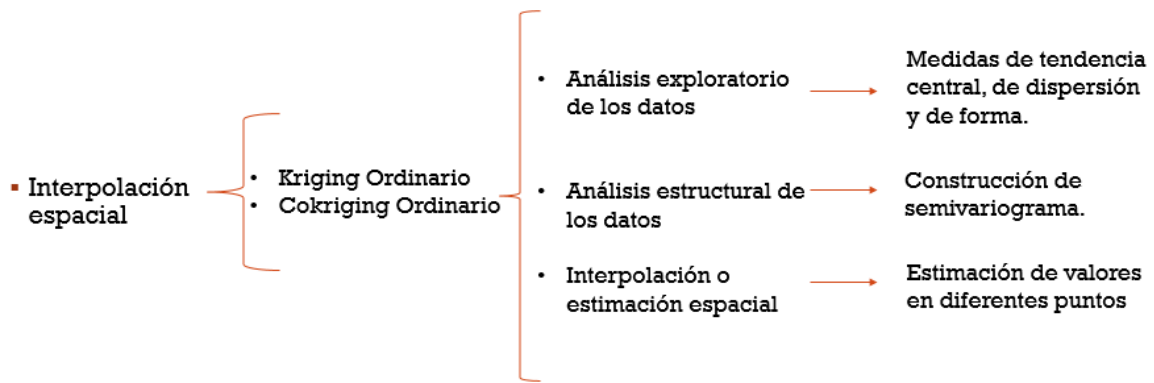


Ilustración 1: Estudio Geostatístico

### Kriging y Cokriging

El Kriging y Cokriging son métodos estadísticos que se utilizan para la interpolación (cartografía y de contorneado). Ambos métodos son formas generalizadas de modelos de regresión lineal univariante y multivariante. Son métodos de promediación ponderada lineal, similar a otros métodos de interpolación; sin embargo, sus pesos no dependen únicamente de la distancia, sino también de las características de los datos vecinos (Ahmadi & Sedghamiz, 2008). Básicamente, van a asignar un ponderado a cada valor de la variable de análisis para poder estimar un valor desconocido.

El Kriging es un método geostatístico de predicción, que designa al "mejor estimador lineal insesgado" de un punto o un bloque. El Kriging predice el valor desconocido basado en modelos estocásticos. El Kriging Ordinario (OK) es el método más utilizado, en parte porque es robusto con respecto a las desviaciones de los supuestos subyacentes, además es aplicable para varias situaciones con hipótesis fácilmente verificables para trabajar con este método solo se necesita los datos de estudio y las coordenadas geográficas a la que pertenece cada dato. OK se basa en la suposición de que todos los datos son aleatorios y que son espacialmente dependientes, en el cual la variable es estacionaria con media y varianza constante, que dependen únicamente de la distancia en que se encuentran los puntos muestreados (Oliver & Webster, 2015).

El Cokriging por otra parte es una estimación conjunta de variables aleatorias regionalizadas. El Kriging utiliza la correlación espacial para determinar los coeficientes en el estimador lineal, el Cokriging utiliza la correlación espacial y la correlación entre funciones aleatorias al mismo tiempo (Díaz, 2002).



## Universidad de Cuenca

La ventaja del método cokriging frente al kriging radica en el hecho de que cuando la variable auxiliar está ampliamente correlacionada con la variable de interés se puede obtener una disminución en la varianza de predicción, no obstante dicha variable tenga menor densidad de muestreo. En situaciones en las que la variable objetivo tiene costos altos de muestreo se recomienda la aplicación de esta metodología (Giraldo, 2002).

### Validación cruzada (leave one out)

La validación cruzada es una herramienta apropiada para determinar la eficiencia de métodos de interpolación: Consiste en dejar de lado un elemento de datos a la vez, y determinar qué tan bien este concepto puede estimarse a partir de los otros datos. Esta validación se utiliza a menudo para las pruebas de modelos Kriging; en este caso, cada valor es desconocido se predice a partir de un pequeño número de datos circundantes (Dubrule, 1983).

### 3. METODOLOGÍA

#### 3.1 Tipo de estudio

Se realizó un estudio de tipo descriptivo puesto que se abarca aspectos que van desde la descripción, visualización y resumen de datos originados a partir de los fenómenos de estudio, hasta la caracterización para establecer la estructura y comportamiento de los mismos.

#### 3.2 Área de estudio

La Provincia del Azuay pertenece a la región de la Sierra Ecuatoriana y está ubicada al sur oeste del Ecuador con las coordenadas  $2^{\circ}53'00''S$  y  $79^{\circ}00'00''O$ . El área de estudio se encuentra ubicada en la región noreste de la provincia (Ilustración 2), esta zona abarca 53 parroquias pertenecientes a las ciudades de: Cuenca, Sigsig, Gualaceo, Paute, El Pan, Guachapala y Sevilla de Oro (Anexo 1). En la zona se distinguen dos estaciones definidas: húmeda y seca, con temperaturas que van desde  $10^{\circ}C$  y  $28^{\circ}C$  en la zona andina (Censos-Ecuador, 2001).

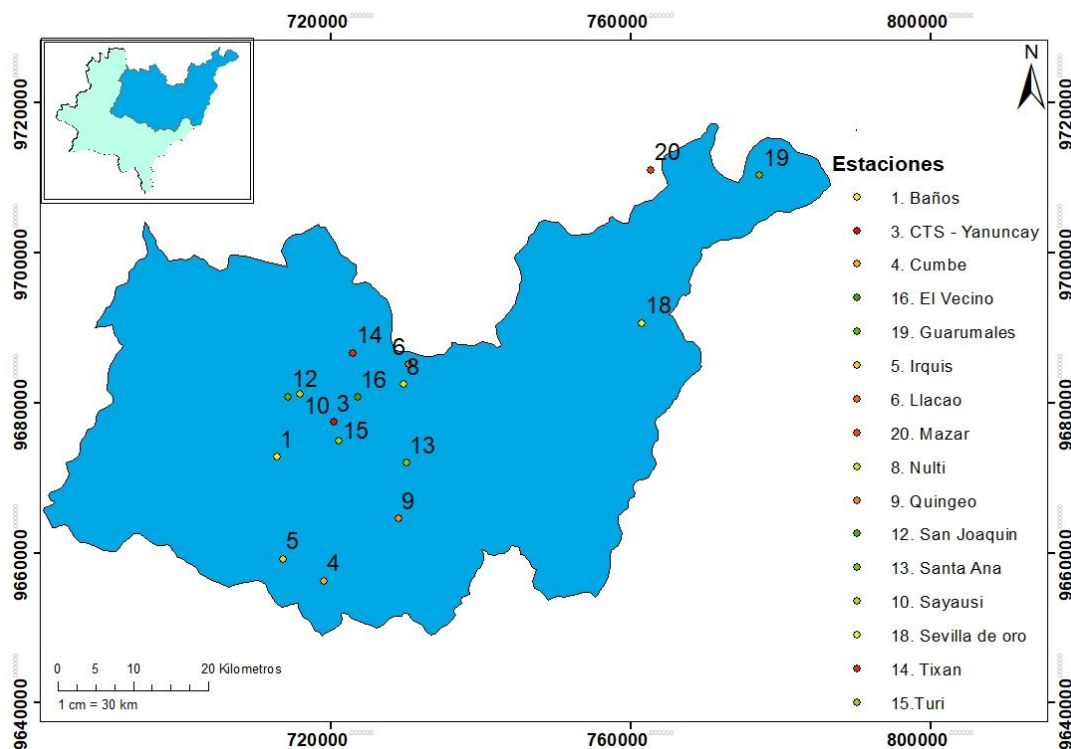


Ilustración 2: Mapa de ubicación del área de estudio. Datum: WGS84. Fuente: INEC-2011



### 3.3 Recopilación de información

#### 3.3.1 Variables de estudio

Tabla 1: Variables meteorológicas de estudio

<b>Parámetro</b>	<b>Unidad</b>
<b>Precipitación</b>	mm
<b>Temperatura</b>	°C

#### 3.3.2 Tipo de información: Secundaria

Los datos meteorológicos fueron recopilados desde los servidores web de la Universidad Politécnica Salesiana y de la Corporación Eléctrica del Ecuador CELEC EP, para ello se realizaron las solicitudes correspondientes para tener acceso a los mismos y proceder a la descarga de la información.

#### 3.3.3 Número de estaciones meteorológicas

Se contó con 20 estaciones localizadas especialmente en la zona noreste de la provincia. En la Tabla 2 se encuentran las estaciones con sus respectivos nombres, dependencia, coordenadas geográficas y la altitud (msnm) la cual varió entre 1896 hasta 3179 msnm.

Tabla 2: Estaciones Meteorológicas UPS-CELEC

<b>ID</b>	<b>Estación</b>	<b>Dependencia</b>	<b>Código</b>	<b>Coord X</b>	<b>Coord Y</b>	<b>Alt Z</b>
<b>1</b>	Baños	UPS	BAN	712899,00	9672817,00	3062
<b>2</b>	Chaucha	UPS	CHU	672859,00	9678690,00	1896
<b>3</b>	Yanuncay CTS	UPS	CTS	720504,00	9677509,00	2561
<b>4</b>	Cumbe	UPS	CUM	719190,00	9656242,00	3179
<b>5</b>	Irquis	UPS	IRQ	713645,00	9659192,00	2665
<b>6</b>	Llacao	UPS	LLA	730418,00	9685180,00	2542
<b>7</b>	Molleturo	UPS	MOL	676435,00	9692152,00	3524





<b>ID</b>	<b>Estación</b>	<b>Dependencia</b>	<b>Código</b>	<b>Coord X</b>	<b>Coord Y</b>	<b>Alt Z</b>
<b>8</b>	Nulti	UPS	NUL	729704,00	9682466,00	2623
<b>9</b>	Quingeo	UPS	QUI	729055,00	9664593,00	2895
<b>10</b>	Sayausí	UPS	SAY	715974,00	9681200,00	2727
<b>11</b>	Sinincay	UPS	SIN	722340,00	9685283,00	2696
<b>12</b>	San Joaquín	UPS	SJQ	714405,00	9680807,00	2764
<b>13</b>	Santa Ana	UPS	SAN	730085,00	9672006,00	2651
<b>14</b>	Tixan- Chiquintad	UPS	TXC	723017,00	9686678,00	2725
<b>15</b>	Turi	UPS	TUR	721103,00	9674971,00	2768
<b>16</b>	El Vecino	UPS	EVC	723584,00	9680788,00	2556
<b>17</b>	Cuenca	CELEC	CUE	729892,00	9684082,00	2492
<b>18</b>	Sevilla de oro	CELEC	SEV	761509,00	9690652,00	2497
<b>19</b>	Guarumales	CELEC	GUA	777205,00	9710390,00	2685
<b>20</b>	Mazar	CELEC	MAZ	762735,00	9711109,00	2643

\*Datum: WGS84

### 3.3.4 Período de análisis

Se pudo acceder a los datos históricos horarios correspondientes a los últimos 3 años (31 de julio del 2014 hasta el 31 de julio del 2017) de las dos fuentes de información. Se realizó un análisis básico de los datos previo a la investigación y se determinó que al existir gran cantidad de datos faltantes (más del 20% en varias estaciones) se limitaba el uso de estaciones para la posterior aplicación de los métodos geoestadísticos y por ello se decidió realizar el estudio durante el período 31 de Julio 2016 a 31 de Julio 2017.

### 3.3.5 Análisis estadístico

Los estadísticos utilizados fueron los siguientes:

- Error de predicción estandarizado medio (MSS).
- Raíz del error cuadrático medio (RMSE)
- Raíz cuadrática media estandarizada (RMSS)



- Media y la mediana
- Coeficiente de sesgo
- Desviación estándar
- Coeficiente de Curtosis
- Test Shapiro Wilk

### 3.3.6 Programa de análisis

Se utilizó el programa estadístico IBM SPSS para el análisis de datos faltantes y para la aplicación de la metodología de relleno de datos (Método de las proporcionalidades y Media Aritmética) se utilizó el programa NetBeans IDE 8.2 con lenguaje de programación JAVA.

Para la aplicación de los métodos Kriging-Cokriging, la validación de los mismos y la elaboración de las cartografías temáticas se utilizó el software ArcGis 10.2.2.

### 3.4 Métodos

#### 3.4.1 Relleno de información meteorológica

En la Ilustración 3, se esquematiza el procedimiento que se siguió para el relleno de información meteorológica faltante.

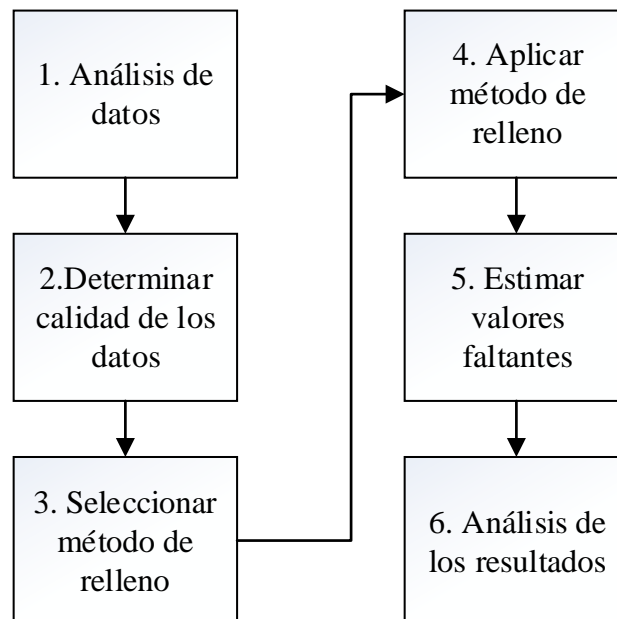


Ilustración 3: Proceso de relleno de datos meteorológicos. Elaboración: Autora



1. Análisis de datos

En la Tabla 3 se detalla las características de las estaciones con las que se realizó el estudio: número de estaciones, cantidad de datos disponibles y el porcentaje de datos vacíos por estación.

Tabla 3: Información precipitación y temperatura de estaciones meteorológicas

ID	Estación	Código	N Datos	Perdidos	
				N	%
1	Baños	BAN	8784	0	0
2	Chaucha	CHU	0	8784	100
3	Yanuncay CTS	CTS	8779	5	0.1
4	Cumbe	CUM	8409	375	4.3
5	Irquis	IRQ	8323	461	5.2
6	Llacao	LLA	7072	1712	19.5
7	Molleturo	MOL	2257	6527	74.3
8	Nulti	NUL	7828	956	10.9
9	Quingeo	QUI	8784	0	0
10	Sayausí	SAY	7853	931	10.6
11	Sinincay	SIN	5625	3159	36
12	San Joaquín	SJQ	8592	192	2.2
13	Santa Ana	SAN	8781	3	0.03
14	Tixan-Chiquintad	TXC	8780	4	0.04
15	Turi	TUR	8784	0	0
16	El Vecino	EVC	8784	0	0
17	Cuenca	CUE	2259	6525	74.28
18	Sevilla de oro	SEV	8784	0	0
19	Guarumales	GUA	8784	0	0
20	Mazar	MAZ	8784	0	0



## 2. Determinar la calidad de los datos

Identificar el patrón de datos faltantes (homogeneidad)

Para determinar la homogeneidad de la serie, hay que tomar en cuenta que las variaciones en los datos sean causadas solo por variaciones en el clima y no por factores externos. El tratamiento estadístico de datos requiere, como condición básica, que sean de la misma naturaleza, del mismo origen, obtenidos mediante observaciones que hayan seguido procedimientos semejantes (Avilés, 2016). De acuerdo a Teegavarapu (2012), identificar cuál es el patrón que tienen los datos perdidos es importante para determinar la mejor metodología a usar para su posterior relleno, y así no introducir mayor margen de error en los datos.

Existen tres mecanismos de pérdida de datos: datos perdidos completamente al azar (MCAR = missing completely at random), datos perdidos al azar (MAR= missing at random) y datos perdidos no ignorables o no debidos al azar (MNI=missing non-ignorable, o MNAR=missing not at random).

Para determinar el patrón de datos faltantes se utilizó el programa estadístico IBM SPSS.

## 3. Seleccionar el método de imputación

Existen un sinnúmero de métodos utilizados para rellenar series de precipitación. De acuerdo a la bibliografía revisada, se determinó que el método que mejores resultados ofrece es la regresión lineal múltiple para el caso de datos de precipitación y la media aritmética para datos de temperatura, sin embargo hay que tomar en cuenta la correlación que existe entre las estaciones auxiliares utilizadas para aplicación de los métodos.

Para conocer la relación existente entre los datos meteorológicos de una estación con los datos de otra, se procedió a eliminar las horas en las que existían datos faltantes en todas las estaciones, construyendo una base con datos conocidos para las dos variables. De estos valores se determinó las correlaciones (Pearson) entre las



estaciones. Para establecer la interpretación de la correlación se utilizó la tabla dada por Bisquerria citado en Abad (2014) (Ilustración 4).

$r=1$	correlación perfecta
$0.8 < r < 1$	correlación muy alta
$0.6 < r < 0.8$	correlación alta
$0.4 < r < 0.6$	correlación moderada
$0.2 < r < 0.4$	correlación baja
$0 < r < 0.2$	correlación muy baja
$r=0$	correlación nula

Ilustración 4: Tabla de interpretación de valores de correlación

Para la variable precipitación los coeficientes de correlación fueron de muy bajos a moderados con valores desde 0.13 a 0.57, por debajo del recomendado (0.70) para la aplicación del método. Por ello se determinó que el método de regresión lineal múltiple no era el adecuado para el relleno de los datos de precipitación. Se decidió utilizar un método de relleno que no utilice estaciones auxiliares para la estimación de los valores faltantes, el método de las proporcionalidades. Este método ofrece resultados satisfactorios cuando no se tiene otra estación con la cual realizar una comparación (Monte Alegre, 1990).

#### Método de las proporcionalidades

Es un método de relleno de datos considerado como confiable pues se trabaja con los valores medios de los datos registrados durante el periodo analizado (1).

$$x_i = p_{mi} \left( \frac{\sum x_{ni}}{p_a} \right) \quad (1)$$

En la cual  $x$  es la variable que representa el dato horario faltante,  $p_{mi}$  es la precipitación media horaria del período para la hora  $i$ ,  $x_{ni}$  son los datos horarios existentes del día en el que se encuentra el dato horario faltante y  $p_a$  es la precipitación media diaria total del período de estudio para los días con datos completos (Loja, 2011).

#### Método de la media aritmética (Temperatura)



Consiste en sacar el valor medio de los datos de las estaciones cercanas o que mayor correlación presentan con la estación de llenado. En este método se toma en cuenta las estaciones con valores a esa hora (Pizarro T, Ramirez B, V, & Pablo, 2003). La ecuación es la siguiente:

$$Pm = \sum \frac{Pi}{n} \quad (2)$$

En la cual Pm es el valor medio del variable de estudio a una hora determinada, Pi son los valores conocidos de la variable a la misma hora y n es el número de estaciones en análisis.

#### 4. Estimar los datos

Para aplicar del método de las proporcionalidades se procedió a eliminar los días que presentaban horas con datos vacíos, de esta manera se construyó una base de datos completa. De esta base de datos se eliminó el 20% de los datos horarios de precipitación de manera aleatoria, ello para posteriormente poder realizar las comparaciones estadísticas y por último se procedió a aplicar el método (Ecuación 1). En el caso del método media aritmética, solo se procedió a establecer las correlaciones, se formaron los grupos de 5 estaciones (1 estación de relleno y 4 estaciones auxiliares) y se aplicó el método (Ecuación 2). El llenado de los datos se realizó mediante el programa NetBeans.

Una vez estimados los valores faltantes se establecieron las comparaciones de los valores reales con los estimados, estos se compararon mediante (Campozano, Sánchez, Avilés, & Samaniego, 2015):

- Error Medio (ME)

$$ME = \frac{\sum_{i=1}^n Ii - Obsi}{n} \quad (3)$$

- Error Absoluto Medio (MAE)

$$MAE = \frac{\sum_{i=1}^n |Ii - Obsi|}{n} \quad (4)$$

- Raíz del error cuadrático medio (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Ii - Obsi)^2}{n}} \quad (5)$$



En donde  $I_i$  es el valor estimado,  $Obs_i$  es el valor observado o real y  $n$  es el tamaño de la muestra.

### 3.4.2 Aplicación de Kriging y Cokriging

- Datos

Con la base de datos tanto de precipitación como de temperatura completas, se procedió a armar un archivo Excel, en las que se colocaron los datos promedio hora pertenecientes a cada una de las estaciones de estudio (Tabla 4).

Tabla 4 Información estaciones: Kriging-Cokriging

<b>ID</b>	<b>Estación</b>	<b>Dependencia</b>	<b>CoordX</b>	<b>CoordY</b>	<b>AltZ</b>	<b>TempAvg</b>	<b>PrecipAvg</b>
1	Baños	UPS	712899	9672817	3062	10.89	0.06
3	CTS - Yanuncay	UPS	720504	9677509	2561	14.71	0.05
4	Cumbe	UPS	719190	9656242	3179	10.23	0.07
5	Irquis - Victoria del Portete	UPS	713645	9659192	2665	12.30	0.04
6	Llacao	UPS	730418	9685180	2542	14.22	0.05
8	Nulti	UPS	729704	9682466	2623	13.99	0.06
9	Quingeo	UPS	729055	9664593	2895	12.03	0.06
10	Sayausí	UPS	715974	9681200	2727	13.37	0.04
12	San Joaquín	UPS	714405	9680807	2764	12.82	0.06
13	Santa Ana	UPS	730085	9672006	2651	13.19	0.06
14	Tixan - Chiquintad	UPS	723017	9686678	2725	13.44	0.07
15	Turi	UPS	721103	9674971	2768	12.93	0.05
16	El Vecino	UPS	723584	9680788	2556	15.13	0.06
18	Sevilla de oro	CELEC	761509	9690652	2497	16.66	0.09
19	Guarumales	CELEC	777205	9710390	2685	19.52	0.20
20	Mazar	CELEC	762735	9711109	2643	18.76	0.09

\*Avg= promedio. Datum: WGS84



Cada estación contó con sus coordenadas en formato WGS84, posteriormente el archivo en formato .xls se convirtió en shapefile, para poder llevar a cabo los métodos en la plataforma ARCGIS.

- Procedimiento computacional

El software ARCGIS 10.2.2 posee la herramienta llamada Geostatistical Analyst, la cual permite ejecutar el procedimiento geoestadístico, el proceso se realiza para varios puntos creando así un mapa de superficie. Además permite efectuar la validación cruzada conjuntamente con el método geoestadístico, lo que ayuda a seleccionar el mejor modelo basándose en dicha validación (Guerra, Gómez, González, & Zambrano, 2006b).

- Análisis exploratorio de los datos

El objetivo de este análisis fue determinar la cantidad, la calidad y la ubicación de los datos con los que se está trabajando, mediante herramientas estadísticas simples como:

- La media y mediana deben ser similares, generalmente se admite una unidad de diferencia entre ellas.
- El coeficiente de sesgo debe ser  $0 < |CS| < 0.5$ .
- Desviación estándar próxima a 0.
- Coeficiente de Curtosis igual a 0.

Una vez determinados los parámetros estadísticos básicos se procedió a verificar la normalidad de los datos, mediante la prueba estadística de Shapiro Wilk para muestras menores a 30 individuos.

En el caso de no cumplir con el supuesto de normalidad se debe realizar una transformación de datos de tipo raíz cuadrada, logarítmica,  $1/x$ , etc. ("Análisis geoestadístico con ArcGIS parte 2. Análisis exploratorio de los datos," 2011).

### **Análisis estructural o semivariográfico**

Para Daniele (2009), este análisis se basa en la creación del semivariograma experimental y posteriormente el teórico, lo cual permite determinar las características de variabilidad y correlación espacial del fenómeno estudiado. Para la construcción del semivariograma empírico se utiliza la siguiente ecuación:





$$2\gamma(s_i, s_j + h) = \frac{1}{n} \sum (Z(s_i) - Z(s_i + h))^2 \quad (6)$$

$$\gamma(s_i, s_j) = \frac{1}{2} \text{var}[Z(s_i) - Z(s_j)] \quad (7)$$

En las ecuaciones (6-7)  $s_i$  es la ubicación de la estación  $i$ ,  $s_j$  es la ubicación para una estación "j" situada a una distancia de "h" de "i", h es la distancia, n son el número de pares de estaciones ( $s_i, s_j$ ) y Z es la función aleatoria distribuida espacialmente (precipitación-temperatura medidas).

A continuación, el semivariograma empírico se ajusta a un modelo con las ecuaciones 8, 9 y 10:

$$\gamma \text{ esférico} = \left\{ \theta_0 + \left( 1 - 5 \left( \frac{h}{\theta_2} \right) - 0.5 \left( \frac{h}{\theta_2} \right)^3 \right) \right\} \quad \begin{array}{l} ; h = 0 \\ ; 0 < h < \theta_2 \\ ; h > \theta_2 \end{array} \quad (8)$$

$$\gamma \text{ exponencial} = \left\{ \theta_0 + \theta_1 \left( 1 - \exp \left( -3 \left( \frac{h}{\theta_2} \right) \right) \right) \right\} \quad \begin{array}{l} ; h = 0 \\ ; h \neq 0 \end{array} \quad (9)$$

$$\gamma \text{ gaussiano} = \left\{ \theta_0 + \theta_1 \left( 1 - \exp \left( -3 \left( \frac{h}{\theta_2} \right)^2 \right) \right) \right\} \quad \begin{array}{l} ; h = 0 \\ ; h \neq 0 \end{array} \quad (10)$$

En donde  $\theta_0$ , h y  $\theta_2$  se denominan nugget, distancia y rango.

El Cokriging es una variación del Kriging en la cual se utiliza una variable auxiliar que guarda relación con la variable de estudio.

La variable auxiliar utilizada para el cokriging fue la elevación (msnm) de las estaciones, para ello primero se determinó cual es la correlación existente entre las dos variables de estudio y la altura a la que se encuentra ubicada la estación meteorológica.

- Proceso

El análisis estructural inició con la determinación de la autocorrelación espacial y direccional de los valores de las variables. Para ello se analizó la variabilidad espacial, distribución o dispersión que tienen los datos. Se construyó primero un semivariograma empírico, es decir el experimental.

Semivariograma teórico



El objetivo de este semivariograma es determinar cual es el mejor modelo geoestadístico que encaja con los datos de análisis. Para iniciar se seleccionó la herramienta Geostatistical Analyst de ArcMap, la cual permite aplicar de los métodos Kriging o Cokrigin, para cada uno de ellos se especificó el tipo de transformación y la tendencia de los datos.

### Interpolación o estimación espacial

Posteriormente se estableció la relación hipotética espacial entre estaciones: entre más cercanas estén las estaciones tienen valores más similares que las que están apartadas. Con la ecuación 11, se obtiene el peso de los datos espaciales para su uso en la ecuación 12, la cual determina el valor del punto interpolado.

$$\left\{ \begin{array}{l} \sum_{i=1}^n \lambda_i \gamma_{ij} - \mu = \gamma_{i0} \\ \sum_{i=1}^{ns} \lambda_i = 1 \end{array} \right\} \quad j = 1, \dots, n \quad (11)$$

$$Z(g) = \sum_i^n \lambda_i Z(s)_i \quad (12)$$

Este procedimiento se repitió para todos los puntos con los cuales se trabajó.

#### 3.4.3 Validación del modelo geoestadístico

Obtenida la estimación de los datos para cualquier punto dentro de la región se realizó la validación. Para ello se utilizó el método de validación cruzada, el cual comprueba la validez del modelo geoestadístico y permite seleccionar el que mejor se ajuste a los datos. El método consiste en quitar un dato y predecirlo mediante el modelo utilizado con el resto de la muestra, luego se compara el valor predicho de ese punto con el valor observado (Haddad, Rahman, A Zaman, & Shrestha, 2013).

Los criterios establecidos para la elección del mejor modelo son los expuestos por Adhikary, Yilmaz, & Muttil, (2015) los cuales son:



- El error estandarizado medio (MSS): Es el promedio de los errores estandarizados. Este valor debe estar cerca de 0.

$$MSS = \frac{\sum_{i=1}^n (I_i - Obs_i / \sigma_i)}{n} \quad (13)$$

- La raíz del error cuadrático medio (RMSE): La variabilidad de los datos estimados se determinó comparando los valores RMSE con los valores tanto KSE (Kriging Estándar Error) y CoSE (Cokriging Estándar Error), si los valores eran cercanos entre sí, la predicción se realizó de manera correcta.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (I_i - Obs_i)^2}{n}} \quad (14)$$

- Raíz Cuadrada de la Media del Error Estandarizado al Cuadrado (RMSS): la cual debe ser cercana a uno para que las varianzas sean consistentes.

$$RMSS = \sqrt{\frac{\sum_{i=1}^n (I_i - Obs_i / \sigma_i)^2}{n}} \quad (15)$$

En las cuales  $I_i$  es el valor predicho,  $z_i$  es el valor real o medido,  $\sigma_i$  es la desviación estándar y  $n$  es el tamaño de la muestra.

### 3.4.4 Publicación de resultados

Los resultados obtenidos en este trabajo de titulación (método de interpolación con su respectivo modelo geoestadístico) se utilizarán para ser incorporados en una plataforma web (Dirección url: <http://201.159.222.53:8080/cedia/mapa.jsp>). El Geoportal es obra de un grupo de investigación conformado por docentes de varias universidades del país (Escuela Superior Politécnica de Chimborazo, Universidad Nacional de Chimborazo y Universidad de Cuenca) y REDCEDIA (Red Nacional de Investigación y Educación del Ecuador). Es de código abierto, por lo que permite que se vaya mejorando constantemente.

Con el geoportal se pretende que cualquier persona que necesite información de temperatura y precipitación pueda acceder a ella en cualquier fecha y hora. Lo que se quiere es que esta información sea de calidad, por ello si existieran vacíos de información por cualquier motivo, al implementar los métodos aquí analizados (método de las proporcionalidades y la media aritmética) se podrá obtener una estimación



estadística del dato faltante a partir de datos históricos, con un porcentaje de error mínimo al momento de interpolar la información. Una vez que la base de datos este completa se procede a realizar la interpolación geoestadística y con ello se obtiene la información meteorológica de cualquier lugar ubicado en la zona noreste de la provincia del Azuay.

El proceso de llenado de datos y de interpolación geoestadística se realizará de manera ajena al usuario que desee la información, de modo que el conocimiento técnico que este debe tener es mínimo. Se espera que esta información sea útil para profesionales, empresas o estudiantes que requieran de esta información (Patiño & Lupercio, 2017).

#### 4. RESULTADOS Y DISCUSIÓN

##### 4.1 Relleno de información meteorológica

###### 4.1.1 Análisis de datos

De acuerdo a lo establecido por Urrutia, Palomino, & Salazar (2010) el porcentaje de datos faltantes no debe sobrepasar el 20% del total de datos, por lo que para la presente investigación se omitió la información de cuatro estaciones: tres pertenecientes a la UPS (Chaucha, Molleturo y Sinincay) y una a CELEC (Cuenca).

Tabla 5: Estaciones meteorológicas (% de datos faltantes)

ID	Estación	Código	N Datos	Perdidos	
				N	%
1	Baños	BAN	8784	0	0
2	Chaucha	CHU	0	8784	100
3	Yanuncay CTS	CTS	8779	5	0.1
4	Cumbe	CUM	8409	375	4.3
5	Irquis	IRQ	8323	461	5.2
6	Llacao	LLA	7072	1712	19.5
7	Molleturo	MOL	2257	6527	74.3
8	Nulti	NUL	7828	956	10.9
9	Quingeo	QUI	8784	0	0
10	Sayausí	SAY	7853	931	10.6
11	Sinincay	SIN	5625	3159	36

12	San Joaquín	SJQ	8592	192	2.2
13	Santa Ana	SAN	8781	3	0.03
14	Tixan-Chiquintad	TXC	8780	4	0.04
15	Turi	TUR	8784	0	0
16	El Vecino	EVC	8784	0	0
17	Cuenca	CUE	2259	6525	74.28
18	Sevilla de oro	SEV	8784	0	0
19	Guarumales	GUA	8784	0	0
20	Mazar	MAZ	8784	0	0

Se trabajó por lo tanto, con 16 estaciones de las cuales 9 de ellas presentaron datos faltantes (Tabla 5). En la Ilustración 6 se presenta un resumen global de los datos perdidos, en total se trabajó con 135,905 valores teniendo un 56,25% de información.

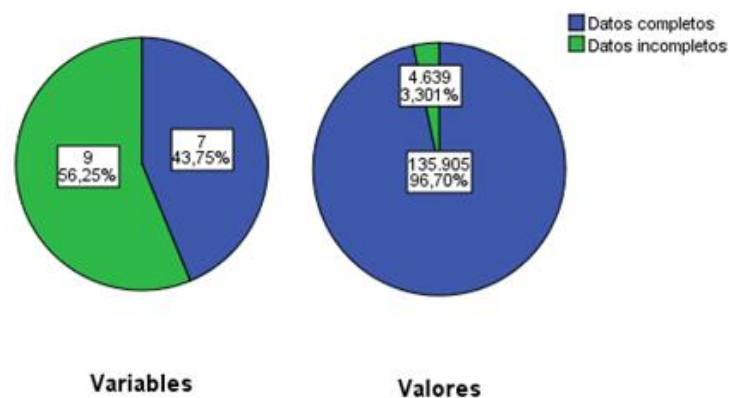


Ilustración 5 Resumen global datos meteorológicos perdidos

#### 4.1.2 Determinación de la calidad de los datos

Para disminuir el efecto de los valores extremos u erróneos en los datos, se consideró como un valor atípico o outlier todo valor que superó las 3 desviaciones típicas del valor medio normal para la hora correspondiente para la variable temperatura y 5 desviaciones típicas para el caso de precipitación, pues esta variable tiene mayor variabilidad que la temperatura (Lago, 2009). En la Tabla 6 se tiene un resumen de los valores atípicos eliminados para precipitación. En el caso de temperatura, no se presentó ningún outlier.



Tabla 6: Estaciones con outliers para variable precipitación

ID	Estación	Outliers
3	Yanuncay CTS	62
4	Cumbe	61
5	Irquis	58
6	Llacao	43
8	Nulti	56
10	Sayausí	61
12	San Joaquín	74
13	Santa Ana	65
14	Tixan-Chiquintad	49

#### 4.1.3 Identificación del patrón de datos faltantes

Para determinar el patrón de datos faltantes se utilizó el programa estadístico IBM SPSS, con la herramienta VMA (Missing Value Analysis), subcomando EC (expectation-maximization). Se aplicó la prueba de hipótesis de Little, que propone un estadístico de prueba que sigue una distribución  $X^2$  (chi-cuadrada) con  $f$  grados de libertad, en donde la hipótesis nula ( $H_0$ ) establece que los datos omitidos siguen un patrón MCAR (Garzon, 2015).

Del análisis se obtuvieron los valores de medias marginales estimadas representadas en la Tabla 7 y Tabla 8, en la que se puede observar que  $p=1$ , por tanto al ser mayor a 0.05, no se rechazó la hipótesis nula, comprobando lo establecido por Teegavarapu (2012), que el patrón que tienen los datos faltantes es MCAR, es decir la propensión a que falte un punto de datos de precipitación o temperatura fue completamente aleatoria y no dependió de los valores de  $X$  e  $Y$  (ni de los propios ni de las otras estaciones). Esto permitió utilizar las estrategias de imputación disminuyendo posibles sesgos (Cañizares, Barroso, & Alfonso, 2004).

Tabla 7: Tabla de Medias marginales- Precipitación

#### Medias marginales estimadas<sup>a</sup>

CTS	CUM	IRQ	LLA	NUL	SAY	SJQ	SAN	TXC
0,1120	0,1186	0,0925	0,1317	0,1162	0,0995	0,1322	0,1160	0,1252

a. Prueba MCAR de Little: Chi-cuadrado = 102,901, DF = 156, Sig. = 1,000

Tabla 8: Tabla de Medias marginales- Temperatura



**Medias marginales estimadas<sup>a</sup>**

CTS	CUM	IRQ	LLA	NUL	SAY	SJQ	SAN	TXC
14.9266	10,1847	12,5830	14,7164	14,2560	13,5160	13,0234	13,3900	13,5803

a. Prueba MCAR de Little: Chi-cuadrado = 95,574, DF = 156, Sig. = 1,000

4.1.3.1 Análisis de correlación

Para conocer la correlación existente entre las estaciones meteorológicas y poder formar los grupos de estaciones para aplicar el método de media aritmética, se utilizó el coeficiente correlación de Pearson (16), el cual mide el grado de variación entre las variables relacionadas linealmente.

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (13)$$

En la cual  $\sigma_{xy}$  es la covarianza (x,y),  $\sigma_x$  es la desviación típica de x y  $\sigma_y$  es la desviación típica de y, siendo x y y los valores de temperatura y precipitación para cada estación.

Se tuvieron correlaciones muy altas desde 0.87 a 0.98 (Anexo 2), es decir existió alta correlación entre la variable dependiente (estación con vacíos) y las variables independientes (estaciones auxiliares).

4.1.4 Estimación de valores faltantes

Para el relleno de datos de precipitación, se construyó una tabla (formato .xls) de las 24 horas del día y los 366 días de los cuales se disponían los datos, ello para poder eliminar los días que tenían al menos un dato horario faltante. Una vez que se tuvo una base de datos completos se procedió a eliminar diferentes horas al azar del período de estudio y se aplicó método para estimar los datos faltantes, a continuación se describe un ejemplo de la aplicación de la fórmula 1:

El día 8 de noviembre del 2016 presenta un dato faltante correspondiente a las 3pm, para la estimación de ese valor se procedió a realizar un promedio de los datos de precipitación existentes a las 3pm de los días que se tiene información ( $p_{mi}$ ), se realizó una sumatoria de los valores existentes ese día ( $x_{ni}$ ) y por último se calculó el



promedio hora día de los días que poseían todos los datos horarios de precipitación ( $p_a$ ), una vez obtenidos estos valores, se aplicó la fórmula correspondiente al método.

Se siguió el mismo proceso para todas las horas que presentaron datos faltantes. De esta manera se obtuvieron los datos faltantes de cada una de las 9 estaciones de relleno

Para el relleno de datos de temperatura se siguió el siguiente procedimiento: se estableció grupos de estaciones que tenían correlaciones altas y que no tenían datos faltantes en las mismas horas que la estación de relleno, para luego aplicar el método de la media aritmética y con ello llenar las horas vacías. En este caso, las estaciones de Cumbe, Irquis, Llaoco y Nulti, no fueron utilizadas como estaciones auxiliares debido a que son las que presentan mayor cantidad de datos faltantes que coinciden con los datos faltantes de las estaciones de relleno. En la Tabla 9, se tiene los grupos con los cuales se trabajó.

Tabla 9: Grupos de estaciones para el método Media Aritmética

	<b>CTS</b>	<b>CUM</b>	<b>IRQ</b>	<b>LLA</b>	<b>NUL</b>	<b>SAY</b>	<b>SJQ</b>	<b>SAN</b>	<b>TXC</b>
<b>BAN</b>		X		X					
<b>CTS</b>							X	X	X
<b>CUM</b>									
<b>IRQ</b>									
<b>LLA</b>									
<b>NUL</b>									
<b>QUI</b>		X	X	X	X			X	
<b>SAY</b>							X	X	
<b>SJQ</b>	X	X	X			X			X
<b>SAN</b>			X			X			
<b>TXC</b>	X				X				
<b>TUR</b>	X	X	X	X	X	X	X		X
<b>EVC</b>	X			X	X	X	X	X	X

Una vez encontrada la información faltante, se comparó los valores reales con los valores estimados mediante el ME, MEA y RMSE (Tabla 10).

Tabla 10: Estadística de relleno: ME, MEA, RMSE





ID	Estación	Precipitación			Temperatura		
		ME	MAE	RMSE	ME	MAE	RMSE
3	Yanuncay CTS	0.002	0.002	0.134	-0.971	0.971	0.063
4	Cumbe	0.0002	0.0002	0.195	1.908	1.908	0.087
5	Irquis	0.0005	0.0005	0.109	1.068	1.068	0.010
6	Llacao	-0.001	0.001	0.213	-1.783	1.783	0.065
8	Nulti	0.007	0.007	0.182	-0.700	0.700	0.028
10	Sayausí	-0.003	0.003	0.129	0.100	0.100	0.011
12	San Joaquín	0.000	0.000	0.179	1.170	1.170	0.077
13	Santa Ana	0.004	0.004	0.146	0.770	0.770	0.063
14	Tixan-Chiquintad	0.006	0.006	0.203	0.347	0.347	5E-17

De acuerdo a lo establecido por Urrutia, Palomino, & Salazar (2010) un método de relleno es excelente cuando el  $MAE=ME=0$ , siendo bueno cuando se aproxima a ese valor. Para Veerasamy et al. (2011), el valor de RMSE debe ser menor a 0.3 para considerar la estimación adecuada. Por otro lado, Draxler & Chai, 2014 manifiestan que el RMSE suele ser mejor revelando el rendimiento de los métodos de relleno de datos. Para temperatura los valores de MAE van desde 0.10 a 1.90, teniendo los valores más altos para las estaciones de Llacao y Cumbe. Sin embargo los valores de RMSE son menores a 0.3 para todas las estaciones. Para la variable precipitación los valores de MAE estuvieron entre 0.0002 y 0.007 para Cumbe y Nulti y los valores de RMSE entre 0.10 a 0.21, menores a 0.3.

En el Anexo 3 se puede apreciar gráficos de los promedio de cada hora del día de los datos rellenos y los datos vacíos tanto para precipitación como para temperatura.

Por lo tanto los dos métodos de relleno cumplieron con el objetivo de relleno estadísticamente. Corroborando que el método de media aritmética resultó satisfactorio para el relleno de datos de temperatura con pocos datos faltantes (Sattari, Rezazadeh-Joudi, & Kusiak, 2016) y que el método de proporcionalidades es adecuado cuando no se tiene correlación con estaciones auxiliares y por tanto se debe utilizar la información de la misma estación de relleno (Lituma & Andreina, 2011).

#### 4.2 Aplicación de Kriging y Cokriging

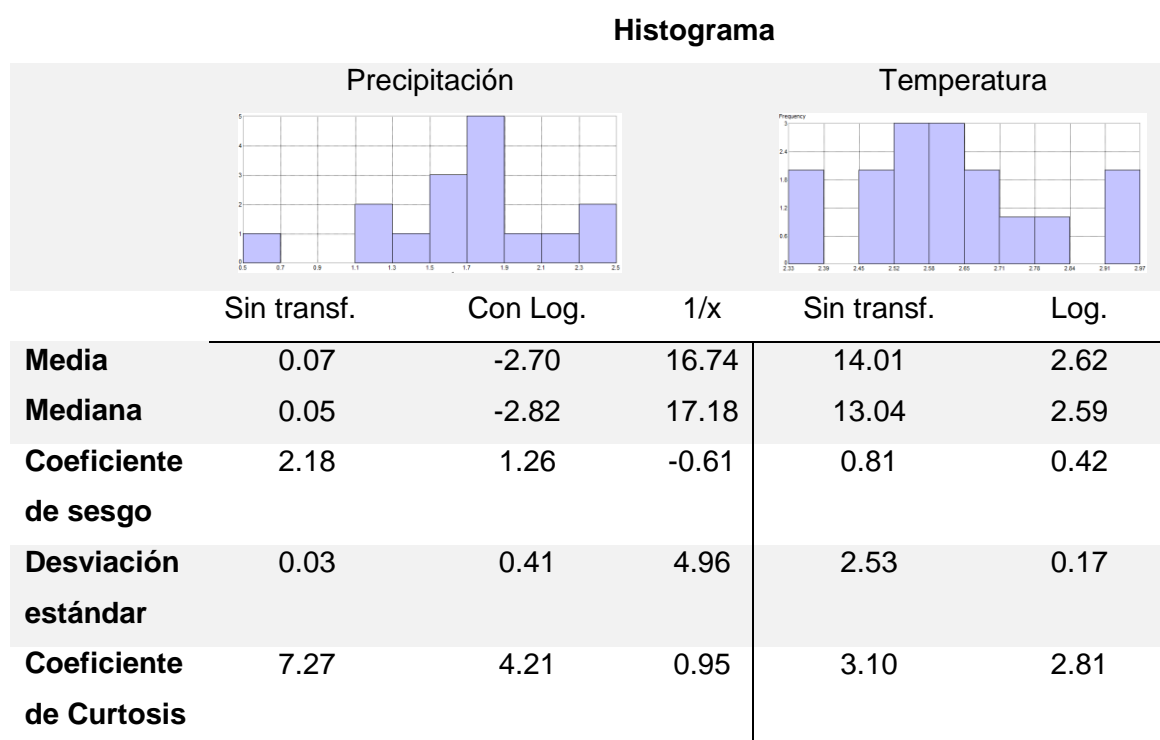


4.2.1 Análisis exploratorio de los datos

Para conocer la distribución de los datos e identificar la tendencia de los datos, se utilizó el histograma y el gráfico de tendencia.

En la Tabla 11 se tienen los parámetros estadísticos para verificar la normalidad de las variables, tanto para los datos sin transformación como transformados logaritmicamente y con transformación 1/x. Para el caso de precipitación se cumplió el primer criterio, media y mediana son casi iguales con una diferencia 0.02 y 0.09; el coeficiente de sesgo fue mayor al recomendado tanto su valor normal como ya transformado, ninguno de los dos casos cumple; la desviación estandar en los dos casos fue mínima y el coeficiente de Curtosis disminuye con la transformación logarítmica. De la misma manera para el caso de la temperatura se cumplió la mayor cantidad de parámetros con la transformación logarítmica.

Tabla 11: Análisis exploratorio: Histograma



Se aplicó el test estadístico Shapiro-Wilk para muestras menores a 30 individuos, para verificar la normalidad de los datos transformados (Tabla 12). El test plantea como hipótesis nula que los datos provienen de una distribución normal, se acepta cuando el nivel de significancia es mayor a 0.05. El nivel de significancia para la variable



temperatura fue de 0.60, es decir mayor a 0.05 por lo que se comprueba la normalidad de los datos. Para el caso de precipitación no se cumple el criterio con la transformación logarítmica por lo que se realizó una transformación de tipo 1/x cumpliendo de esta manera con la hipótesis de normalidad (Tabla13).

Tabla 12: Prueba de normalidad-transformación logarítmica. a)Precipitación, b)Temperatura

Pruebas de normalidad			
a)	Shapiro-Wilk		
	Estadístico	gl	Sig.
LogP	.794	16	.002
b)	Shapiro-Wilk		
	Estadístico	gl	Sig.
LogT	.960	16	.670

Tabla 13: Prueba de normalidad- transformación 1/x. Precipitación

Pruebas de normalidad			
	Shapiro-Wilk		
	Estadístico	gl	Sig.
1/x	.959	16	.644

Por lo tanto, se utilizó transformación logarítmica para temperatura y la transformación 1/x para precipitación.

#### 4.2.2 Análisis estructural de los datos

El semivariograma permitió analizar la autocorrelación que tuvieron los puntos de muestreo bajo la premisa que las estaciones más cercanas tendrán valores mucho más similares que los valores de las estaciones más lejanas. Cada punto rojo del semivariograma representa un par de puntos correlacionados (Ilustración 6). Mientras estos puntos se encuentren más cerca del origen mayor autocorrelación se tendrá.

Las estaciones más correlacionadas para precipitación fueron: San Joaquín-Sayausí, Yanuncay-Turi, Yanuncay-El Vecino, Yanuncay-Sayausí, Nulti-El Vecino, Irquis-Cumbe y Santa Ana-Quingeo. Mientras que los pares San Joaquín-Sayausí, Yanuncay-Turi, Llacao-Nulti y Yanuncay-El Vecino para temperatura. Por lo que los valores predichos para las áreas circundantes a estas zonas tendrán menor margen de error (Oliver & Webster, 2015).

Mientras que las autocorrelaciones más bajas fueron para los pares conformados por Estación UPS-Estación CELEC, principalmente Baños-Mazar para temperatura y Santa Ana-Guarumales para precipitación.

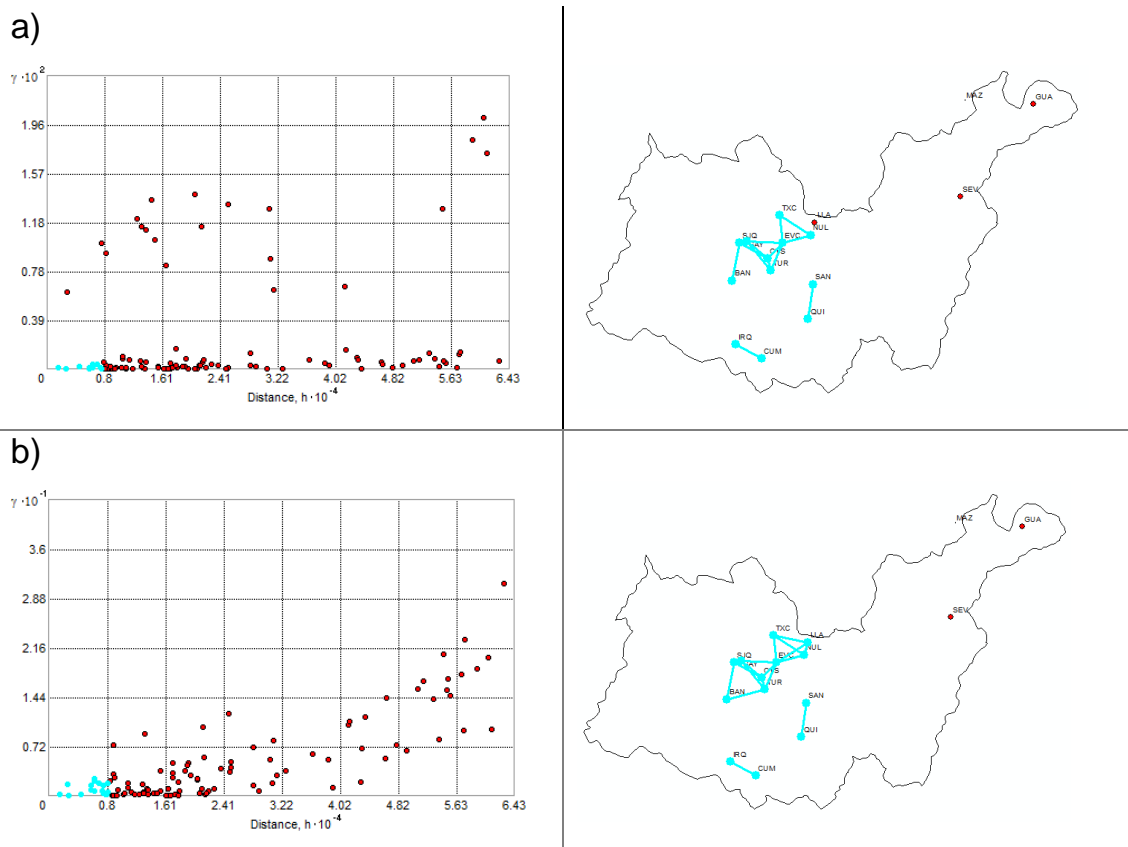


Ilustración 6: Semivariograma empírico: a) Precipitación b) Temperatura

### Anisotropía

Para verificar la existencia de anisotropía se realizó un análisis gráfico de los datos. El semivariograma experimental es calculado en diferentes direcciones y si este



presenta distintos comportamientos con la variación de la distancia, existe anisotropía (Zucarelli, Paris, & Macor, 2014). Generalmente para determinar este comportamiento se construyen diagramas de rosa, sin embargo el software utilizado en esta investigación crea una superficie a partir de los semivariogramas experimentales, por lo que no fue necesario la elaboración de los diagramas. Si el gráfico generado tiene forma circular no existe anisotropía direccional debido a que existe cierta organización en la información, si presenta forma de elipse se confirma la existencia de anisotropía, no hay organización en los datos, es decir, la información varía con la dirección, el semivariograma se comporta diferente cuando se calcula en diferentes direcciones (Dubrule, 1983).

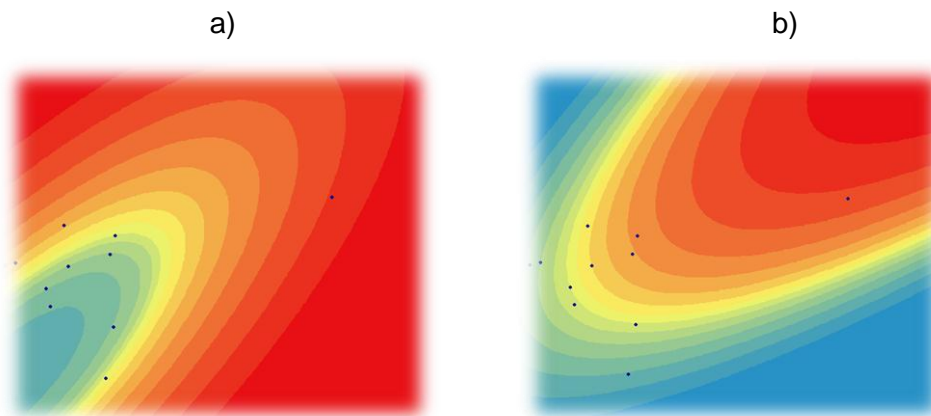


Ilustración 7: Anisotropía a) Precipitación b) Temperatura

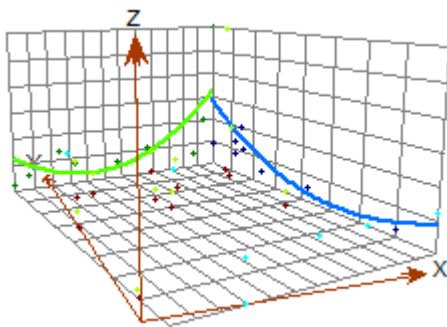
En la Ilustración 7 se tiene las superficies creadas a partir de los variogramas experimentales o empíricos. Se determinó la anisotropía de los datos para las dos variables estudiadas, es decir que la variabilidad espacial cambia con la dirección. En estos casos se debe utilizar variogramas que no dependan solo de la distancia, sino también de la dirección (García, 2004).

#### Estacionariedad

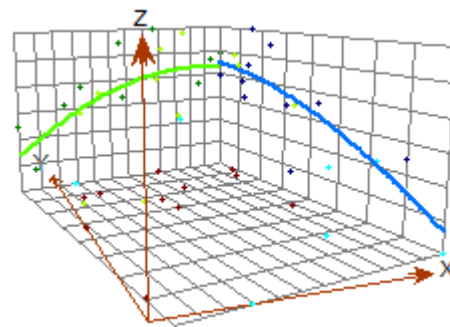
Otro supuesto del método implica que el proceso espacial debe asumirse como de estacionariedad intrínseca, es decir que los datos de entrada no cambiarán con el tiempo o espacio y por lo tanto la media y la varianza son constantes. Este supuesto se convierte en desventaja para la aplicación de los métodos geoestadísticos, debido a que las variables meteorológicas rara vez presenta esta estacionariedad (Abad Encalada, 2014).

Para resolver esta situación López (2006), a planteado que se debe determinar la tendencia que siguen los datos mediante un análisis del comportamiento de las variables distribuidas en el espacio. Para ello se gráfica cada punto en el espacio (x,y,z) y se construye una línea de tendencia. El análisis se realiza a través de modelos de regresión de primero, segundo y tercer orden que relacionan una variable de estudio con sus coordenadas geográficas de tal manera que los resultados representan las tendencias de la variable en alguna dirección (Torres, 2017). En la Ilustración 8 se puede observar que existe una tendencia noreste y noroeste en las dos variables (temperatura y precipitación) y que la línea de tendencia tiene una proyeccion con forma de U (upside-down), es decir existe una distribución de segundo orden polinomial. Para eliminar esta tendencia se realizó una diferenciación simple de segundo orden, que es una resta consecutiva de los valores de la variable en la serie, donde cada valor se le resta el anterior, repitiendo d veces el proceso (López, 2006).

Precipitación



Temperatura



*Ilustración 8: Análisis exploratorio\_ tendencia*

Una vez eliminada la tendencia se comprobó el supuesto de estacionariedad de los datos mediante el test Dickey-Fuller, el cual permite determinar la presencia de raíces unitarias (característica de los procesos que evolucionan en el tiempo). En el test la hipótesis nula plantea que existen raíces unitarias y por tanto no existe estacionariedad. Para temperatura se tuvo una significancia de 0.023 y para precipitación de 0.038, siendo las dos menores a 0.05. Rechazando de esta manera la hipótesis nula y confirmando la estacionariedad de la serie de datos. En el caso de no existir estacionariedad, se recomienda utilizar otros métodos geostatísticos que no requieran de este supuesto como el Kriging Universal (Torres, 2017).

#### Semivariogramas teóricos

Los semivariogramas teóricos se elaboran a partir del ajuste de los semivariogramas empíricos a una ecuación matemática, que explica mejor el comportamiento espacial



(Lamorey & Jacobson, 1995). Existen varios modelos teóricos, sin embargo se han utilizado los semivariogramas esférico, exponencial y gaussiano, pues estos toman en cuenta la distancia y dirección y son los que mejores resultados han ofrecido para variables climáticas (Abad Encalada, 2014).

En base a los semivariogramas obtenidos, se determinó que el rango o distancia a partir de cual ya no existe correlación entre el valor de las variables fue de 38,275.26m. Y que el lag o radio de búsqueda entre las estaciones para el calculo de las varianzas fue de 3,189.26m para precipitación y 1,871.05 para temperatura, es decir los modelos utilizaron todos los valores de las variables que se encontraban dentro de ese radio.

	<b>Efecto pepita</b>	<b>Sill</b>	<b>Semivariograma</b>
Precipitación Modelo Esférico	0.01	0.049	
Precipitación Modelo Exponencial	0.27	0.78	
Precipitación Modelo Gaussiano	0	0.74	

Tabla 14: Semivariograma teórico Precipitación. +=Promedios

Además todos los modelos tuvieron un efecto pepita menor a 1, por lo que se asumen que se tuvieron errores experimentales que no afectaron la estimación o que no existe



variabilidad espacial por debajo del radio de búsqueda determinado (Antón, Cárdenas, & Freire, 2013). En las Tablas 14 y 15 se tienen los modelos de semivariogramas generados, en los cuales en el eje x se grafica la distancia (m) y en el eje y la varianza.

Los semivariogramas de temperatura (Tabla 15), tienden a la horizontalidad mientras que los de precipitación se presentan de manera creciente hasta llegar al sill.

Una vez conocidos y analizados todos los parámetros de interpolación geostatística para cada estación con promedios horarios de temperatura y precipitación, se obtuvieron mapas raster de las variables.

	<b>Efecto pepita</b>	<b>Sill</b>	<b>Semivariograma</b>
<i>Temperatura Modelo Esférico</i>	0.0038	0.0061	
<i>Temperatura Modelo Exponencial</i>	0.0025	0.0053	
<i>Temperatura Modelo Gaussiano</i>	0.0039	0.0052	

Tabla 15: Semivariograma teórico Temperatura. +=Promedios

Los mapas generados se clasificaron en 5 rangos, en los cuales las zonas con tonalidades bajas de azul son zonas de menor precipitación y aquellas con tonalidades



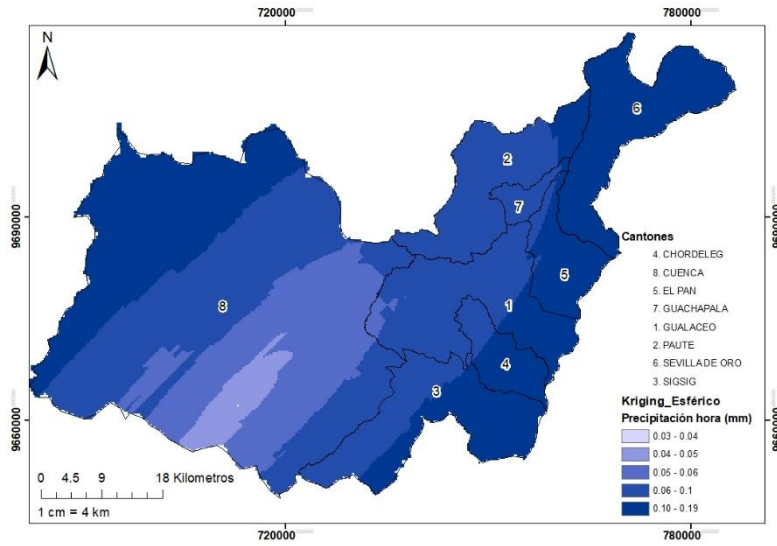


## Universidad de Cuenca

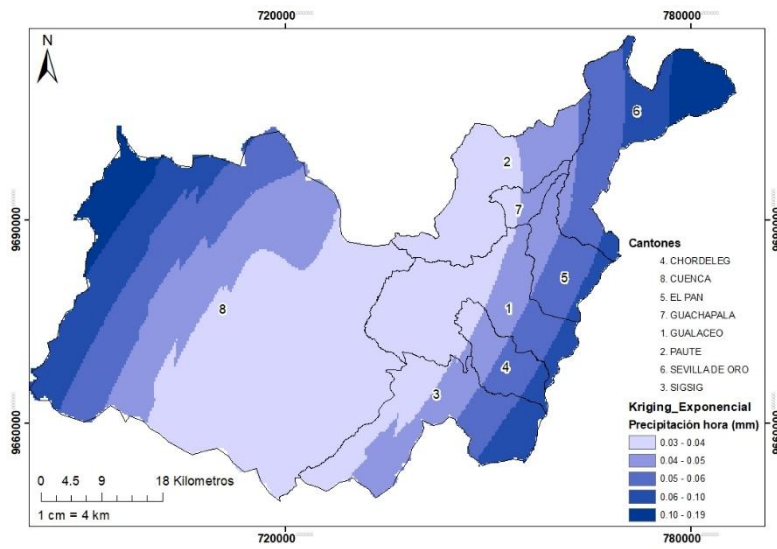
de azul fuerte zonas en las que se produjo mayor precipitación. En la Ilustración 9 se puede observar que los valores estimados de la variables se encontraron en un rango de 0.03 a 0.19mm en los tres modelos empleados. Se identifica que existe mayor precipitación en el este y oeste de la zona de estudio, mientras que la zona central tiene menor cantidad de precipitación, se observa el mismo comportamiento en los tres mapas generados pero con distintos valores. Los cantones de Sevilla de Oro, El Pan y Chordeleg y el oeste del cantón Cuenca son las áreas con mayor precipitación horaria para el año de análisis.



a)



b)



c)

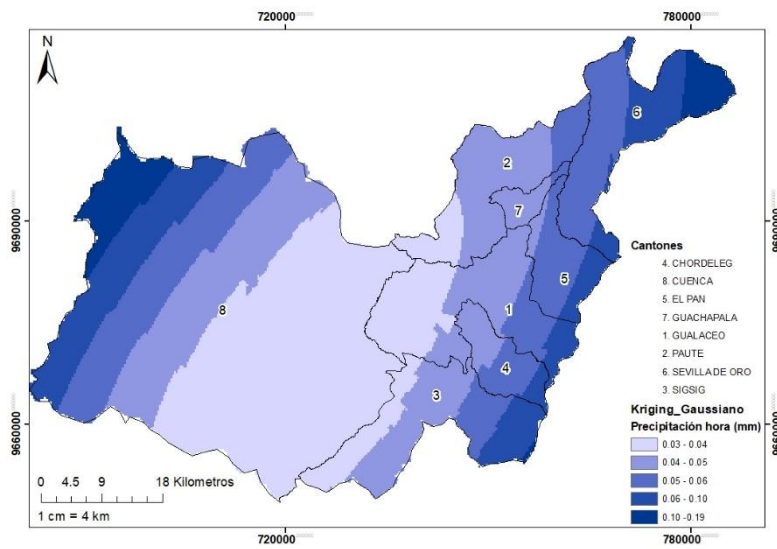




Ilustración 9: Modelos de semivariograma Kriging-Precipitación. a) Mapa Kriging-Modelo Esférico, b) Mapa Kriging-Modelo Exponencial y c) Mapa Kriging-Modelo Gaussiano

Los modelos teóricos para el caso de temperatura tuvieron la misma tendencia que en la precipitación, los valores se encontraron en los siguientes rangos: 5.78 a 20.22 °C para el modelo esférico y exponencial y de 5.89 a 20.26 para el modelo gaussiano (Ilustración 10). Se determinó que las temperaturas más bajas se encuentran en las zona este y sur del área de estudio que abarcan el este del cantón Cuenca y el norte del cantón Sigsig y las temperaturas más altas en la zona noroeste de la misma, cantones de Sevilla de Oro, El Pan, Paute y Chordeleg.

Para Cokriging (Ilustración 11 y 12) se estableció la correlacion existente entre las variables precipitación-altura y temperatura-altura, teniendo una correlación baja-nula de 0.06 y una correlación -0.65, respectivamente. Ello para determinar la influencia que da agregar la variable auxiliar a los modelos. Las dos correlaciones se encontraron por debajo del límite 0.75 recomendado por Guan, Wilson, & Makhnin (2005) para uso de variables auxiliares para el método Cokriging. Sin embargo se utilizó esta variable para ver que resultados se obtenían en el caso de datos a escala horaria. A la variable altura se le realizó una transformación logarítmica para que cumpla con el supuesto de normalidad establecido por los métodos geoestadísticos (Tabla 16).

Tabla 16: Prueba normalidad-altura

	Pruebas de normalidad		
	Shapiro-Wilk		
	Estadístico	gl	Sig.
Altura	.876	16	.033
log_altura	.896	16	.069

En cuanto a los valores que resultaron al agregar la variable altura a la interpolación en el caso de precipitación, los valores mínimos y máximos de precipitación estimados fueron los mismos que los determinados por Kriging para los tres modelos (Ilustración 11), ello puede atribuirse a la baja correlación que presentan las variables de estudio con la variable auxiliar (menor al recomendado).

Para la temperatura (Ilustración 12), se tuvieron valores de 5.64 a 20.31°C para el modelo esférico, 5.65 a 20.31 para el exponencial y 5.86 a 20.23°C para el modelo



## Universidad de Cuenca

gaussiano. En cuanto a distribución espacial se observa la misma tendencia que los mapas generados por Kriging.

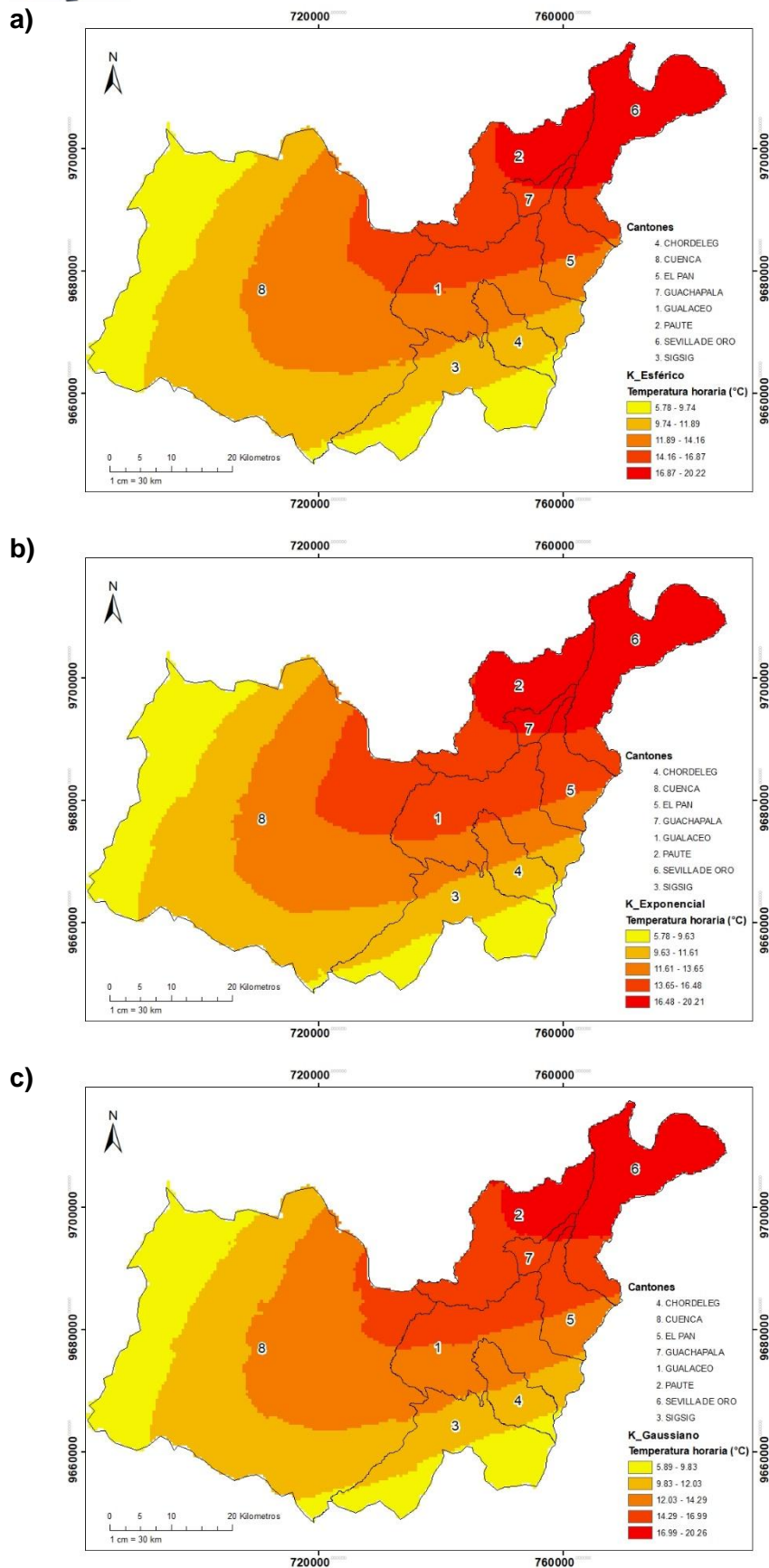
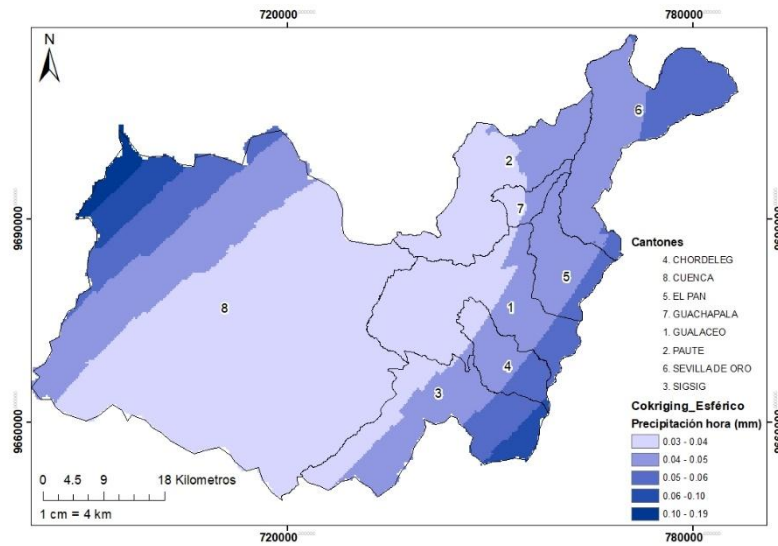
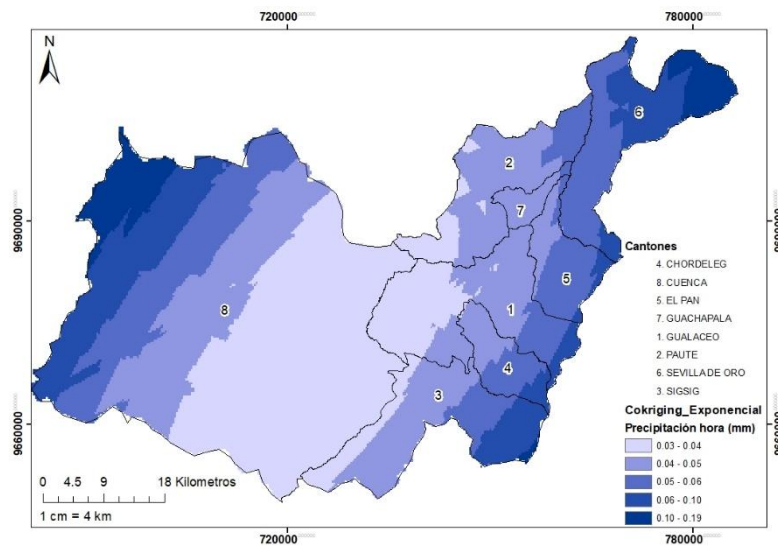


Ilustración 10: Modelos de Semivariograma Kriging-Temperatura. a) Mapa Kriging-Modelo Esférico, b) Mapa Kriging-Modelo Exponencial y c) Mapa Kriging-Modelo Gaussiano

a)



b)



c)

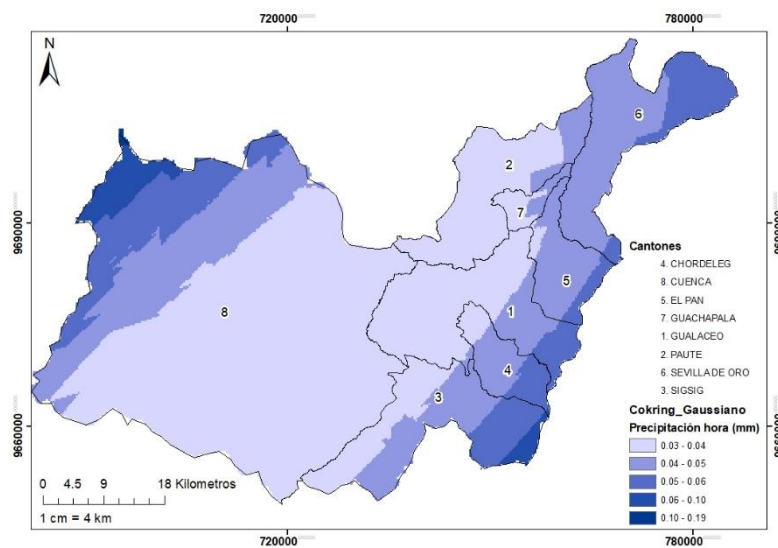


Ilustración 11  
Modelos de Semivariograma Cokriging-Precipitación.  
a) Mapa Cokriging-Modelo Esférico,  
b) Mapa Cokriging-Modelo Exponencial  
c) Mapa Cokriging-Modelo Gaussiano

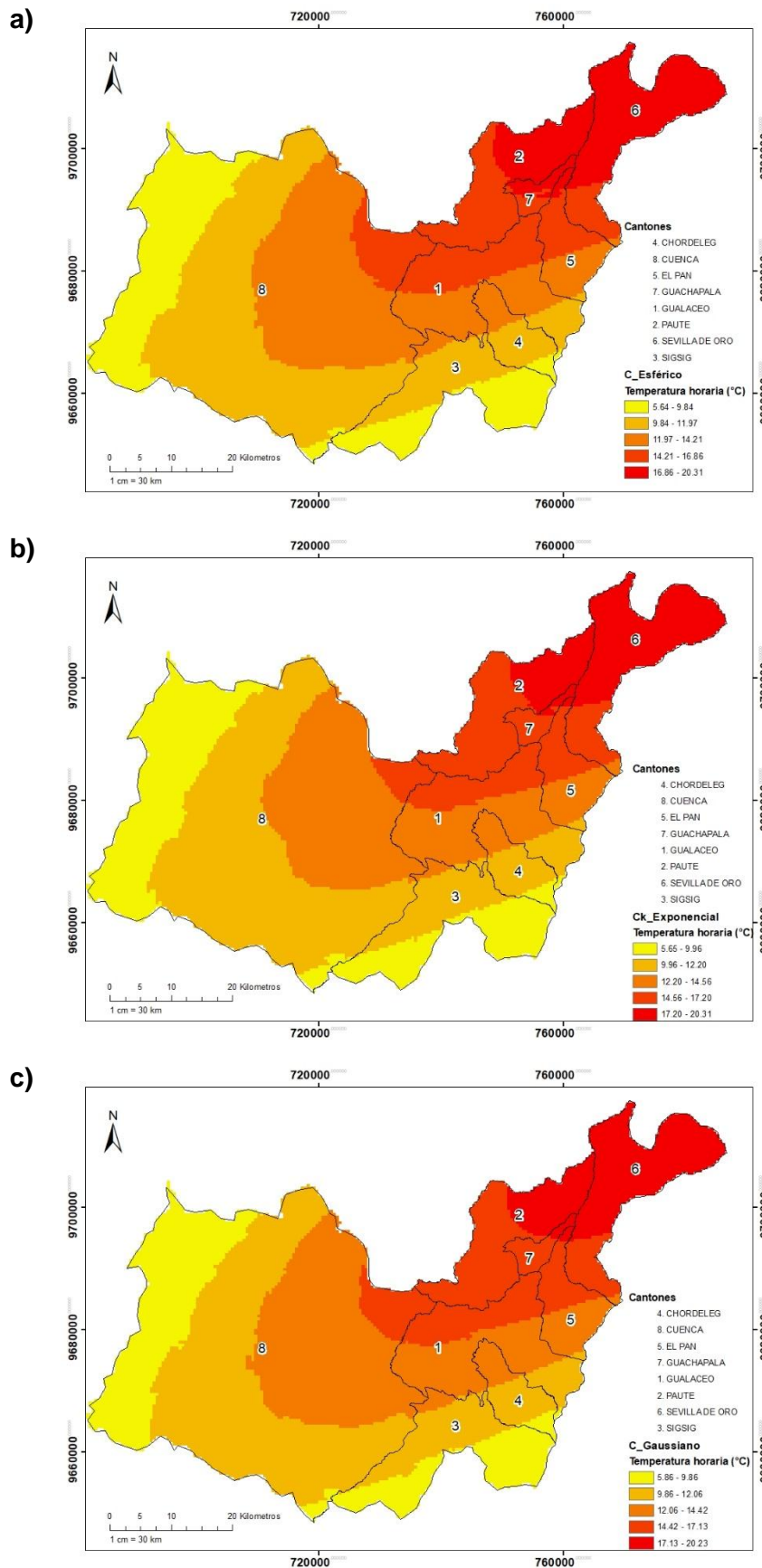


Ilustración 12: Modelos de semivariograma Cokriging-Temperatura. a) Mapa Cokriging-Modelo Esférico, b) Mapa Cokriging-Modelo Exponencial y c) Mapa Cokriging-Modelo Gaussiano

#### 4.2.3 Validación del modelo Kriging y Cokriging

Los resultados obtenidos para la validación cruzada se muestran en la Tabla 17. Se obtuvieron errores menores para los modelos de Kriging, con ello se ratifica lo manifestado por Ly, Charles, & Degre (2011), en cuya investigación determinaron que el mejor método geostadístico para el manejo de datos de precipitación es Kriging Ordinario, debido a que Cokriging no mejoró la precisión de la interpolación. Y también lo expuesto por Mahir & Fares (2011), la información secundaria mejora la predicción del método cuando la correlación entre las variables de estudio (precipitación-temperatura) y la variable auxiliar (elevación) es alta de 0.75.

Tabla 17: Parámetros estadísticos Validación. a) Precipitación, b) Temperatura

	<b>Método</b>	<b>Modelo</b>	<b>Media</b>	<b>MSS</b>	<b>RMSE</b>	<b>KSE</b>	<b>RMSE- KSE</b>	<b>RMSS</b>
a)	Kriging	Esférico	0.007	-0.180	0.054	0.025	0.029	1.979
		Exponencial	0.004	-0.027	0.040	0.020	0.020	1.204
		Gausiano	0.004	-0.183	0.053	0.002	0.051	1.227
	Cokriging	Esférico	0.006	-0.180	0.055	0.025	0.029	1.979
		Exponencial	0.004	-0.090	0.050	0.023	0.037	1.976
		Gausiano	0.005	-0.122	0.051	0.035	0.016	1.378
b)	Kriging	Esférico	-0.013	-0.045	1.096	1.009	0.087	1.232
		Exponencial	-0.010	-0.041	1.079	1.006	0.073	1.222
		Gausiano	-0.009	-0.029	1.064	0.996	0.068	1.196
	Cokriging	Esférico	-0.027	-0.042	0.843	0.986	-0.142	1.072
		Exponencial	-0.022	-0.048	0.854	0.968	-0.114	0.984
		Gausiano	-0.029	-0.054	1.058	1.114	0.143	1.172

Para escoger el modelo que mejor se ajusta al semivariograma experimental se analizó MSS, RMSE y RMSS (Adhikary et al., 2015), se determinó que el modelo gaussiano es el que sin duda se ajustó de mejor manera al semivariograma de la variable temperatura, pues este cumplió con todos la mayoría de parámetros estadísticos. Mientras que en el caso de la variable precipitación el modelo exponencial se ajustó adecuadamente.

Se tienen valores MSS -0.027 y -0.183 para el modelo exponencial y gaussiano para la variable precipitación y de -0.029 para Kriging Gaussiano de temperatura, al ser





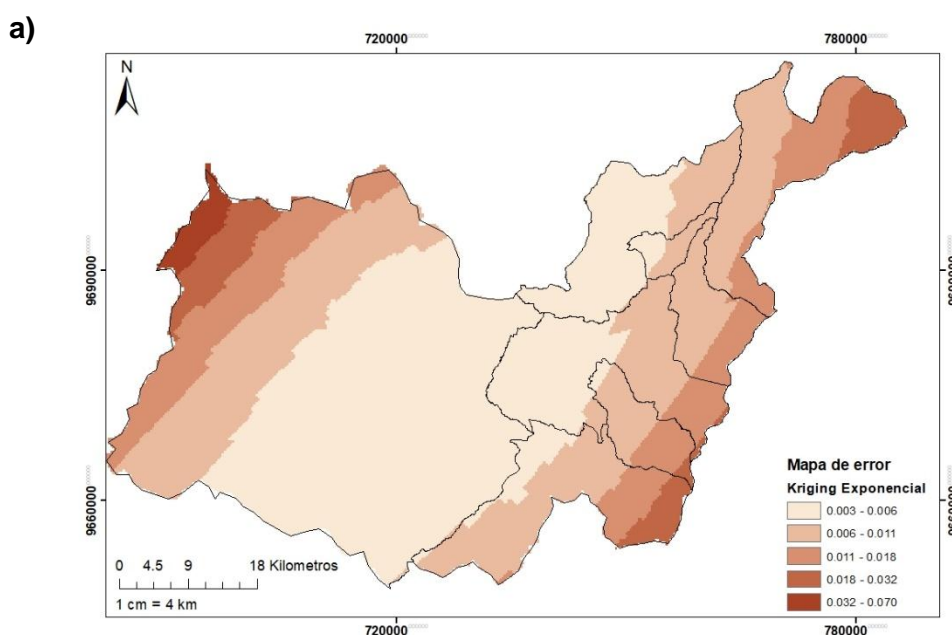
cercanos a 0, se obtuvieron estimaciones imparciales (no desviadas) y no sesgadas (Adhikary et al., 2015) . El RMSE para los modelos de precipitación fue de 0.040 y para temperatura de 1.06, por lo que en el primer caso se muestran predicciones más cercanas a los valores reales que en el segundo caso (Tabla 17).

Este resultado se respalda en las investigaciones realizadas por Zucarelli, Paris, & Macor (2014) quienes en su estudio indicaron que se logró el mejor ajuste con un variograma de tipo exponencial para la variable precipitación.

Una vez establecidos los modelos geostadísticos que mejor se ajustaron se realizó un análisis del mapa de error (Ilustración 13). Los mapas de error muestran la medida del error de estimación de la precipitación promedio horaria, en ellos se puede observar como entre más cerca los puntos se encuentran de las estaciones meteorológicas menor será el error de estimación y a mayor distancia aumentan las semivarianzas (Earls & Dixon, 2007).

Así, en Kriging Exponencial de Precipitación se obtuvieron errores de estimación entre 0.003 y 0.07 y para Kriging Gausiano en temperatura los errores de estimación fueron entre 0.09 y 0.40.

Se observa que los límites de la zona de estudio no se encuentran bien definidos, ello debido a que en estas zonas no se presentaron mediciones de ninguna de las variables por carecer de las estaciones meteorológicas, por lo que la precisión de la estimación en estas zonas siempre será más débil.



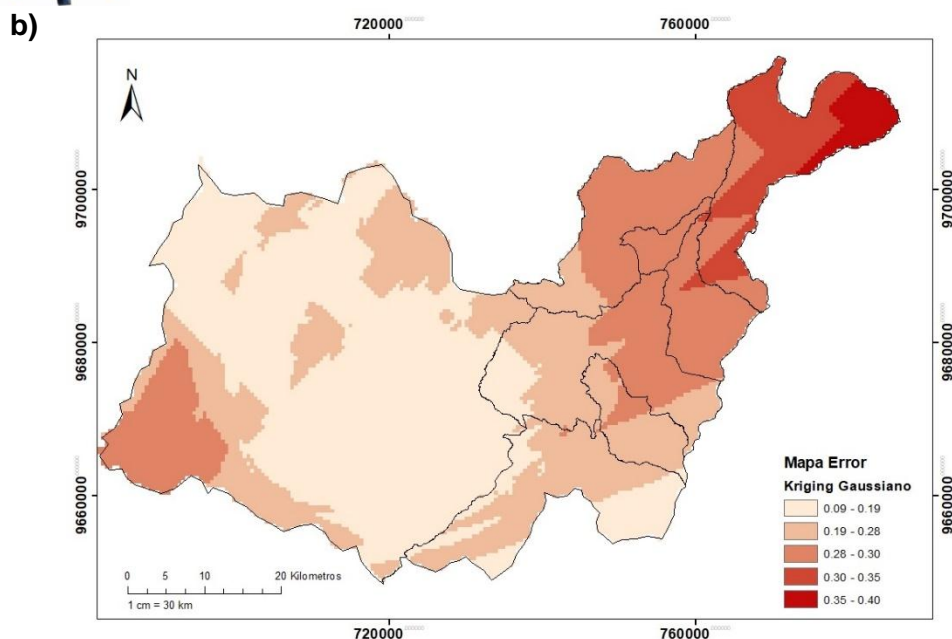


Ilustración 13: Mapas de error. a) Mapa de error Kriging Exponencial-Precipitación y b) Mapa de error Kriging Gaussiano-Temperatura

Además, se realizó una validación cartográfica, mediante el uso de la herramienta Extract Multi Values to Points de ArcMap, la cual extrae los valores de las celdas de un ráster en ubicaciones específicas. Para ello se utilizó los puntos geográficos en los cuales se encuentran las estaciones y los ráster pertenecientes a cada modelo, teniendo como resultado valores semejantes a los valores medio-hora de precipitación y temperatura (Tabla 16).

Tabla 16: Estadística de validación cartográfica. a) Temperatura, b) Precipitación

ID	Estación	Código	Valores reales		Valores estimados	
			a)	b)	a)	b)
11	Nulti	NUL	13.99	0.06	12.57	0.057
12	Santa Ana	SAN	13.19	0.05	13.43	0.048
13	Llacao	LLA	14.22	0.07	10.61	0.075
14	Tixan - Chiquintad	TXC	13.44	0.04	11.28	0.051
15	CTS - Yanuncay	CTS	14.71	0.05	14.86	0.053
16	El Vecino	EVC	15.13	0.06	14.66	0.063
17	Irquis	IRQ	12.30	0.06	12.05	0.058
18	Quingeo	QUI	12.03	0.04	12.98	0.046
19	Cumbe	CUM	10.23	0.06	12.85	0.062



## Universidad de Cuenca

20	Baños	BAN	10.89	0.06	13.35	0.051
21	San Joaquín	SJQ	12.82	0.07	13.68	0.078
22	Sayausí	SAY	13.37	0.05	13.35	0.057
23	Turi	TUR	12.93	0.06	13.78	0.027
18	Sevilla de oro	SEV	16.66	0.09	16.19	0.121
19	Guarumales	GUA	19.52	0.20	19.54	0.183
20	Mazar	MAZ	18.76	0.09	18.60	0.112

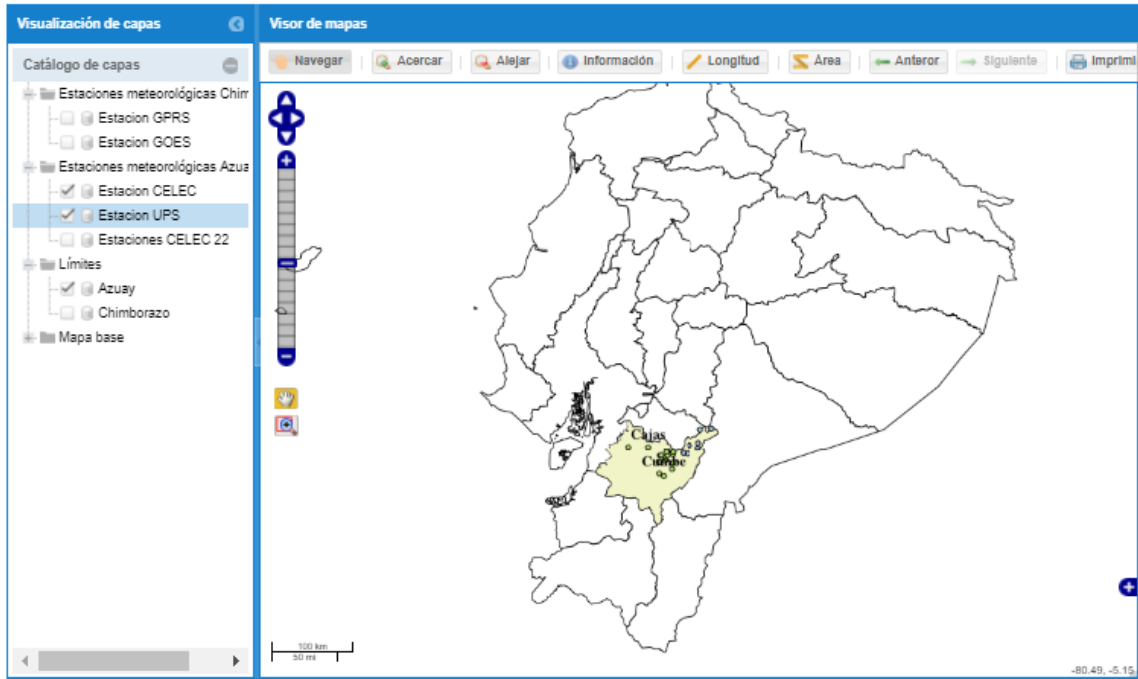
Se confirma que mediante el uso de métodos geoestadísticos Kriging-Cokriging, se obtienen resultados satisfactorios con errores mínimos y muy cercanos a los valores de las estaciones, siendo una metodología confiable para el procesamiento de información meteorológica (Choudhury, Rahman, & Ferdousi, 2014).

Además, los métodos geoestadísticos utilizados a pesar de demandar mayor tiempo y ser complejos, los resultados son mejores en cuanto a la generación de superficies continuas de precipitación o temperatura con menores errores (Abad, 2014).

Adicionalmente, se crearon mapas de superficie de temperatura y precipitación horaria mensual (Anexo 4), para ello se utilizó los dos métodos seleccionados como los que mejores se ajustan a cada variable.

#### 4.2.4 Publicación de resultados

Ya determinada la confiabilidad de los métodos de relleno y seleccionado el método geostadístico para la interpolación (Kriging Gausiano para las dos variables), los métodos serán incorporados a la plataforma web. La cuál recibe información de las estaciones meteorológicas cada hora. El mecanismo mediante el cual se obtiene los datos son con Web Services. Estos solicitan dos Web Service, uno que da los datos históricos y un segundo web service que da los datos en tiempo real. Cuando el usuario solicite los datos a una determinada hora, el web service pedirá los datos al servidor de la Empresa proveedora de los mismos (UPS-CELEC), una vez obtenidos los datos de las diferentes estaciones inmediatamente se aplicará el método Kriging Exponencial y Gausiano para generar el mapa de estimación y dar la respuesta al usuario.



*Ilustración 14: Geoportal-Provincia del Azuay*

Hasta la fecha (Septiembre-2017) ya se ha creado el Geoportal, en el que se puede observar el mapa del Ecuador, destacándose la provincia del Azuay, en la cual se encuentran las estaciones a las cuales se tiene acceso. Sin embargo, falta todavía implementar los métodos de relleno y el geostatístico.



## 5. CONCLUSIONES

- A través de los métodos geoestadístico empleados en esta investigación (Kriging y Cokriging) se pudo determinar una temperatura media-horaria estuvo en un rango de 5.86 a 20.23 °C y una precipitación promedio-hora en un rango de 0.03 a 0.19mm, durante el período de un año (31 de julio 2016-2017) en la zona noreste de la provincia del Azuay.
- Los métodos de relleno analizados, método de proporcionalidades y media aritmética, resultan satisfactorios para llenar series de precipitación y temperatura horaria respectivamente, pues valores faltantes estimados presentaron medidas estadísticas adecuadas, aunque no excelentes. La metodología de la media aritmética presenta el limitante de la obligatoriedad de existencia de datos en las horas requeridas, lo que en ciertos casos hace necesario que se utilicen los datos de estaciones con las que se tiene menos correlación. Mediante la aplicación de los dos métodos se obtuvo una base de datos completa para la aplicación posterior de los métodos geoestadísticos.
- El Kriging ha resultado una buena metodología de interpolación para estimar los valores de las variables meteorológicas de estudio en zonas en las que no se conoce esta información, por carecer de una estación meteorológica, a pesar de no contar con la cantidad de puntos (estaciones) recomendadas para la aplicación del método. Se comprobó la importancia de entender y manejar correctamente todos los conceptos relacionados con la geoestadística para la correcta aplicación del método y obtener valores predichos muy cercanos a los reales.

Entre los tres modelos de variograma utilizados, el modelo exponencial fue el que mejor se ajustó a precipitación y el gaussiano para temperatura, debido a lo cual será utilizado para la interpolación espacial de los datos horarios, en el geoportal.

- El rendimiento de los dos métodos geoestadísticos no tuvo variaciones grandes, por lo que se concuerda con otros autores, en que al aumentar otra



## Universidad de Cuenca

variable a la interpolación (Cokriging) no se obtienen mejores resultados, sin embargo hay que tomar en cuenta el grado de correlación de las variables.

- Los métodos de relleno y geostadísticos analizados son lo suficientemente robustos y simples para implementarlos en la plataforma web, lo que permitirá obtener información a tiempo real de temperatura o precipitación, de acuerdo a las necesidades del usuario.

### 6. RECOMENDACIONES Y LIMITACIONES

- No existen suficientes estudios en los que se analicen los métodos de relleno o geostadísticos para variables meteorológicas a escala horaria, por lo que no se tiene valores para realizar una comparación. Debido a lo cual, la mayoría de las comparaciones estadísticas y de modelamiento, se realizaron con estudios a escala diaria y anual. Sin embargo, esta investigación puede ser una base para posteriores estudios.
- La cantidad de estaciones disponibles para el estudio (16) permitieron obtener resultados satisfactorios para la zona noreste de la provincia, sin embargo es importante que más instituciones compartan la información meteorológica que poseen para obtener resultados de toda la provincia en estudio.
- En la presente investigación se pudo trabajar solo con un año de información meteorológica debido a la necesidad de tener mayor cantidad de estaciones para la aplicación de los métodos geoestadísticos, sin embargo se recomienda utilizar mayor cantidad de datos debido a que se trabaja a una escala horaria y la precipitación es muy variable.



## 7. BIBLIOGRAFÍA

- Abad Encalada, I. O. (2014). Estimación de precipitación espacial mediante correlación con variables secundarias y la ayuda de tecnologías de información geográficas. Retrieved from <http://dspace.ucuenca.edu.ec/handle/123456789/19857>
- Adhikary, S. K., Yilmaz, A. G., & Muttil, N. (2015). Optimal design of rain gauge network in the Middle Yarra River catchment, Australia. *Hydrological Processes*, 29(11), 2582–2599. <https://doi.org/10.1002/hyp.10389>
- Ahmadi, S. H., & Sedghamiz, A. (2008). Application and evaluation of kriging and cokriging methods on groundwater depth mapping. *Environmental Monitoring and Assessment*, 138(1–3), 357–368. <https://doi.org/10.1007/s10661-007-9803-2>
- Análisis geoestadístico con ArcGIS parte 2. Análisis exploratorio de los datos. (2011). Retrieved August 30, 2017, from <http://www.aguaysig.com/2011/03/analisis-geoestadistico-con-arcgis.html>
- Cañizares, M., Barroso, I., & Alfonso, K. (2004). Datos incompletos: una mirada crítica para su manejo en estudios sanitarios. *Gaceta Sanitaria*, 18(1), 58–63.
- Carrasco, J. A. (2014). Description: Maestría en Sistemas de Información Geográfica y Teledetección Aplicada al Medio Ambiente V3.0. Retrieved August 16, 2017, from <http://postgradofcauagrm.com/aulavirtual/course/index.php?categoryid=9>
- Censos (Ecuador), I. N. de E. y. (2001). *VI censo de población y V de vivienda, 2001: Provincia del Azuay (3 v.)*. República del Ecuador, Instituto Nacional de Estadística y Censos.
- Choudhury, N. H., Rahman, A., & Ferdousi, S. (2014). Kriging Infill of Missing Data and Temporal Analysis of Rainfall in North Central Region of Bangladesh. *Journal of Climatology & Weather Forecasting*, 1–5. <https://doi.org/10.4172/2332-2594.1000141>



- Congreso (, A. E. de C. (2004). *El clima, entre el mar y la montaña*. Ed. Universidad de Cantabria.
- Daniele, D. (2009). *Aplicación de sistemas de información geográfica al estudio de acuíferos costeros complejos. Caso del campo de Dalías*. Universidad Almería.
- Díaz Viera, M. A. (2002). *Geoestadística Aplicada*. Cuba.
- Draxler, R., & Chai, T. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature.
- Dubrule, O. (1983). Cross validation of kriging in a unique neighborhood. *Journal of the International Association for Mathematical Geology*, 15(6), 687–699.  
<https://doi.org/10.1007/BF01033232>
- Earls, J., & Dixon, B. (2007). Spatial Interpolation of Rainfall Data Using ArcGIS: A Comparative Study. *ResearchGate*. Retrieved from [https://www.researchgate.net/publication/255574434\\_Spatial\\_Interpolation\\_of\\_Rainfall\\_Data\\_Using\\_ArcGIS\\_A\\_Comparative\\_Study](https://www.researchgate.net/publication/255574434_Spatial_Interpolation_of_Rainfall_Data_Using_ArcGIS_A_Comparative_Study)
- Gallardo, A. (2006). *Geostatística*.
- García, F. J. M. (2004). Aplicación de la geoestadística en las ciencias ambientales. *Ecosistemas*, XIII(1), 0.
- Garzon, D. (2015). Missing Values Analysis.
- Geoestadística - EcuRed. (2010). Retrieved October 23, 2016, from <https://www.ecured.cu/Geoestad%C3%ADstica>
- Giraldo Henao, R. (2002). *Introducción a la Geoestadística, Teoría y Aplicación*. Colombia.
- González, J., Guerra, F., & Gómez, H. (2007). Conceptos Básicos De Geoestadística En Geografía Y Ciencias De La Tierra: Manejo Y Aplicación. Retrieved October 16, 2016, from <http://www.redalyc.org/articulo.oa?id=36014577008>
- Guan, H., Wilson, J. L., & Makhnin, O. (2005). Geostatistical Mapping of Mountain Precipitation Incorporating Autosearched Effects of Terrain and Climatic





Characteristics. *Journal of Hydrometeorology*, 6(6), 1018–1031.

<https://doi.org/10.1175/JHM448.1>

Guerra, F., Gómez, H., González, J., & Zambrano, Z. (2006a). Uso actual de métodos y técnicas para el estudio de la precipitación incluyendo plataformas SIG. *Geoenseñanza*, 11(1). Retrieved from <http://www.redalyc.org/resumen.oa?id=36012424009>

Guerra, F., Gómez, H., González, J., & Zambrano, Z. (2006b). Uso actual de métodos y técnicas para el estudio de la precipitación incluyendo plataformas SIG. Retrieved May 17, 2017, from <http://www.redalyc.org/articulo.oa?id=36012424009>

Guzman, V., & León, A. (2012). IMPACTO DE LA CALIDAD DE DATOS SOBRE EL CIERRE DEL BALANCE HÍDRICO EN CUENCAS DE PÁRAMO Y BOSQUES (MAZAR Y SOLDADOS).

Haddad, K., Rahman, A., A Zaman, M., & Shrestha, S. (2013). Applicability of Monte Carlo cross validation technique for model development and validation using generalised least squares regression. *Journal of Hydrology*, 482, 119–128. <https://doi.org/10.1016/j.jhydrol.2012.12.041>

Hidrología, I. N. de M. e. (2016). Instituto Nacional de Meteorología e Hidrología. Retrieved May 26, 2017, from <http://www.serviciometeorologico.gob.ec>

Lamorey, G., & Jacobson, E. (1995). Estimation of semivariogram parameters and evaluation of the effects of data sparsity. *Mathematical Geology*, 27(3), 327–358. <https://doi.org/10.1007/BF02084606>

Lituma, L., & Andreina, M. (2011). Estudio hidrológico, diseño del alcantarillado pluvial y corrección hidrológica de las quebradas que atraviesan la cabecera parroquial de Pachicutza, cantón El Pangui. Retrieved from <http://dspace.utpl.edu.ec/handle/123456789/678>

Loja, M. (2011). “ESTUDIO HIDROLÓGICO, DISEÑO DEL ALCANTARILLADO PLUVIAL Y CORRECCIÓN HIDROLÓGICA DE LAS QUEBRADAS QUE



ATRAVIESAN LA CABECERA PARROQUIAL DE PACHICUTZA, CANTÓN EL PANGUI.”

- López, C. P. (2006). *Problemas resueltos de econometría*. Editorial Paraninfo.
- Ly, S., Charles, C., & Degre, A. (2011). Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrology and Earth System Sciences*, 15(7).  
<https://doi.org/10.5194/hess-15-2259-2011>
- Machado, D., Puche, M., Vilorio, J., Parra, M., & Primavera, M. (2007). Geoespacialización de información climática mediante la utilización de sistemas de información geográfica. Caso cuenca alta del Río Guárico. Retrieved from <http://ceer.isa.utl.pt/cyted/venezuela2008/Geoespacializacion.pdf>
- Mahir, A., & Fares, A. (2011). Comparison of Rainfall Interpolation Methods in a Mountainous Region of a Tropical Island. Retrieved August 30, 2017, from <https://www.colwiz.com>
- María, L. (2009). Informe sobre evidencias en Córdoba (Andalucía).
- Martinez-Vargas, A., & Ramírez-García, J. (2005). Desarrollo Actual De La Geoestadística En El Mundo. Retrieved October 23, 2016, from <http://www.redalyc.org/articulo.oa?id=223516052001>
- Monte Alegre, J. (1990). *Instituto Colombiano de Hidrología, Meteorología y adecuación de tierras*.
- Oliver, M. A., & Webster, R. (2015). *Basic Steps in Geostatistics: The Variogram and Kriging*. Springer.
- Patiño, E., & Lupercio, L. (2017). Geoprocesamiento Web (WPS “Web Processing Service”) para la generación de mapas hidrometeorológicos a partir de las estaciones automáticas de las provincias de Azuay y Chimborazo.
- Pizarro T, R., Ramirez B, C., V, F., & Pablo, J. (2003). Análisis comparativo de cinco métodos para la estimación de precipitaciones areales anuales en períodos



Universidad de Cuenca

extremos. *Bosque (Valdivia)*, 24(3), 31–38. <https://doi.org/10.4067/S0717-92002003000300003>

Ross, S. M. (2007). *Introducción a la estadística*. Reverté. Retrieved from <https://books.google.com.ec/books?id=pPM2TgQsx8wC>

Teegavarapu, R. S. V. (2012a). *Floods in a Changing Climate: Extreme Precipitation*. Cambridge University Press. Retrieved from <https://books.google.com.ec/books?id=KJa6tEGMzwQC>

Teegavarapu, R. S. V. (2012b). *Floods in a Changing Climate: Extreme Precipitation*. Cambridge University Press. Retrieved from <https://books.google.com.ec/books?id=KJa6tEGMzwQC>

Torres, L. (2017). ANÁLISIS DE LA EVAPORACIÓN DE CORRIENTES HÍDRICAS USANDO MÉTODOS DE INTERPOLACIÓN ESPACIAL EN ANTIOQUIA. Retrieved from [http://ridum.umanizales.edu.co:8080/xmlui/bitstream/handle/6789/3018/Torres\\_Lina\\_Maria\\_2017.pdf?sequence=3&isAllowed=y](http://ridum.umanizales.edu.co:8080/xmlui/bitstream/handle/6789/3018/Torres_Lina_Maria_2017.pdf?sequence=3&isAllowed=y)

Urrutia, J. A., Palomino, R., & Salazar, H. D. (2010). Metodología Para La Imputación De Datos Faltantes En Metereología. Retrieved October 23, 2016, from <http://www.redalyc.org/articulo.oa?id=84920977010>

Veerasamy, R., Rajak, H., Jain, A., Sivadasan, S., Christopher, P. V., & Agrawal, R. K. (2011). Validation of QSAR Models - Strategies and Importance. *Int J Drug Design and Discov*, 2, 511–519.

Zhang, H., & Wang, Y. (2010). Kriging and cross-validation for massive spatial data. *Environmetrics*, 21(3–4), 290–304. <https://doi.org/10.1002/env.1023>

Zucarelli, A., Paris, M., & Macor, J. (2014). Utilización de kriging para la elaboración de curvas isohietas de precipitación mensual en la Provincia de Santa Fe, Argentina. Retrieved from <http://www.tvdigitalargen.ina.gov.ar/ifrh-2014/Eje3/3.39.pdf>

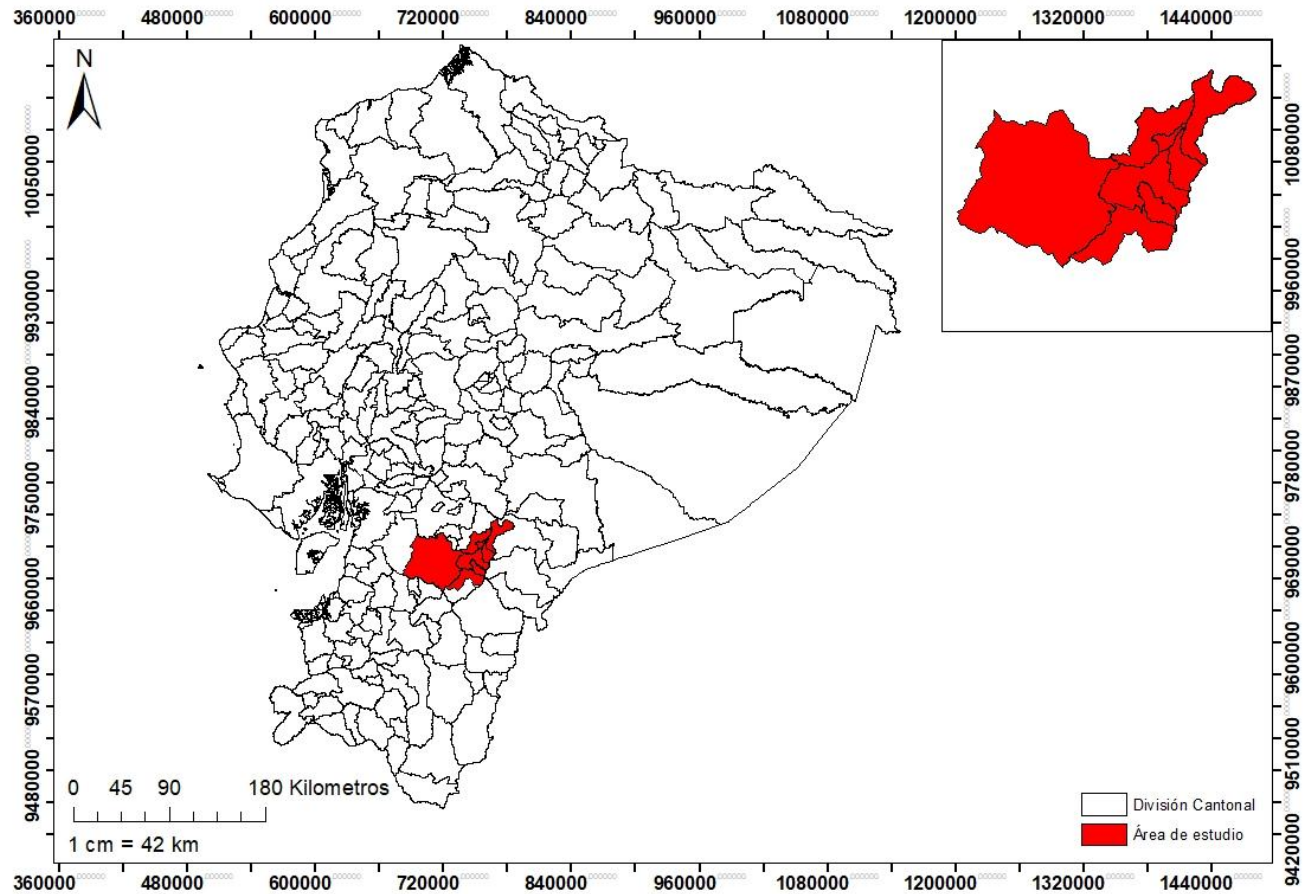


Universidad de Cuenca



### 8. ANEXOS

#### Anexo 1: Zona de estudio



*Ilustración 15: Mapa de la República del Ecuador (área de estudio)*

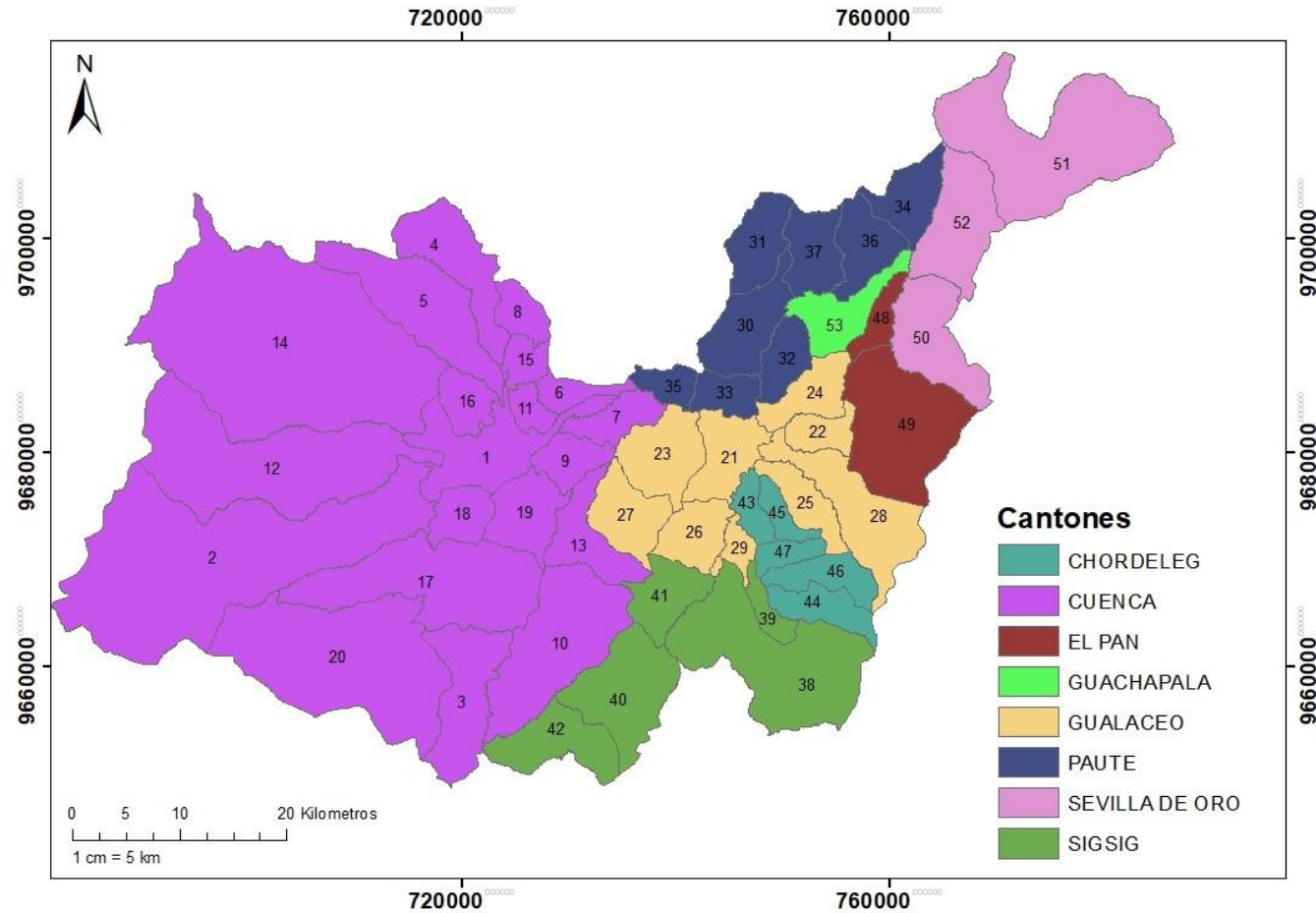


Ilustración 16: División cantonal del área de estudio



ID	Parroquia	ID	Parroquia	ID	Parroquia
1	CUENCA	19	VALLE	37	DUG DUG
2	BAÑOS	20	VICTORIA DEL PORTETE	38	SIGSIG
3	CUMBE	21	GUALACEO	39	GUEL
4	CHECA (JIDCAY)	22	DANIEL CORDOVA	40	LUDO
5	CHIQUINTAD	23	JADAN	41	SAN BARTOLOME
6	LLACAO	24	MARIANO MORENO	42	SAN JOSE DE RARANGA
7	NULTI	25	REMIGIO CRESPO	43	CHORDELEG
8	OCTAVIO CORDERO	26	SAN JUAN	44	PRINCIPAL
9	PACCHA	27	ZHIDMAD	45	LA UNION
10	QUINGEO	28	LUIS CORDERO VEGA	46	LUIS GALARZA
11	RICAU RTE	29	SIMON BOLIVAR	47	SAN MARTIN PUZHIO
12	SAN JOAQUIN	30	PAUTE	48	EL PAN
13	SANTA ANA	31	BULAN	49	SAN VICENTE
14	SAYAUSI	32	CHICAN	50	SEVILLA DE ORO
15	SIDCAY	33	EL CABO	51	AMALUZA
16	SININCAY	34	GUARAINAG	52	PALMAS
17	TARQUI	35	SAN CRISTOBAL	53	GUACHAPALA
18	TURI	36	TOMEBAMBA		

*Parroquias pertenecientes al área de estudio*



Anexo 2: Correlaciones entre las estaciones

		Correlaciones-Precipitación															
est		BAN	CTS	CUM	IRQ	NUL	SAY	SJQ	SAN	TXC	TUR	EVC	SEV	GUA	MAZ	LLA	QUI
Correlación de Pearson	LLA	.146	.211	.103	.135	.668	.130	.188	.207	.345	.180	.203	.097	.062	.083	1.000	.132
	NUL	.173	.248	.128	.133	1.000	.304	.299	.281	.406	.271	.374	.083	.092	.089	.668	.178
	TXC	.163	.237	.131	.123	.406	.239	.283	.299	1.000	.170	.530	.077	.061	.081	.345	.181
	CTS	.370	1.000	.173	.187	.248	.338	.431	.243	.237	.575	.402	.111	.079	.121	.211	.212
	SAN	.272	.243	.273	.159	.281	.151	.171	1.000	.299	.257	.318	.090	.081	.086	.207	.279
	EVC	.261	.402	.123	.133	.374	.488	.347	.318	.530	.334	1.000	.118	.082	.093	.203	.160
	SJQ	.361	.431	.167	.196	.299	.573	1.000	.171	.283	.349	.347	.136	.158	.152	.188	.157
	TUR	.443	.575	.237	.216	.271	.288	.349	.257	.170	1.000	.334	.105	.102	.129	.180	.211
	BAN	1.000	.370	.282	.255	.173	.322	.361	.272	.163	.443	.261	.135	.127	.174	.146	.252
	IRQ	.255	.187	.370	1.000	.133	.116	.196	.159	.123	.216	.133	.132	.074	.161	.135	.236
	QUI	.252	.212	.240	.236	.178	.089	.157	.279	.181	.211	.160	.095	.070	.110	.132	1.000
	SAY	.322	.338	.138	.116	.304	1.000	.573	.151	.239	.288	.488	.077	.109	.094	.130	.089
	CUM	.282	.173	1.000	.370	.128	.138	.167	.273	.131	.237	.123	.079	.093	.104	.103	.240
	SEV	.135	.111	.079	.132	.083	.077	.136	.090	.077	.105	.118	1.000	.314	.354	.097	.095
	MAZ	.174	.121	.104	.161	.089	.094	.152	.086	.081	.129	.093	.354	1.000	.492	.083	.110
GUA	.127	.079	.093	.074	.092	.109	.158	.081	.061	.102	.082	.314	.492	1.000	.062	.070	

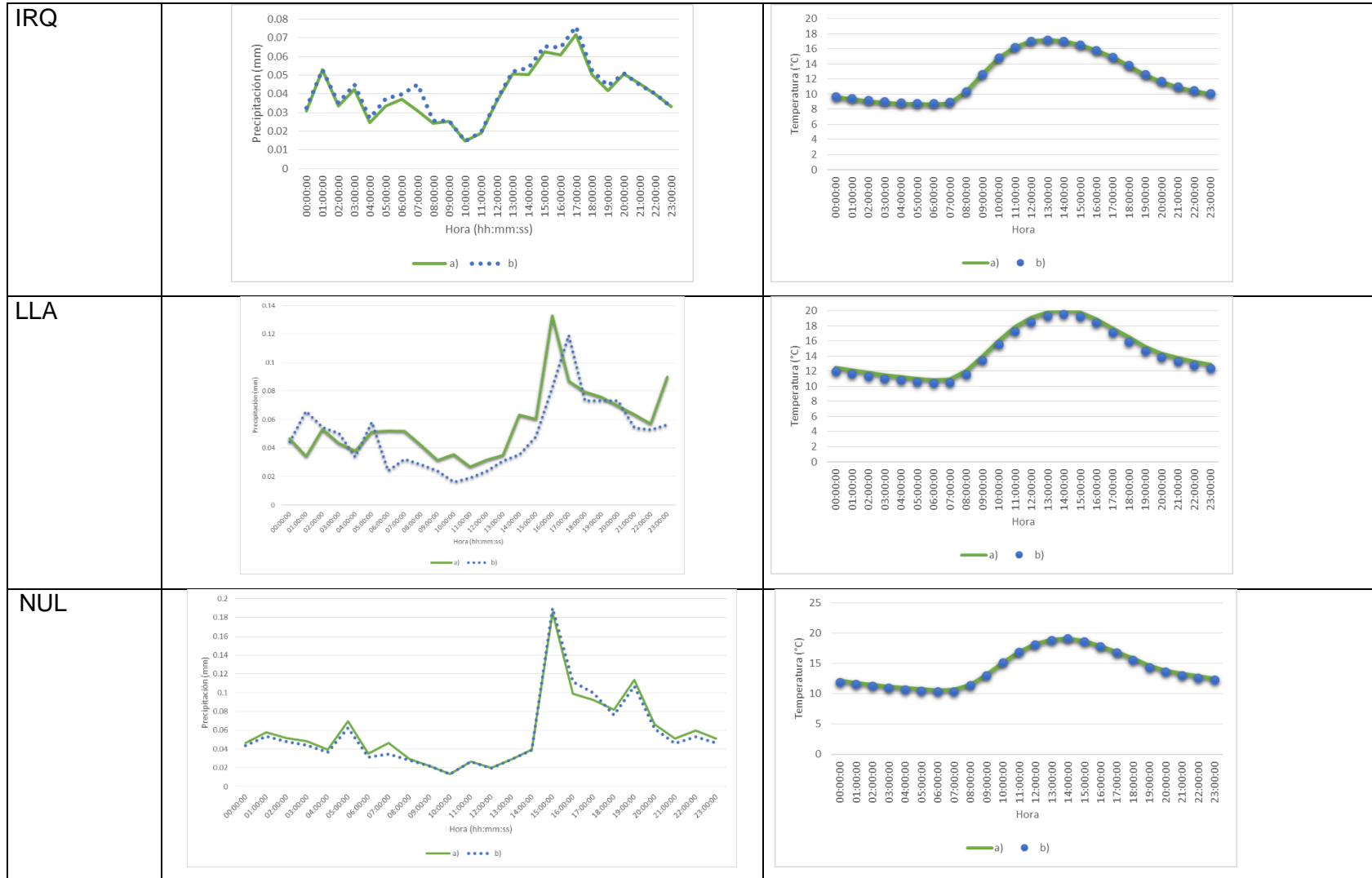
		Correlaciones-Temperatura															
		LLA	BAN	CTS	CUM	IRQ	NUL	QUI	SAY	SJQ	SAN	TXC	TUR	EVC	SEV	GUA	MAZ
Correlación de Pearson	TXC	.983	.949	.964	.899	.852	.977	.948	.977	.972	.946	1.000	.959	.982	.699	.624	.912
	LLA	1.000	.936	.969	.903	.869	.990	.949	.966	.960	.954	.983	.963	.983	.679	.601	.910
	EVC	.983	.940	.983	.893	.858	.984	.943	.973	.966	.948	.982	.969	1.000	.711	.615	.905
	SAY	.966	.957	.974	.905	.870	.963	.949	1.000	.993	.949	.977	.967	.973	.667	.591	.903
	NUL	.990	.936	.973	.902	.866	1.000	.945	.963	.958	.947	.977	.970	.984	.661	.584	.904
	SJQ	.960	.959	.968	.908	.876	.958	.949	.993	1.000	.945	.972	.965	.966	.655	.583	.903
	CTS	.969	.934	1.000	.897	.891	.973	.938	.974	.968	.947	.964	.979	.983	.661	.562	.890
	TUR	.963	.952	.979	.920	.882	.970	.943	.967	.965	.939	.959	1.000	.969	.625	.556	.905
	BAN	.936	1.000	.934	.928	.823	.936	.946	.957	.959	.917	.949	.952	.940	.661	.617	.901
	QUI	.949	.946	.938	.930	.877	.945	1.000	.949	.949	.966	.948	.943	.943	.671	.606	.902
	SAN	.954	.917	.947	.897	.898	.947	.966	.949	.945	1.000	.946	.939	.948	.666	.580	.883
	MAZ	.910	.901	.890	.867	.798	.904	.902	.903	.903	.883	.912	.905	.905	.730	.671	1.000
	CUM	.903	.928	.897	1.000	.861	.902	.930	.905	.908	.897	.899	.920	.893	.564	.533	.867
	IRQ	.869	.823	.891	.861	1.000	.866	.877	.870	.876	.898	.852	.882	.858	.484	.404	.798
	SEV	.679	.661	.661	.564	.484	.661	.671	.667	.655	.666	.699	.625	.711	1.000	.820	.730
	GUA	.601	.617	.562	.533	.404	.584	.606	.591	.583	.580	.624	.556	.615	.820	1.000	.671

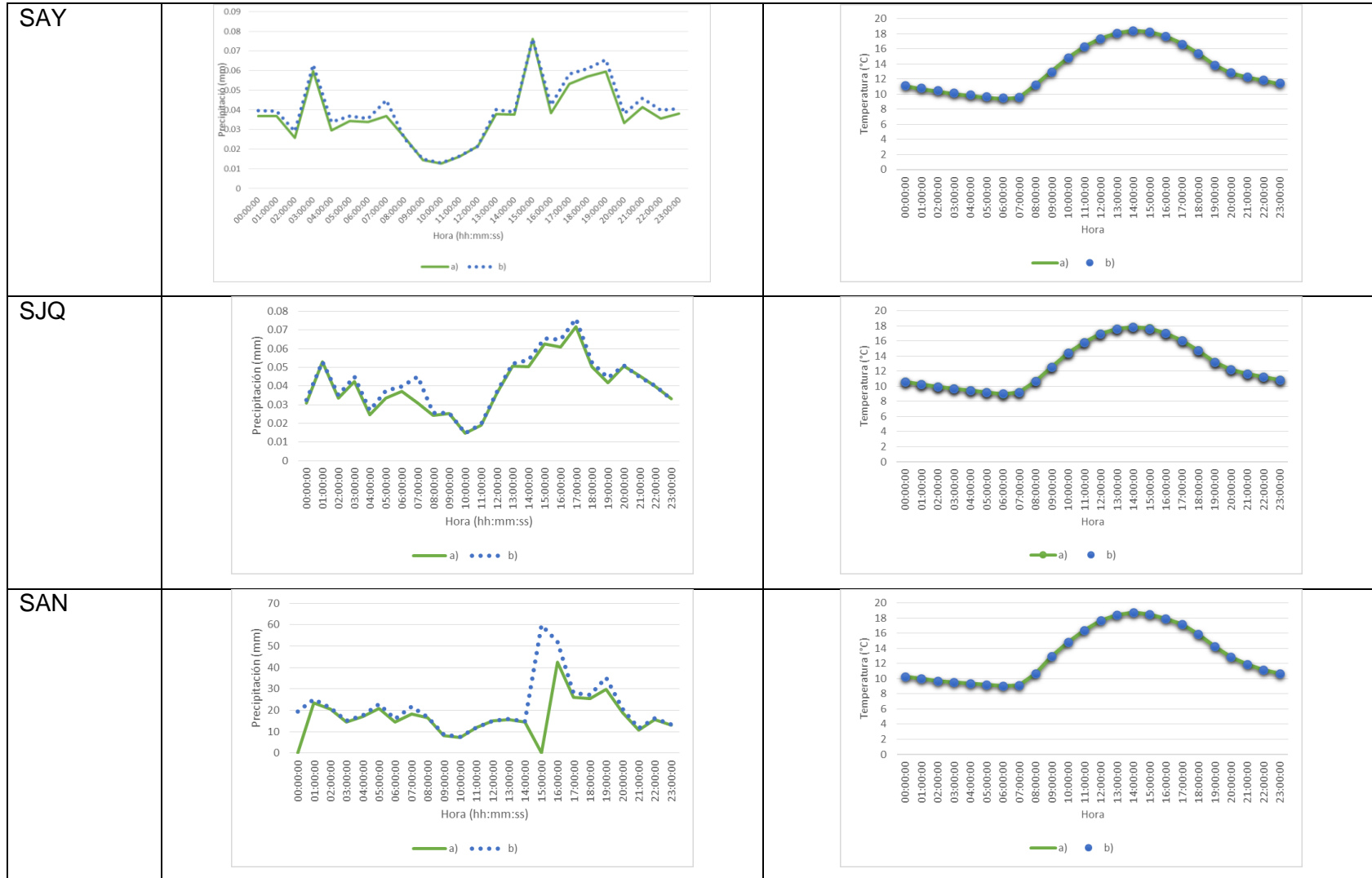




Anexo 3: Estaciones con datos faltantes vs Estaciones con datos completos

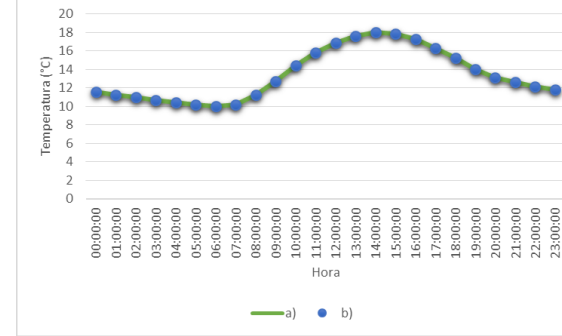
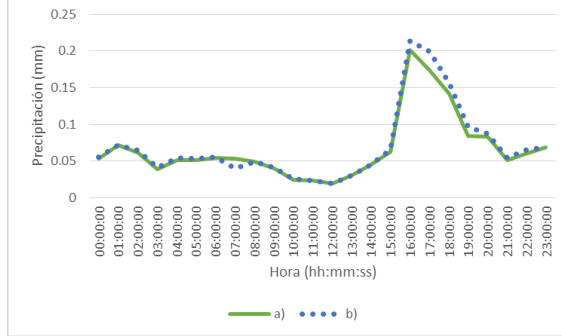
ID	Precipitación	temperatura
CTS		
CUM		







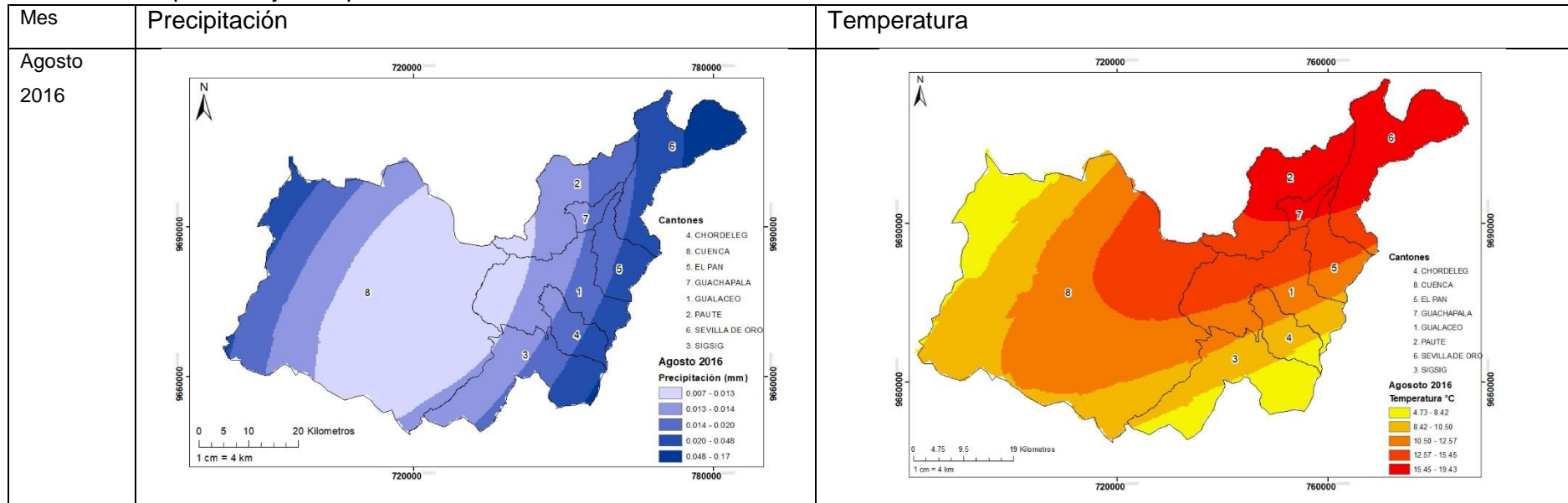
TXC



a) Estaciones con datos faltantes. b) Estaciones con datos completos.

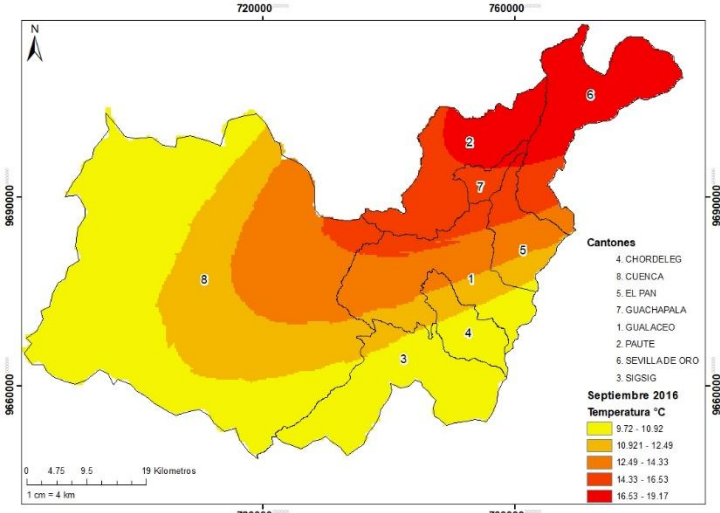
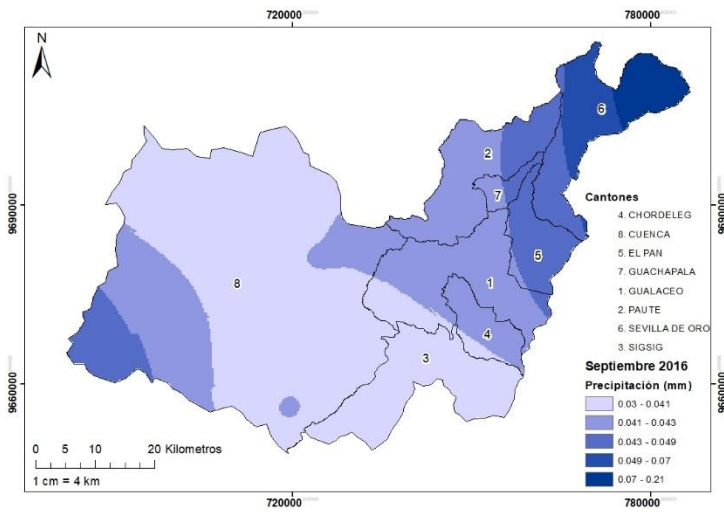


Anexo 4: Temperatura y Precipitación mensual

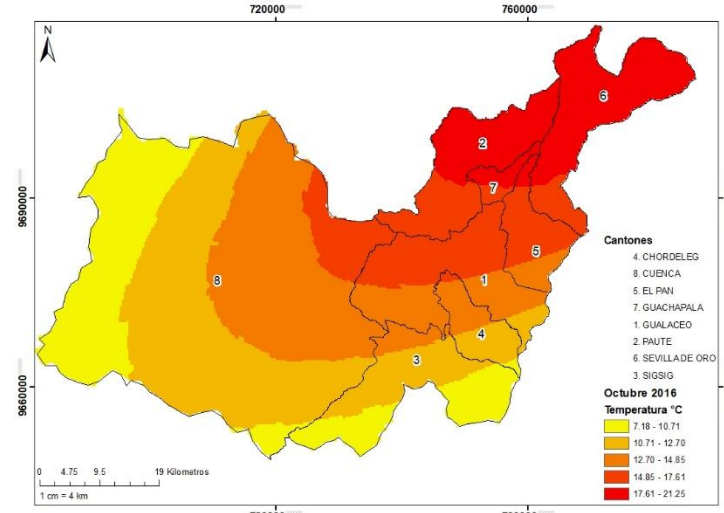
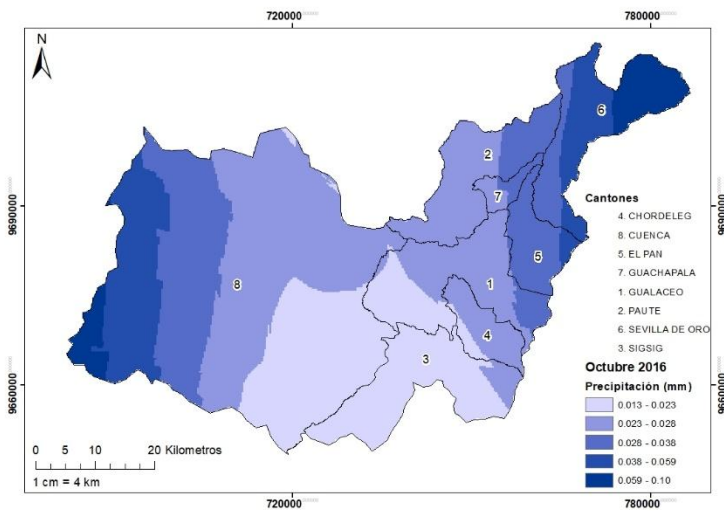




Septiembre 2016

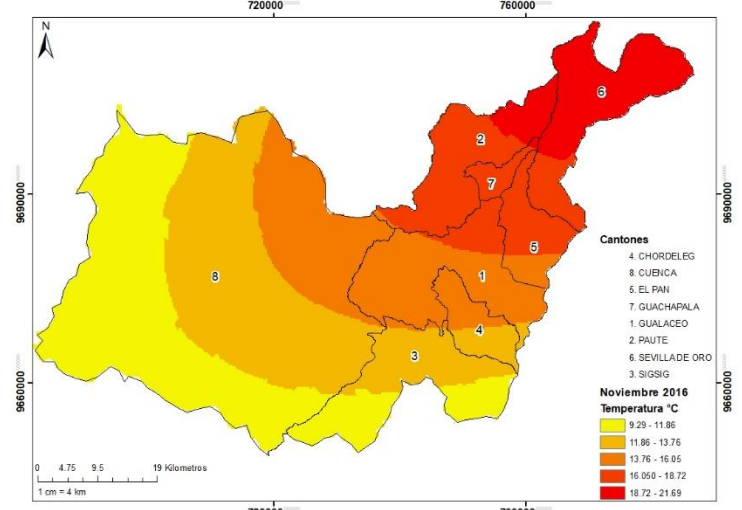
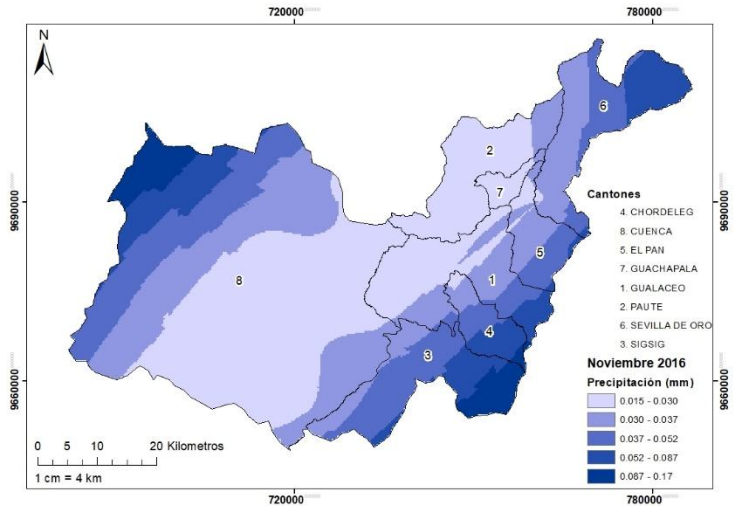


Octubre 2016

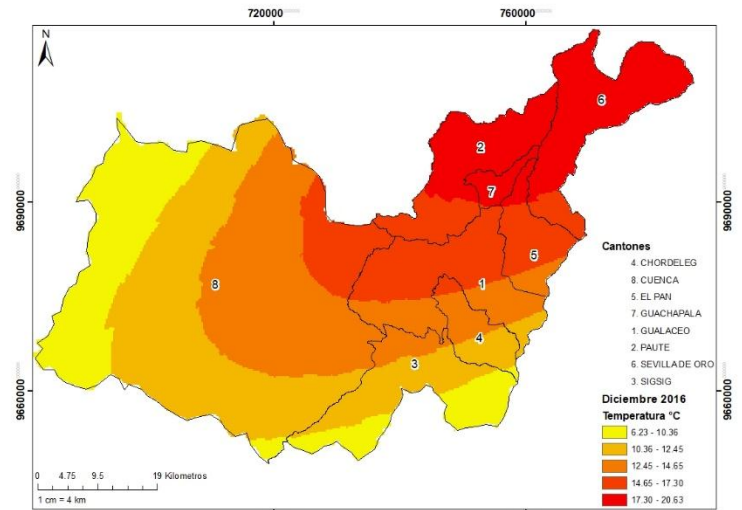
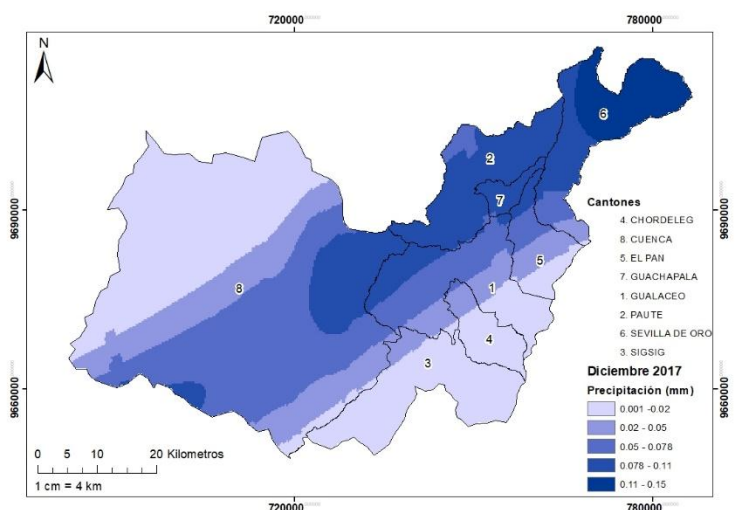


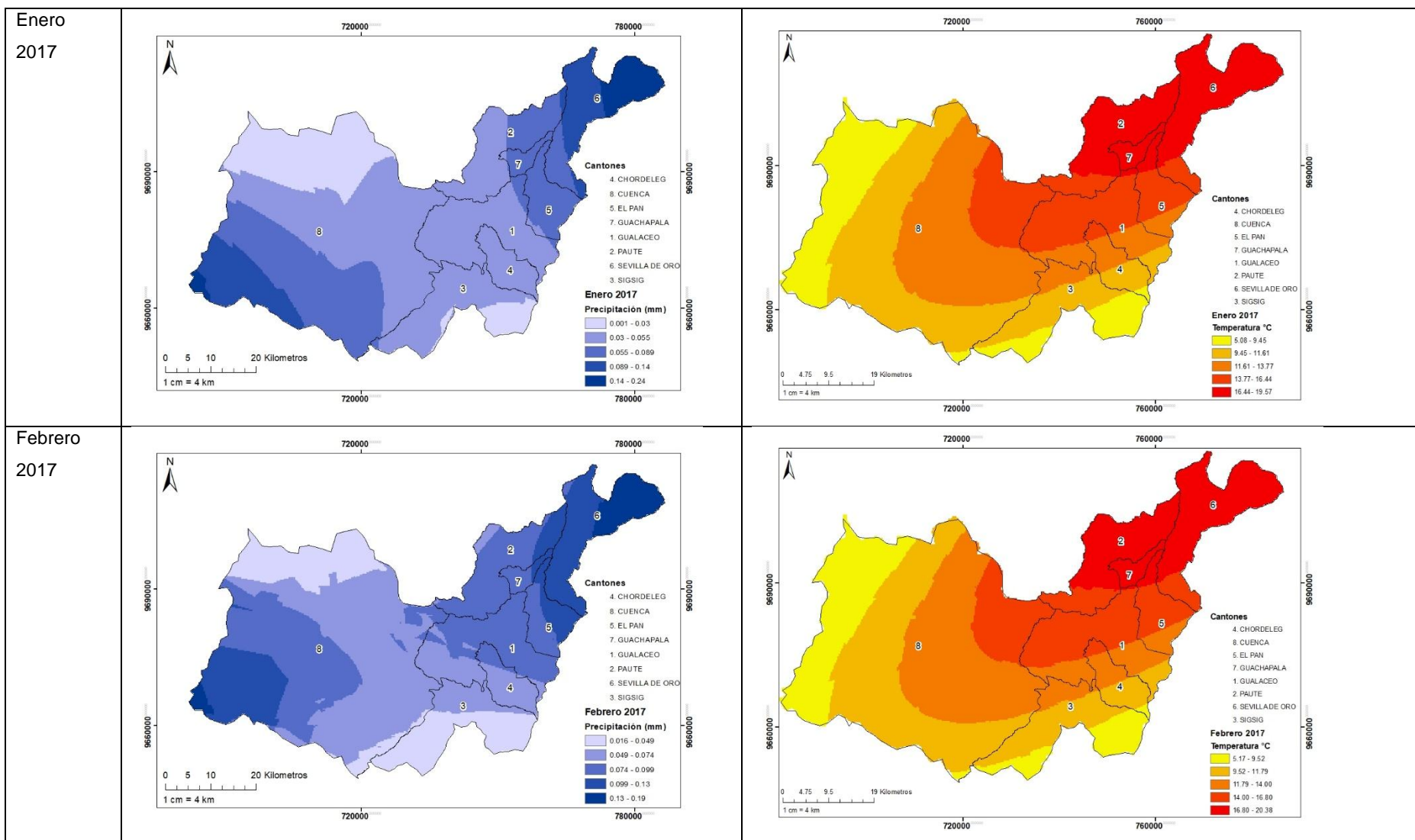


Noviembre 2016



Diciembre 2016

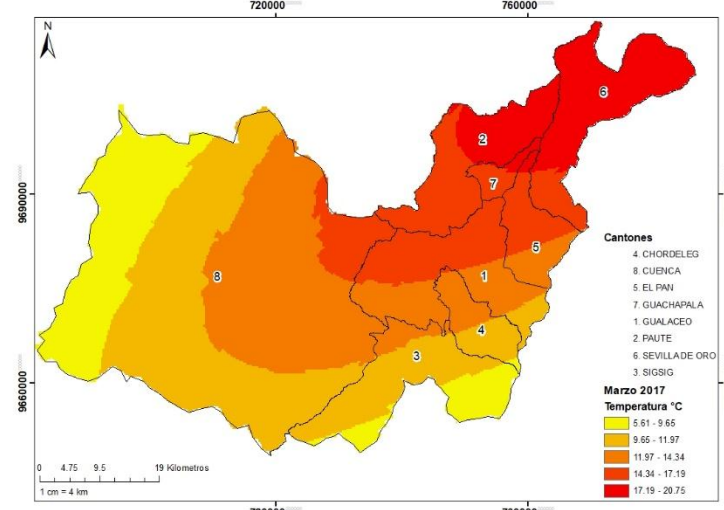
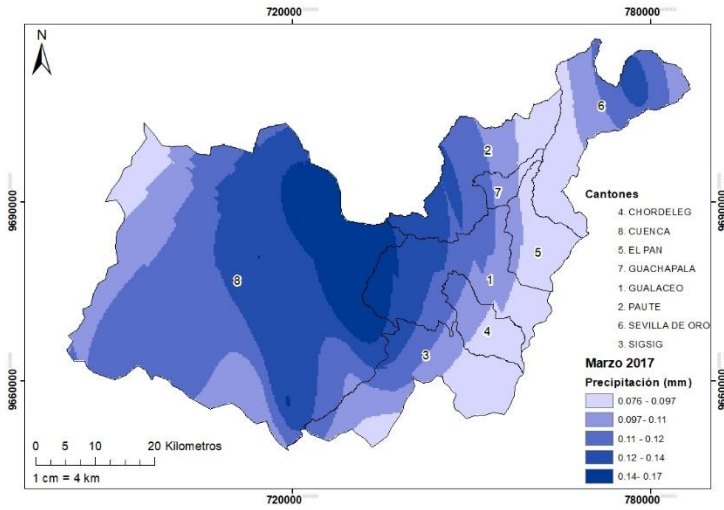




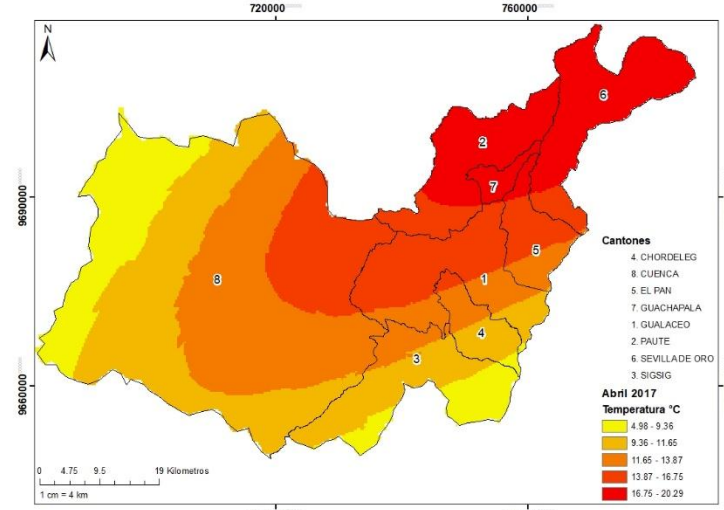
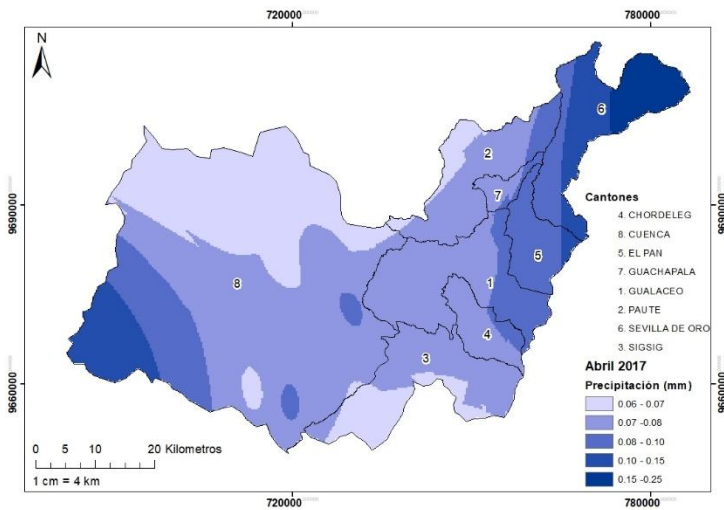




Marzo  
2017

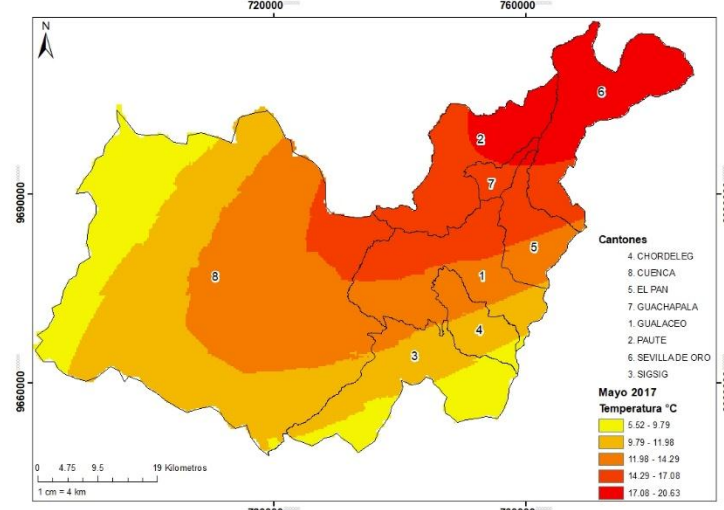
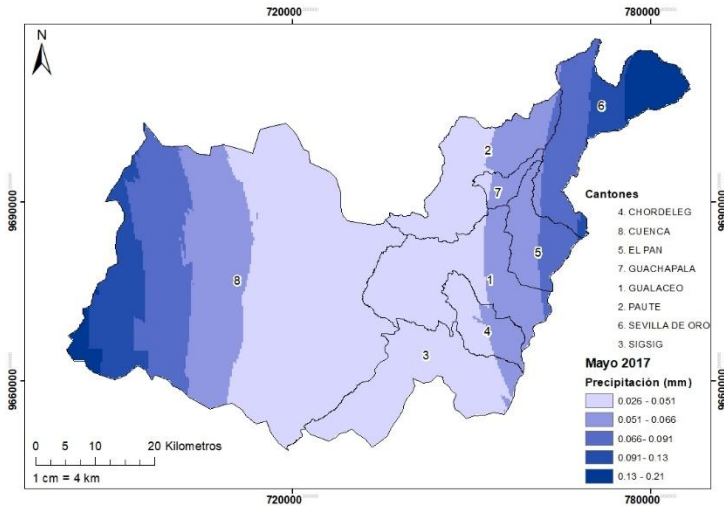


Abril  
2017

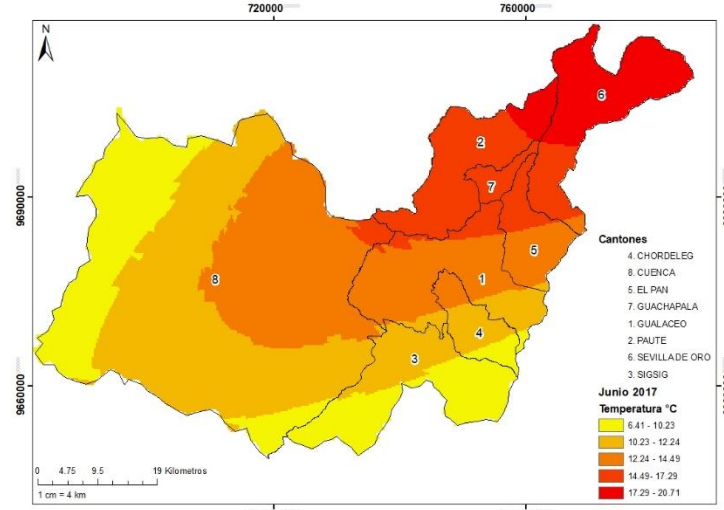
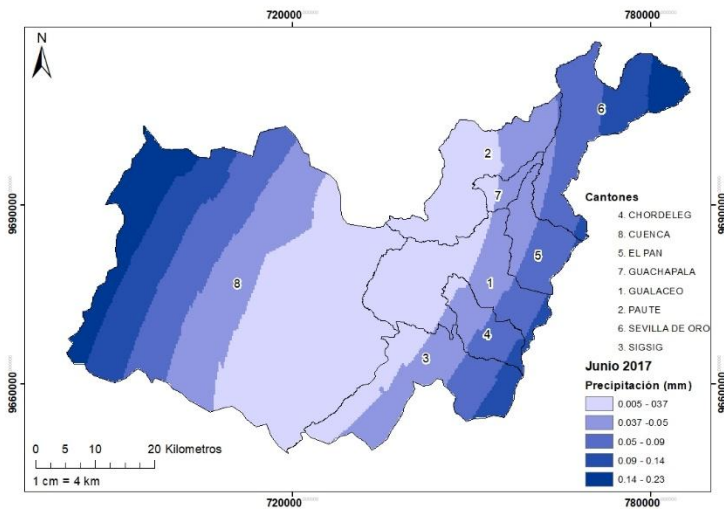


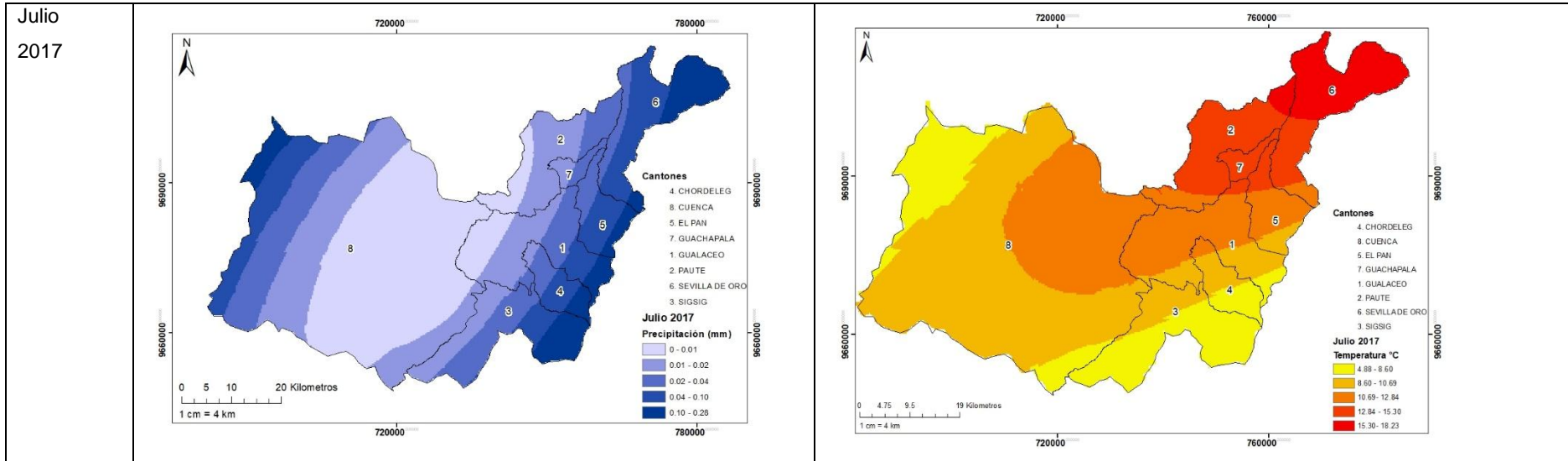


Mayo  
2017



Junio  
2017





\*Los mapas generados se realizaron utilizando cartografía del INEC 2011 y Datum WGS84.