



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

FACULTAD DE CIENCIAS

CARRERA ESTADÍSTICA

**ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES DE
CONTAGIOS DE SARS-COV-2 (COVID-19), EN LA PROVINCIA DE
CHIMBORAZO PERÍODO 2020-2021, A TRAVÉS DEL CÁLCULO,
MODELACIÓN Y COMPARACIÓN DE VARIOGRAMAS**

Trabajo de Titulación

Tipo: Proyecto de Investigación

Presentado para optar al grado académico de:

INGENIERO/A ESTADÍSTICO/A

AUTORES: EDISON RIGOBERTO SAYAY TENELEMA

JOHANA ELIZABETH GUZMAN LEON

DIRECTORA: ING. PAULINA FERNANDA BOLAÑOS LOGROÑO, MGS.

Riobamba Ecuador

2023

©2023, Johana Elizabeth Guzman Leon y Edison Rigoberto Sayay Tenelema

Autorizamos la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento, siempre y cuando se reconozca el Derecho de Autor.

Nosotros, Johana Elizabeth Guzman Leon y Edison Rigoberto Sayay Tenelema, declaramos que el presente Trabajo de Titulación es de nuestra autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autores asumimos la responsabilidad legal y académica de los contenidos de este Trabajo de Titulación; el patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 13 de diciembre de 2023



Johana Elizabeth Guzman Leon

060544404-1

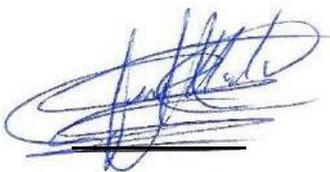
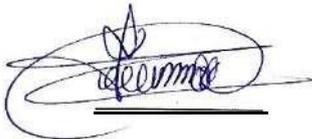


Edison Rigoberto Sayay Tenelema

060432601-7

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA ESTADÍSTICA

El Tribunal del Trabajo de Titulación, certifica que: El Trabajo de Titulación; Tipo: Proyecto de Investigación, **ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES DE CONTAGIOS DE SARS-COV-2 (COVID-19), EN LA PROVINCIA DE CHIMBORAZO PERÍODO 2020-2021, A TRAVÉS DEL CÁLCULO, MODELACIÓN Y COMPARACIÓN DE VARIOGRAMAS**, realizado por los señores: **EDISON RIGOBERTO SAYAY TENELEMA Y JOHANA ELIZABETH GUZMAN LEON**, ha sido minuciosamente revisado por los Miembros del Tribunal del Trabajo de Titulación, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal autoriza su presentación.

	FIRMA	FECHA
Biof. Tania Paulina Morocho Barrionuevo, Mgs. PRESIDENTE DEL TRIBUNAL		2023-12-13
Ing. Paulina Fernanda Bolaños Logroño, Mgs. DIRECTORA DEL TRABAJO DE TITULACIÓN		2023-12-13
Ing. Johanna Enith Aguilar Reyes, Mgs. ASESORA DEL TRABAJO DE TITULACIÓN		2023-12-13

DEDICATORIA

El presente trabajo de investigación se lo dedico a mi padre José María Tenelema Tenelema en el cielo, por todo el esfuerzo realizado para brindarme todos los recursos necesarios cuando estaba con vida. A mi madre Mariana Chinlli por todo el amor y la paciencia. A mi madre de sangre Piedad Tenelema por darme la vida, cariño y valores. A mi hija Eliana Sayay que ha sido fuente de inspiración para culminar con mi trabajo de titulación. A mi esposa por ser mi compañera de vida. A mis hermanos Luis, Jaime, Manuel por sus palabras de motivación que han ayudado a continuar con mi carrera profesional. Y a mis hermanos de sangre por ser el soporte y motivación para no rendirme.

Edison

Este trabajo va dedicado a mi madre, por ser el pilar más importante de mi vida, por su cariño y apoyo incondicional, a mis abuelitos Beatriz y Agustín quienes confiaron en mi esfuerzo, constancia y dedicación, a mi familia por la confianza y aliento que depositaron en mi para cumplir mis metas y anhelos que me he propuesto, a mi hermano que me ha dado la fortaleza para permanecer firme en todas las circunstancias buenas y malas de mi vida, a mis profesores por su tiempo, enseñanza y toda la sabiduría que me transmitieron en el proceso de mi formación.

Johana

AGRADECIMIENTO

En primer lugar, quiero agradecer a Dios por todas las bendiciones derramadas sobre mi, a mi padre en el cielo quien me ha brindado todo el apoyo necesario para poder cumplir uno de mis sueños, a mi familia y a mi esposa quienes me han proporcionado la fuerza necesaria para no rendirme a mitad del camino, agradezco a la Escuela Superior Politécnica de Chimborazo a la Facultad de Ciencias por haberme permitido ser parte de tan noble institución para poder estudiar mi carrera, así como también a los diferentes docentes quienes con su conocimiento y experiencia me ayudaron en el transcurso de mi vida universitaria. Mi agradecimiento también a la Ing. Johanna Aguilar y la Ing. Paulina Bolaños por el tiempo y paciencia que han tenido para asesorarme en el transcurso de este proyecto de investigación.

Edison

Mi agradecimiento más sincero a la Escuela Superior Politécnica de Chimborazo, Facultad de Ciencias por haberme brindado sus sabios conocimientos para mi formación académica. A la Ing. Paulina, directora de Tesis y a la Ing. Johanna miembro del tribunal por su paciencia guía y asesoramiento en la culminación en el presente proyecto de investigación. Agradezco a mi madre y familiares que me ayudaron tanto económicamente como sentimentalmente, a mis maestros y compañeros con quienes compartimos buenos y duros momentos en el transcurso de la carrera muy agradecida con la escuela de Estadística quienes se han encargado de nuestra formación gracias a su experiencia y conocimientos. A todos quienes de manera directa e indirecta me alentaron a cumplir este propósito, con palabras de aliento y motivación.

Johana

ÍNDICE DE CONTENIDO

ÍNDICE DE TABLAS	xi
ÍNDICE DE ECUACIONES.....	xii
ÍNDICE DE ILUSTRACIONES	xiii
ÍNDICE DE ANEXOS	xv
RESUMEN.....	xvi
ABSTRACT	xvii
INTRODUCCIÓN.....	1
CAPÍTULO I	
1. PROBLEMA DE INVESTIGACIÓN	5
1.1. Planteamiento del problema	5
1.2. Limitaciones y delimitaciones	5
1.3. Problema general de investigación	5
1.4. Problemas específicos de investigación	6
1.5. Objetivos	6
1.5.1. <i>Objetivo general</i>	6
1.5.2. <i>Objetivos específicos</i>	6
1.6. Justificación	7
1.6.1. <i>Justificación teórica</i>	7
1.6.2. <i>Justificación metodológica</i>	7
1.6.3. <i>Justificación práctica</i>	7
1.6.4. <i>Hipótesis</i>	8

CAPÍTULO II

2.	MARCO TEÓRICO	9
2.1.	Referencias teóricas	9
2.1.1.	<i>Análisis exploratorio de datos espaciales</i>	9
2.1.2.	<i>Datos espaciales</i>	10
2.1.3.	<i>Sistemas de información geográficas</i>	10
2.1.4.	<i>Autocorrelación espacial</i>	10
2.1.5.	<i>Autocorrelación espacial (LISA)</i>	11
2.1.6.	<i>Índices de autocorrelación espacial</i>	11
2.1.6.1.	<i>Índices global de moran</i>	11
2.1.6.2.	<i>Índice local de autocorrelación espacial (LISA)</i>	12
2.1.7.	<i>Distribución espacial</i>	13
2.1.8.	<i>Base cartográfica espacial</i>	14
2.1.9.	<i>Geometrías puntuales</i>	14
2.1.10.	<i>Patrones puntuales</i>	15
2.1.11.	<i>Análisis estructural espacial</i>	15
2.1.12.	<i>Interpolación espacial</i>	15
2.1.12.1.	<i>Interpolación (IDW) distancia inversa ponderada</i>	16
2.1.12.2.	<i>Interpolación krigging</i>	17
2.1.13.	<i>Variograma</i>	17
2.1.14.	<i>Semivariograma</i>	19
2.1.15.	<i>Función umbral</i>	20
2.1.16.	<i>Error cuadrático medio</i>	20
2.1.17.	<i>Modelación de variogramas</i>	21
2.1.17.1.	<i>Modelo lineal</i>	22
2.1.17.2.	<i>Modelo esférico</i>	22
2.1.17.3.	<i>Modelo exponencial</i>	23
2.1.17.4.	<i>Modelo gaussiano</i>	23
2.2.	Referencias Conceptuales	24
2.2.1.	<i>Indicadores de morbilidad</i>	24
2.2.2.	<i>Sarc-cov-2 (COVID-19)</i>	24

2.2.3.	<i>Contagios</i>	25
2.2.4.	<i>Contagio de una enfermedad</i>	25
2.2.5.	<i>Variables geográficas</i>	25
2.2.6.	<i>Latitud</i>	25
2.2.7.	<i>Longitud</i>	26
2.2.8.	<i>Unidades espaciales</i>	26
2.2.9.	<i>Áreas geográficas en un (SIG)</i>	26

CAPÍTULO III

3.	MARCO METODOLÓGICO	27
3.1.	Enfoque de investigación	27
3.2.	Nivel de Investigación	27
3.3.	Diseño de investigación	27
3.3.1.	<i>Según la manipulación o no de la variable independiente</i>	27
3.3.2.	<i>Según las intervenciones en el trabajo de campo</i>	28
3.4.	Población de estudio	28
3.4.1.	<i>Localización de estudio</i>	28
3.4.2.	<i>Técnicas de recolección de datos</i>	28
3.5.	Métodos, técnicas e instrumentos de investigación	28
3.6.	Operacionalización de las variables	29

CAPÍTULO IV

4.	MARCO DE ANÁLISIS E INTERPRETACIÓN DE RESULTADOS	31
4.1.	Análisis descriptivo	31
4.2.	Análisis exploratorio espacial	38
4.2.1.	<i>Mapa temático de cuantiles</i>	40
4.2.2.	<i>Medidas centrográficas</i>	41
4.2.3.	<i>Mapa de intensidad</i>	42
4.3.	Análisis estructural espacial	43
4.3.1.	<i>Análisis del Índice de Moran</i>	43
4.3.2.	<i>Análisis de autocorrelación espacial (LISA)</i>	44
4.3.3.	<i>Análisis de vecinos más próximos</i>	45

4.3.4.	<i>Ejecución modelación de variogramas</i>	46
4.3.4.1.	<i>Modelación lineal</i>	47
4.3.4.2.	<i>Modelación lineal to still</i>	49
4.3.4.3.	<i>Modelación exponencial</i>	50
4.3.4.4.	<i>Modelación esférico</i>	52
4.3.4.5.	<i>Modelación gaussiana</i>	54
4.4.	Comparación de los modelos de variogramas, en función del umbral y el menor error cuadrático medio	56
	CONCLUSIONES	59
	RECOMENDACIONES	61
	BIBLIOGRAFÍA	
	ANEXOS	

ÍNDICE DE TABLAS

Tabla 1-3 : Operacionalización de las variables.	29
Tabla 1-4 : Valores aplicados en la modelación de variogramas, respecto a la meseta, rango y efecto pepita.	56
Tabla 2-4 : Resultados de los modelos de variogramas, en función del umbral y el menor error cuadrático medio	57

ÍNDICE DE ECUACIONES

Ecuación 1-2 : Autocorrelación espacial a partir del índice de Moran.....	12
Ecuación 2-2 : Índice de Moran Local.....	12
Ecuación 3-2 : Conjunto de puntos en el espacio.....	15
Ecuación 4-2 : Estimación de interpolación espacial.....	16
Ecuación 5-2 : Interpolación distancia inversa ponderada.....	16
Ecuación 6-2 : Interpolación valor estimado de observaciones.....	17
Ecuación 7-2 : Interpolación Krigging.....	17
Ecuación 8-2 : Variograma generalizado.....	17
Ecuación 9-2 : Estimación de un variograma.....	17
Ecuación 10-2 : Semivariograma experimental.....	19
Ecuación 11-2 : Semivarianza.....	19
Ecuación 12-2 : Variograma ajustado.....	20
Ecuación 13-2 : Error cuadrático medio.....	21
Ecuación 14-2 : Fórmula de un variograma modelo lineal.....	22
Ecuación 15-2 : Fórmula de un variograma modelo esférico.....	22
Ecuación 16-2 : Fórmula de un variograma modelo exponencial.....	23
Ecuación 17-2 : Fórmula de un variograma modelo gaussiano.....	23

ÍNDICE DE ILUSTRACIONES

Ilustración 1–2 : Criterios de autocorrelación espacial vinculados a un I de Moran.....	12
Ilustración 2–2 : Los mapas y geometrías del espacio en Sistemas de Información Geográfica	14
Ilustración 3–2 : Parámetros básicos de un variograma	18
Ilustración 4–2 : Principales modelos de variogramas espaciales.....	21
Ilustración 1–4 : Base de datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021	31
Ilustración 2–4 : Diagrama de barras que presenta la edad de pacientes contagiados.....	32
Ilustración 3–4 : Diagrama de barras de las frecuencias absolutas de la variable sexo	33
Ilustración 4–4 : Diagrama de barras frecuencias absolutas de la variable Cantón.....	33
Ilustración 5–4 : Series de tiempo mensual de la cantidad de contagios	34
Ilustración 6–4 : Diagrama de barras de las frecuencias absolutas de la variable condición final paciente.....	35
Ilustración 7–4 : Diagramas de barras de pacientes muertos según edades	36
Ilustración 8–4 : : Diagrama de barras de pacientes muertos según cantones	37
Ilustración 9–4 : Diagrama de barras del número de contagios por sexo y cantón.....	38
Ilustración 10–4 : Representación espacial cantones de la Provincia de Chimborazo.....	39
Ilustración 11–4 : Representación de la capa vectorial de datos SARS-COV-2 (COVID-19)	39
Ilustración 12–4 : Mapa de cuantiles de los casos SARS-COV-2 (COVID-19)	40
Ilustración 13–4 : Centro medio simple, ponderado y desviación estándar de los datos.....	41
Ilustración 14–4 : Mapa de intensidad o de calor de los datos SARS-COV-2 (COVID-19).	42
Ilustración 15–4 : Resultado I de Moran de casos SARS-COV-2 (COVID-19)	43
Ilustración 16–4 : Mapa de clusteres, mapa de significancia (LISA), valor - p, e índice de morán	44
Ilustración 17–4 : Resultados del análisis de vecinos más próximos generado en Qgis	45
Ilustración 18–4 : Máxima distancia casos SARS-COV-2 (COVID-19) en Chimborazo	46
Ilustración 19–4 : Modelo lineal de los datos SARS-COV-2 (COVID-19).....	47
Ilustración 20–4 : Krigging modelo lineal de los datos SARS-COV-2 (COVID-19).....	48
Ilustración 21–4 : Modelo lineal to still de los datos SARS-COV-2 (COVID-19)	49

Ilustración 22-4 : Krigging modelo lineal to still de los datos SARS-COV-2 (COVID-19)	50
Ilustración 23-4 : Modelo exponencial de los datos SARS-COV-2 (COVID-19)	50
Ilustración 24-4 : Krigging modelo exponencial de los datos SARS-COV-2 (COVID-19).	51
Ilustración 25-4 : Modelo esférico de los datos SARS-COV-2 (COVID-19)	52
Ilustración 26-4 : Krigging modelo esférico de los datos SARS-COV-2 (COVID-19)	53
Ilustración 27-4 : Modelo gaussiano de los datos SARS-COV-2 (COVID-19)	54
Ilustración 28-4 : Krigging modelo gaussiano de los datos SARS-COV-2 (COVID-19)	55
Ilustración 29-4 : Validación cruzada del modelo de variograma esférico)	57

ÍNDICE DE ANEXOS

ANEXO A: PROYECTO QGIS

ANEXO B: PROYECTO GEODA

ANEXO C: PRUEBAS DE VARIOGRAMAS

ANEXO D: CÓDIGO DE R

RESUMEN

En Ecuador los registros de contagios COVID 19 se emitieron únicamente mediante boletines epidemiológicos no existían análisis estadísticos espaciales razón por la cual no se conocieron los focos reales de contagios, por lo tanto el objetivo de la presente investigación fue determinar el patrón espacial de la morbilidad del total de casos de SARS-CoV-2 (COVID-19), en la provincia de Chimborazo período 2020-2021 mediante un análisis exploratorio espacial, aplicando análisis estructural y modelación de variogramas, la metodología aplicada tuvo un enfoque cuantitativo, se utilizó un diseño no experimental con una población total de 16114 personas contagiadas, la información se obtuvo de la cuenta oficial del ministerio de salud pública, posteriormente se realizó un análisis estadístico espacial donde el cantón con un mayor número de contagios es Riobamba con un 67.4 % y los cantones que presentan un valor mínimo de contagios son Pallatanga, Penipe que no superan el 2 %, con respecto al análisis estructural se determinó que existe autocorrelación espacial dada la formación de clústeres, según la ubicación geográfica las zonas están asociadas con valores elevados y mínimos respectivamente clasificados en dos grupos, zonas más propensas de afección Riobamba, Colta, Guamate, Chambo, Guano, zonas menos propensas de afección Pallatanga, Cumandá, Chunchi, Alausí, por lo observado se realizó un modelado de variogramas donde el 99 % de los modelos tienen el mismo comportamiento con un RMSE del 3.056 y R^2 0.229, se valoró un ajuste adecuado en la representación gráfica, considerando un buen corte en el umbral donde el rango alcanza la varianza con tendencia regionalizada descendente, evidenciado en el modelo esférico, por lo tanto se concluye que el patrón de morbilidad de distribución espacial es agrupado dado que los eventos de contagios tienden a concentrarse en ciertas áreas presentando relaciones altas de proximidad geográfica en los diferentes cantones de la provincia de Chimborazo.

Palabras clave: <ESTADÍSTICA ESPACIAL>, <AUTOCORRELACIÓN>, <ÍNDICE DE MORÁN>, <INTERPOLACIÓN>, <INFORMACIÓN GEOGRÁFICA>, <VARIOGRAMAS>.

2238-DBRA-UPT-2023



ABSTRACT

In Ecuador, COVID 19 infection records were issued only through epidemiological bulletins there were no spatial statistical analyses, so the objective of this research was to determine the spatial pattern of morbidity of the total number of SARS-CoV-2 (COVID-19) cases in the province of Chimborazo for the period 2020-2021 by means of an exploratory spatial analysis, structural analysis and variogram modeling, the methodology applied had a quantitative approach, a non-experimental design was used with a total population of 16114 infected people, the information was obtained from the official account of the Ministry of Public Health, then a spatial statistical analysis was performed where the canton with the highest number of infections is Riobamba with 67.4 % and the cantons that have a minimum number of infections are Pallatanga, and Penipe which do not exceed 2 %, compared to the structural analysis determined that there is spatial autocorrelation due to the formation of clusters, depending on the geographic location, the zones are associated with high and low values respectively, classified in two groups, Riobamba, Colta, Guamote, Chambo, Guano, are the areas most prone to be affected, while Pallatanga, Cumandá, Chunchi, and Alausí are the areas least prone to be affected. Consequently, a variogram modeling was performed where 99 % of the models have the same behavior with an RMSE of 3.056 and R^2 0.229, an adequate adjustment was assessed in the graphical representation, considering a good cutoff at the threshold where the range reaches the variance with a downward regionalized trend, evidenced in the spherical model, therefore it is concluded that the spatial distribution morbidity pattern is grouped since the contagion events tend to concentrate in certain areas presenting high relationships of geographical proximity in the different cantons of the province of Chimborazo.

Keywords: <SPATIAL STATISTICS>, <AUTOCORRELATION>, <MORAN'S INDEX>, <INTERPOLATION>, <GEOGRAPHIC INFORMATION>, <INTERPOLATION>.



Edgar Mesias Jaramillo Moyano
0603497397

INTRODUCCIÓN

La propagación de un virus conocido como SARS-CoV-2 (COVID-19) generado en la ciudad China de Wuhan, provocó el inicio de una pandemia en todo el mundo, según la Organización Mundial de la Salud (OMS) en febrero de 2020 denominó al síndrome respiratorio agudo severo coronavirus 2 una enfermedad que transmite Covid-19 (OMS, 2020a:p.1), con el posterior número de casos positivos en todos los países, regiones y ciudades se presentó una distribución rápida de contagios dado que se desconocía el incremento impactante a nivel de transmisión del virus, denotando un aumento acelerado de casos sin registro (Dey, Chakraborty y Mukhopadhyay 2020: p.1).

El 30 de enero del 2020 la OMS declaró alerta sanitaria por el brote COVID-19 que provocó una emergencia de salud pública Internacional (ESPII), por causa del incremento alarmante del número de casos, revisado desde una expansión geográfica, se declara evento característico de emergencia dando cumplimiento al reglamento sanitario internacional (RSI), además se ha propagado a la Región de las Américas en EEUU, Brasil, Chile, Argentina desde el 26 de febrero del 2020, se notificó los primeros casos en Latinoamérica provocando consecuencias sanitarias, económicas y sociales por la emergencia de coronavirus SARS-CoV-2 (COVID-19), para el 7 de diciembre se denota un alarmante aumento de casos a nivel mundial (Dabanch 2021a: pp. 12-17).

Las actividades de investigación tuvieron un impacto sin precedentes a causa de la enfermedad por coronavirus, llegando a más de 89 millones de personas, provocando más de 1,8 millones de muertes (Sohrabi et al. 2021). En mayo de 2020, solo dos meses después de que el SARS-CoV-2 (COVID-19), fuera declarada por la Organización Mundial de la Salud (OMS) una pandemia, se confirmaron más de 22 millones de casos de Covid-19, Latinoamérica se volvió el epicentro acumulando más del 27 por ciento de muertes causadas por Covid-19 a nivel mundial, cada vez se presentaban nuevas opciones de aparición SARS-CoV-2 con altos indicadores de infección, las pruebas de diagnóstico sin registro presentaban inconsistencias para un país causando subestimación en el verdadero número de casos y muertes por el Covid-19 asociado con esta enfermedad (Anaya, Covarrubias et al., 2022).

Los equipos de gestión de incidentes locales y nacionales en todo el país activado por la OPS/OMS respondió directamente al Ministerio de Salud y otras autoridades de monitoreo nacionales, para ayudar al laboratorio, a los servicios de apoyo, a la atención médica, prevención de infecciones, gestión clínica, comunicación de riesgos y filas prioritarias. La organización ha detallado, publicado documentos técnicos y publicitarios, basado en evidencia para guiar la estrategia y la política de la pandemia nacional (OMS, 2020b:p.1).

El Sistema Nacional de Salud de Ecuador se vio severamente abrumado, la ciudad capital Quito, la provincia de Pichincha, la provincia de Guayas, Manabí, Azuay, Imbabura, el Oro, Tungurahua, Chimborazo entre otras, fueron afectadas por la pandemia , resultando en miles de casos positivos (Alava y Guevara, 2021).

En la provincia de Chimborazo, se demostró que existe una fuerte concentración espacial de infección en áreas urbanas y rurales con mayor densidad de población en los cantones Guano, Guamote, Penipe, Chambo, Colta, Pallatanga y Riobamba verificado al aplicar el análisis del centralidad y dispersión espacial a nivel territorial (Bolaños et al.,2022:pp. 1-19). A pesar de las noticias alentadoras acerca de la evolución de las vacunas no se ha controlado en su totalidad, pero se ha evitado un extremo de muertes en todo el mundo.

El artículo de investigación de (Barrantes, Sotela y Mayorga, 2020) titulado “Comportamiento espacial de la pandemia COVID-19 en Costa Rica durante los meses de marzo y abril de 2020 mediante un análisis de autocorrelación espacial” menciona que actualmente, con el uso de sistemas de información geográfica (SIG), se pueden hacer varios modelos cósmicos para estudiar el comportamiento espacial de un evento en particular. La imagen siempre ha sido una herramienta importante en el monitoreo de diversas enfermedades y su distribución espacial, con el uso de SIG, el análisis espacial adquiere una mayor relevancia, puesto que manipular datos geográficos y sus diversos modelos es mucho más fácil para la ejecución y que se puede realizar en el más corto tiempo posible además la cartografía es un medio de transmitir conocimientos de manera espacial.

Según el artículo de (Buffalo y Ryzewski, 2021) titulado “Dinámica territorial de pandemia COVID-19 en la provincia de Córdoba Argentina” Se analizó la distribución de casos en Argentina entre marzo y noviembre de 2020, el diseño metodológico del análisis espacial se realizó a partir de sistemas de información geográfica en función del uso de software libre (QGIS) asociado la posición, el tiempo de los casos Covid-19, el estado estadístico de la población, las características sociales y económicas de la población.

En el proyecto de titulación de (Escobar Rivera, 2022) titulado “Análisis espacial y temporal del Covid-19 en Chile” se ha evaluado la asociación espacial a través del índice Morán (local y global) entre junio de 2020 y junio de 2021, lo que indica la existencia de patrones espaciales y de tiempo de acuerdo a la ubicación geográfica de las personas.

En la presente investigación se estudia el análisis exploratorio de datos espaciales de contagios de SARS-CoV-2 (COVID-19), en la provincia de Chimborazo período 2020-2021, a través del cálculo y modelación de variogramas, se ha planteado este estudio por la falta de información y procesamiento estadístico de la dispersión de los contagios en esta provincia. En la justificación teórica se aplicarán técnicas y modelos estadísticos espaciales con la finalidad de producir conocimientos válidos, confiables y verídicos sobre la presencia de contagios, respecto al patrón de morbilidad y zonas más afectadas en la provincia de Chimborazo.

Metodológicamente podemos decir, que es una investigación cuantitativa debido a que se analiza la cantidad o número de contagios de SARS-CoV-2 (COVID-19), en cuanto a la manipulación de las variables es no experimental dado que se obtuvo la base de datos de fuentes ya establecida como es el MSP (Ministerio de salud pública). En el desarrollo de la propuesta se definió en base al cuadro de mando integral las actividades para atender las perspectivas de este trabajo de investigación, donde se genera un presupuesto para su ejecución.

La ejecución del análisis a partir del AEDE, para conocer la distribución espacial de una variable, depende de ciertos parámetros, en donde se debe tener en cuenta la formación de patrones a través de relaciones de proximidad entre los diferentes cantones, pueden influenciar otras características, es parte esencial del proceso, dado que se puede mencionar la influencia que tiene la proximidad de dos espacios en la diseminación de enfermedades o patrones en este caso de contagios covid. El método general proviene del I de Morán para obtener un índice cuyos valores toman un rango entre -1 y 1. Cuando el valor del índice es igual a 0 no existe autocorrelación y la distribución espacial de la variable en cuestión es un fenómeno aleatorio; -1 y 1 representan los valores de perfecta correlación negativa y positiva respectivamente. Además, este índice permite analizar el nivel de significancia del resultado (Saastamoinen et al., 2022).

La aplicación de varios métodos como son los de autocorrelación espacial LISA nos permitió identificar en cada ubicación el valor más similar a sus vecinos inmediatos, que sería un caso al azar, también podrá caracterizar los grupos en el sentido de valores tanto altos como bajos o existencia de atípicos, hay otro aspecto de estas estadísticas locales, se sabe, que están conectados a una estadística global en donde verificamos que la varianza es constante, por lo que la estructura del

vecino más cercano es la relación entre una medida de la fuerza de asociación entre el valor de una ubicación y el valor en otra localización aproximada comparando la ubicación promedio de sus vecinos, dado que mediante la generación de clústeres se identifica los grupos de contagios (Stein, 2022).

Análisis de patrones de puntos nos permite caracterizar la estructura espacial de un conjunto de puntos en función de parámetros como la densidad o las distancias entre puntos y su configuración en el espacio, otro grupo de procedimientos con componente estadística, los relativos a las técnicas de interpolación, la aplicación de la prueba de recuento de cuadrantes es uno de los primeros métodos estadísticos espaciales desarrollados. Se puede utilizar para comprobar si los puntos son completamente aleatorios desde el punto de vista espacial, es decir, si son uniformemente aleatorios o agrupados en toda el área de interés refiriéndonos al mapeo de geometrías puntuales, el análisis espacial dispone de excelentes posibilidades para el trabajo conjunto que permite unir a la ciencia pura con la ciencia aplicada a través del uso de instrumentos con bases sólidas teórico metodologías que permiten actuar en la resolución de problemáticas en las que se relaciona la sociedad con su medio ya sea natural o antrópico (Technologies, 2020).

Para la modelación de variogramas es importante recalcar que esta es una herramienta geoestadística, que permite analizar el comportamiento espacial de una variable sobre un área definida, la información georreferenciada se recoge en muchas aplicaciones y no utilizar esta información puede obstruir las características importantes del mecanismo de generación de datos en este caso se trabajó en QGIS. Técnicas como kriging simple y ordinario consideran una media constante de la variable regionalizada que es modelada, conocida y desconocida, respectivamente, además, asumen con respecto a esta media condiciones de estacionalidad y la existencia de una varianza obteniendo como resultado un variograma experimental que refleja la distancia máxima y la forma en que un punto tiene influencia sobre otro punto a diferentes distancias. El resultado de este análisis no puede ser aplicado directamente en los diferentes métodos de interpolación que lo ocupan como información base, por lo que una vez calculado se debe ajustar los valores de alcance, meseta, nugget para definir un modelo que represente el patrón espacial de los datos (Wulder y Boots, 2016).

CAPÍTULO I

1. PROBLEMA DE INVESTIGACIÓN

1.1. Planteamiento del problema

Se identifica la falta de información y procesamiento estadístico de la dispersión de los contagios de SARS-CoV-2 (COVID-19) en la provincia de Chimborazo, período 2020-2021, generada durante la pandemia; desde el último trimestre del año 2019 los datos registrados en el Ecuador son únicamente informativos emitidos mediante boletines epidemiológicos, infografías e informes de situación (SITREP) desde el ministerio de salud pública del Ecuador, sin embargo no existen análisis estadísticos espaciales de los casos registrados, no se proporciona la información del procesamiento de datos adecuado, por tanto, es imprescindible conocer las zonas con mayor número de contagios, permitiendo aplicar análisis exploratorio, estructural espacial y modelación de variogramas, siendo prioritarios los casos de morbilidad en la provincia de Chimborazo.

1.2. Limitaciones y delimitaciones

Una de las limitaciones que se presentan son las ubicaciones geográficas con latitud y longitud exactas de los lugares de cada uno de los pacientes de la provincia, puesto que únicamente se puede registrar de modo general por cantón, además, cabe recalcar que al inicio de la pandemia no había un manejo de datos adecuado por la magnitud de contagios, pero contamos con una base sólida de cifras de covid, de fuentes ya establecidas, posiblemente no en su totalidad pero confiables, dado que en Ecuador al inicio de la pandemia no se tomó la precaución de tratar los registros suficientes, dentro de las delimitaciones se considera a toda la población de personas que se contagiaron y murieron a causa del SARS-CoV-2 (COVID-19) en la provincia de Chimborazo en los años 2020, 2021.

1.3. Problema general de investigación

¿Cuál es el patrón de morbilidad total de los casos SARS-CoV-2 (COVID-19) existentes en la provincia de Chimborazo en el período 2020-2021?

1.4. Problemas específicos de investigación

- ¿Cuáles son las zonas con mayor número de pacientes con covid 19 que se pueden determinar a partir de los indicadores Locales de Autocorrelación Espacial (LISA)?
- ¿Cómo identificar la distribución espacial de los casos de morbilidad en la provincia de Chimborazo aplicando el mapeo de geometrías puntuales, aplicando análisis exploratorio espacial?
- ¿Cuáles son los patrones puntuales mediante el análisis estructural y modelamiento de variogramas?
- ¿Cómo comparar y determinar cuál es el mejor modelo de variograma, en función del umbral, y el menor error cuadrático medio?

1.5. Objetivos

1.5.1. *Objetivo general*

- Determinar el patrón espacial de la morbilidad del total de casos de SARS-CoV-2 (COVID-19), en la provincia de Chimborazo período 2020-2021 mediante un análisis exploratorio espacial, aplicando análisis estructural y el modelamiento de variogramas.

1.5.2. *Objetivos específicos*

- Determinar las zonas con mayor número de pacientes con covid 19 a partir de los indicadores Locales de Autocorrelación Espacial (LISA).
- Identificar la distribución espacial de los casos de morbilidad en la provincia de Chimborazo mediante el mapeo de geometrías puntuales, aplicando análisis exploratorio espacial.
- Análisis de patrones puntuales mediante el análisis estructural y modelamiento de variogramas.
- Comparar y determinar cuál es el mejor modelo de variograma, en función del umbral, y el menor error cuadrático medio.

1.6. Justificación

1.6.1. Justificación teórica

Para el desarrollo del presente proyecto de investigación se aplicarán técnicas y modelos estadísticos espaciales respecto al número de contagios y zonas más afectadas en la provincia de Chimborazo que permita conocer el número de contagios en determinadas zonas de la provincia, estas herramientas geoestadísticas de distribución espacial, permitirán visualizar diversos datos geográficos focalizados en las áreas de contagios del virus es decir las zonas de mayor concentración estimada de contagios y morbilidad.

1.6.2. Justificación metodológica

Con la finalidad de dar cumplimiento al estudio del patrón espacial de morbilidad del total de casos SARS-CoV- 2 (COVID-19), en la provincia de Chimborazo período 2020-2021 mediante un análisis exploratorio espacial, se considera determinar las zonas con mayor número de pacientes COVID a partir de los indicadores Locales de Autocorrelación espacial (LISA) dado que este tipo de estudios están aplicados a sistemas de información geográfica son de gran utilidad para conocer de manera más detallada los sitios de mayor aglomeración, acumulación de personas contagiadas, además usando el software QGIS se hace el mapeo de geometrías puntuales, análisis estructural y modelación de variogramas dado que radica la posibilidad de aplicar estrategias de una manera mucho más perspectiva, dinámica y visual en las zonas más acomplexadas de la presencia de este virus u otro tipo de enfermedades, en este estudio se trabaja con la base de datos publicados en la página del Ministerio de Salud Pública del Ecuador para conocer la distribución espacial de los casos de morbilidad, mediante QGIS, lo cual permite visualizar los resultados de manera geográfica, puntualizar detalles relevantes y trabajar con información manejable para su tratamiento, explicación y conocimiento general.

1.6.3. Justificación práctica

Debido a la falta de análisis exploratorio espacial, modelación de variogramas, análisis estructural, falta de reportes estadísticos, se ve la necesidad de desarrollar este estudio que permita conocer el patrón de morbilidad, se analizará los casos de contagios para encontrar el mejor modelo de variograma que explique el patrón real de morbilidad de la provincia, toda la información resultante permitirá analizar el comportamiento de los casos a través de documentación exploratoria espacial, dado que actualmente no se puede tratar a cabalidad toda la información de contagios, sigue

siendo un reto el manejo, obtención, manipulación y tratamiento de datos totales el cual se ha vuelto un problema no solo social, económico o sanitario sino también investigativo, por eso se pretende aplicar este tipo de estudio con un enfoque determinado de análisis, siendo optimo el trabajo funcional que se realice como iniciativa de trabajos investigativos espaciales debido a que actualmente son poco evidentes las zonas que presentan mayor impacto epidemiológico.

1.6.4. Hipótesis

Existen diferencias en el modelamiento de variogramas que influyen en el análisis estructural, la influencia del umbral versus el error cuadrático medio.

CAPÍTULO II

2. MARCO TEÓRICO

2.1. Referencias teóricas

2.1.1. *Análisis exploratorio de datos espaciales*

El análisis exploratorio de datos se define como el grupo de técnicas que describen distribuciones espaciales, identifican posiciones, reconocen esquemas de uno de los componentes más relevantes del (AEDE) que es el análisis gráfico. En combinación con las técnicas de análisis estadístico, permite que toda información sea posible y eficiente cuando se trabaja con grandes bases de datos al mismo tiempo, crea técnicas gráficas con la capacidad de trabajar con todas las observaciones o si lo desea analizar una cierta serie de datos (Sicilia y Rivera, 2017).

Diseñado para un tratamiento específico de datos geográficos, un análisis exploratorio de datos espaciales (AEDE) se usa para identificar relaciones sistemáticas entre variables o cuando no hay un conocimiento claro sobre su distribución en el espacio geográfico al intentar analizar los datos de georreferencia, puede asociarse con el hecho de que las relaciones pueden ser diferentes a través del espacio, tal vez porque existen variaciones espaciales en las actitudes o preferencias de la población o hay contextos administrativos, políticos u otros que producen diferentes respuestas al mismo estímulo, es que probablemente la representación de una o más variables relevantes para la explicación del modelo del análisis de exploración de datos espaciales está relacionado con la identificación de excepciones locales o tendencias generales, tanto en datos como en relaciones (Bohórquez y Ceballos, 2008: p.8).

Desde un software GIS se puede efectuar el análisis de los datos del espacio de exploración permite encontrar o conocer el comportamiento, anomalías en la estructura de datos, a través de distintas técnicas como el análisis global y local que permite conocer la asociación espacial uno de ellos muy aplicativo en el área es el índice de Moran posibilita conocer las existencia de autocorrelación espacial donde se tienen una visión más amplia de ubicación general de variable, muy aplicativo en la representación de redes, áreas, superficies (Marrugo, 2020: p.9).

2.1.2. Datos espaciales

Los datos espaciales pueden definirse como una observación de variables asociadas con el espacio geográfico, la naturaleza de los datos espaciales también permite un enfoque colaborativo porque son ampliamente utilizados en diferentes sectores en planes y estrategias ayudando a evaluar la efectividad de decisiones (Bohórquez y Ceballos, 2008).

Permiten una representación real mediante modelos dinámicos, esto depende de las ranuras temporales y espaciales, una gran cantidad de estudios han permitido aumentar la capacidad y las condiciones de desarrollo para herramientas que realizan un análisis espacial permitiendo modelar y comprender el desarrollo de fenómenos geográficos, se cree que las transformaciones del área geográfica dependen de los tiempos y fenómenos espaciales (Rodríguez y Risquet, 2020).

2.1.3. Sistemas de información geográficas

Estas son herramientas diseñadas para administrar, analizar y presentar la información que se obtiene de un territorio, ya que al principio tiene una gran capacidad para componer con grandes volúmenes de datos, las aplicaciones SIG se centraron en la geográfica cuantitativa y espacial. Tenga en cuenta que SIG le permite trabajar con datos y hacerlo útil, se ha convertido en una herramienta de trabajo fundamental para numerosas disciplinas: prevención de desastres naturales, enfermedades u otros (Gómez et al., 2022: pp.11-14).

2.1.4. Autocorrelación espacial

La autocorrelación del espacio es una correlación entre las variables de un valor cercano o similar definido como un espacio de escala del grado cuantitativo con alta similitud de la variable categoría entre los puntos adyacentes del espacio, el concepto de autocorrelación espacial hace referencia a comprender la distribución espacial de los fenómenos analizados, existe cuando el valor de una variable depende del valor de la misma variable, es fuerte si las variables están estrechamente relacionada entre sí, es débil o no existe si el valor de la variable se distribuye en un espacio al azar (Bravo y López, 2021: pp. 51-53).

Muchas interpretaciones se relacionan con el mapa, ciertos sistemas de información geográfica las herramientas de diagnóstico, el reemplazo de variables faltantes, el exceso de información, el proceso espacial, el desbordamiento espacial y el resultado de la delimitación de las unidades esto se puede determinar cuantitativamente con varios índices, incluida la relación de Geary y el coeficiente de Moran, estadísticamente más potente de dos medidas, si estos dos índices están correlacionados o conectados negativamente (Griffith, 2020).

2.1.5. Autocorrelación espacial (LISA)

La autocorrelación espacial es la concentración del valor de la variable en el mapa que refleja actividades geográficas similares a otros objetos o actividades, se produce una dependencia de una unidad espacial cuando el valor de una misma variable es frecuente en el análisis geográfico de las variables que terminan siendo dependientes, es común que los valores de las variables se correlacionen de manera espacial o sean dependientes (Bohórquez y Ceballos 2008: pp. 14-16).

2.1.6. Índices de autocorrelación espacial

Es importante especificar el índice de autocorrelación porque la escala del patrón espacial observado se resume en una sola cantidad (Celemín, 2009: p.3).

- Es positiva si las unidades espaciales vecinas tienen valores cercanos entre sí lo cual indica la tendencia a agrupar unidades espaciales.
- Es negativa si las unidades espaciales vecinas presentan valores muy disímiles por lo tanto indica una tendencia a la dispersión de las unidades espaciales.
- Sin autocorrelación esto sucede cuando no ocurre ninguna de las dos condiciones anteriores. Por lo tanto, los valores de las unidades espaciales vecinas muestran valores procedentes en forma aleatoria (Celemín, 2009: p.3).

2.1.6.1. Índices global de moran

Posibilita encontrar una dependencia espacial de las variables dadas en diferentes zonas de lo contrario mostrará una distribución aleatoria espacial, las unidades territoriales se miden con análisis de valores entre -1 y +1 los positivos indican análisis similares, en cambio los valores negativos indican análisis distintos independientemente si son valores altos o bajos, si el valor de este índice es cercano a cero indica que no hay auto correlación (Vilca, 2019: p.9).

Autocorrelación espacial a partir del índice de Morán

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n W_{ij}(x_i - \bar{x})^2} \quad (1-2)$$

Donde:

n es el número de áreas

x_i valor de la variable considerada en el área i

i, x_j es el valor de la misma variable en otra ubicación donde i no es igual a j

\bar{x} es el número medio de la variable en la región de estudio

w_{ij} es el peso de la matriz W aplicado a la comparación entre las ubicaciones i, j (Vilca, 2019).

El dominio del I de Morán se vincula a los patrones espaciales permitiendo identificar la existencia de autocorrelación LISA que describe el comportamiento de las unidades de análisis, a partir del criterio de vecindad se puede generar los clústeres, estos varían según la distribución de los fenómenos agrupados en el espacio geográfico.

Patrones espaciales		
Autocorrelación espacial positiva	Autocorrelación espacial negativa	Ausencia de autocorrelación
Patrón clúster	Patrón disperso	Patrón aleatorio
Índice de Morán teórico		
$I > 0$	$I < 0$	$I = 0$
Índice de Morán en la práctica		
$I > 0,35$	$I < -0,35$	$-0,35 \leq I \leq 0,35$

Ilustración 1–2: Criterios de autocorrelación espacial vinculados a un I de Moran

Fuente: (Siabato et al. 2019)

Cuando $I = 0$, en el caso teórico, existe una distribución aleatoria que varía dependiendo de la distribución de unidades, se sugiere pensar en la aleatoriedad del fenómeno aplicado al experimento cuando el índice está en el rango indicado en la ilustración 1-2. Los valores más altos y más bajos deberían dar como resultado patrones de tipo clúster o patrones dispersos (Siabato et al. 2019).

2.1.6.2. Índice local de autocorrelación espacial (LISA)

Esto permite la identificación de los patrones de la asociación de espacio local, evalúa el impacto de las ubicaciones individuales en las estadísticas globales a partir del índice Moran y amplían la capacidad de visualización de los valores analizados mediante el uso de sistemas de información geográfica expresando con un valor significativo los indicadores estadísticos de la asociación que advierte la presencia de puntos espaciales de alta temperatura mismos que dependen de la asociación

significativa de los datos estadísticos que se ha analizado. Usando el software QGIS se generan dos tipos de mapas, el primer mapa se denomina mapa significativo valor P, y en base al procedimiento de aleatorización, la probabilidad de que la relación ocurra aleatoriamente de cada unidad espacial, el segundo mapa se llama grupo o clúster esto indica que cada unidad espacial es diferente de la unidad espacial adyacente (Bucheli, 2019).

El análisis de la auto correlación o LISA, el indicador local de la asociación espacial es un índice de Moran calculado para cada región se puede usar el índice Moran local para encontrar contribuciones globales y detectar los valores extremos y hacerlos en un LISA. Oficialmente, el índice Moran local es el siguiente:

$$I_i = \frac{z_i}{\frac{\sum_j z_j^2}{N_j}} \sum_j W_{ij} z_j \quad (2-2)$$

Donde:

z_i es el valor de la variable proporcionada a la región i .

N_j es el conjunto de vecinos a i .

Los indicadores locales (LISA) consideran escenarios específicos (de ahí el nombre), lo que permite identificar las subáreas en las que se agrupa o dispersa el fenómeno. Aunque los cálculos básicos de LISA también tienen en cuenta los promedios globales, estos indicadores proporcionan una medida de autocorrelación para cada unidad de análisis permitiendo identificar el grado de agrupamiento o distribución de cada unidad en relación con sus vecinos o con respecto a las unidades del criterio con un valor eminente, positivo o negativo y significativo del estadístico da lugar a la presencia de un clúster en torno a la región i de valores similares elevados o bajos (Viramontes y Hernández, 2021).

2.1.7. Distribución espacial

Se considera como la cantidad de unidades del mismo tipo se distribuye de cierta manera sobre el área geográfica, pueden ser puntos, líneas, polígonos con diferentes atributos asociados en el sistema vectorial o ubicaciones, se puede representar las áreas en el sistema ráster, las diferentes características medidas geográficamente son poco probables que tengan una distribución homogénea es común que presenten variaciones dado que mediante la distribución espacial se brinda la oportunidad de analizar una diferenciación inicial en el espacio geográfico son poco relativas las variables individuales del área de investigación (Baxendale y Buzai, 2011: pp.8-9).

2.1.8. Base cartográfica espacial

La base cartográfica representa una serie de datos geográficos que admite los sistemas de información geográfica (SIG) y se puede utilizar en todo tipo de aplicaciones, propósito y expresiones de superficie plana. El método analógico convencional para fabricar mapas ha sido reemplazado por un sistema digital que puede producir un mapa interactivo dinámico que puede operarse de manera digital (Moreno y López 2015: pp.4-23).

La composición para la creación de mapas apropiados se basa en un SIG haciendo posible que sea tratada la información determinada al uso e investigación del usuario, el elemento más relevante es una expresión real, no una realidad misma, generalmente significa un volumen, por lo tanto, puede representar un cambio de tres dimensiones, la escala del mapa es la relación entre la representación espacial con la realidad, longitud y latitud hace referencia a las ubicaciones geográficas lo cual permite conocer la distribución espacial de dicho espacio, área geográfica. (Dominguez 2000: pp.9-11).

2.1.9. Geometrías puntuales

Se identifica como una de las técnicas de ubicación de puntos para elementos poco relativos y separados mediante un definido tiempo o ubicación de dimensiones ciertamente puntuales para fines de representación de un mapa (Gil y Beltrán, 2020).

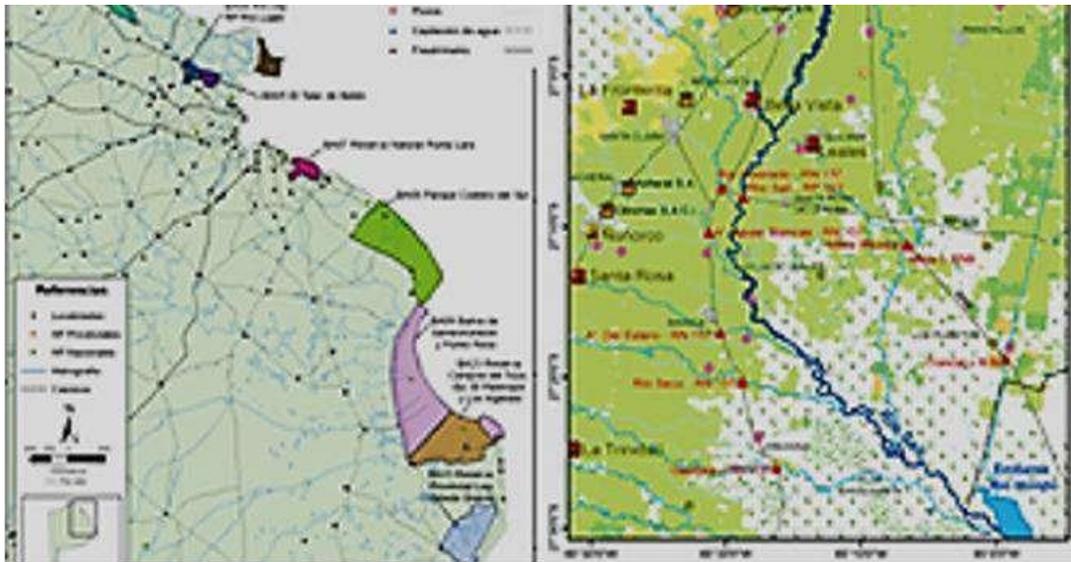


Ilustración 2–2: Los mapas y geometrías del espacio en Sistemas de Información Geográfica

Fuente: (Louis, 2017)

2.1.10. Patrones puntuales

Es una elaboración completa de un procedimiento puntual $X \subset R^d$ en un conjunto unido de puntos en el espacio:

$$X = \{X_1, X_2, \dots, X_n\} \quad (3-2)$$

Donde:

X_i son las ubicaciones

Mención d dimensión

Este análisis se fundamenta en realizar una gráfica sobre el espacio todo dependerá de la distribución del patrón de los puntos lo cual genera un evento contable en el espacio puede ser aleatorio, agrupado, disperso (Moncada y Fuentes 2018: pp.6-7).

2.1.11. Análisis estructural espacial

Esta fase es responsable de la caracterización del espacio temporal y estructura espacial de una propiedad o fenómeno regionalizado. Es el proceso dentro del marco desde el cual se conserva un modelo para la función aleatoria examinada. Consiste en la estimación y modelado de una función esto refleja la correlación espacial y temporal de las variables regionalizadas a partir de la adquisición justificada de la hipótesis, más suficiente sobre su variabilidad. Esta variabilidad se lleva a cabo mediante un análisis de los puntos de población objeto de estudio que permite la definición del tipo de estructura espacial y espacio tiempo que organiza la ubicación y un análisis de dependencia o autocorrelación, el tiempo espaciales de la distribución que hace posible medir el grado la similitud de una variable regionalizada del tipo cuantitativo entre puntos o eventos vecinos en un estudio de muestra (Alvarado, 2011).

2.1.12. Interpolación espacial

Consiste en usar puntos con un valor conocido, también llamado puntos control, para estimar las variables en un lugar donde se desconoce; también se considere cómo cambiar la información a tiempo a la información de superficie, con el objetivo de combinarla con otros datos para facilitar el análisis y espacio de modelado (Niño 2011:pp.2-4).

La interpolación espacial es un problema clásico de estimación de una función:

$$F(x_p) = \sum_{i=1}^m W_i * F(x_i) \quad (4-2)$$

Donde:

$F(x)$ en un punto X

p de un plano a partir de valores conocidos de $F X_i$

en un cierto número m de puntos regionales

W_i es la ponderación de los valores de cada punto regional (Gutiérrez y López et al., 2011a:p.2).

2.1.12.1. Interpolación (IDW) distancia inversa ponderada

El método no geoestadístico ampliamente utilizado para interpolar ocurre a partir de los pesos ponderados según la distancia (IDW), como primer punto se calcula el promedio ponderado para obtener el estimador de posición, el peso asignado a la observación se calcula de acuerdo con la distancia a la ubicación de interpolación (Porras 2017:pp.1-14).

El valor dado por la interpolación normalmente es igual a la observada, similar al método de krigging se usa varias cosas como un efecto Pepita pequeño o nulo en contraste con el método que considera la distancia a la posición de predicción e ignoran la composición espacial de la observación. Este hecho puede tener efectos innecesarios si la posición de la observación está muy agrupada, otra diferencia es garantizar que el método tenga un peso de 0 a 1 como resultado, se puede obtener un valor de interpolación que no estén dentro del rango del valor (Velásquez, Alberto 2017:pp.1-14).

$$Z^*(S_0) = \frac{\sum_{i=1}^n W(S_i)Z(S_i)}{\sum_{i=1}^n w(S_i)} \quad (5-2)$$

Donde:

$Z^*(S_0)$ es el valor estimado en el punto interpolado S_0

N es el número de observaciones vecinas usadas para la estimación

S_0 son los pesos asignados de acuerdo con la distancia

$Z(S_i)$ es el valor observado de las estimaciones vecinas

2.1.12.2. Interpolación krigging

El método krigging es óptimo en el sentido de interpolación y estadísticas de extrapolación de evaluación más preciso viceversa a todos los demás métodos, esto nos permite calcule el error de evaluar las variables entre puntos basados en la distancia que lo separa se designa como una condición que es indispensable para usar el promedio y variante de la función dada que puede ser estacionaria o no estas estadísticas no dependen de la posición del punto de medición, sino solo en la distancia que la separa (Gutiérrez y López et al., 2011a).

$$\hat{Z}(X_0) = \lambda_1 Z(X_1) + \lambda_2 Z(X_2) + \dots + \lambda_n Z(X_n) \quad (6-2)$$

$$\hat{Z}(S_0) = \sum_{i=1}^n \lambda_i Z(S_i) \quad (7-2)$$

Donde:

$\hat{Z}(X_0)$ Es el valor que se estima a través del punto interpolado S_0

n es el número de observaciones vecinas usadas para el valor de estimación

λ_i es el peso que se da al valor observado $Z(S_i)$ en las proximidades del valor S_0

2.1.13. Variograma

Este es un valor determinado de estimación de varianza de la dispersión de la población, por lo que la población debe tener una tendencia en el estacionariedad mismo que está asociado con la dirección y la distancia h , es la principal herramienta que admite los métodos de Krigging y le permite cuantificar la variación del fenómeno regionalizado en el espacio los valores en el espacio o el tiempo se colocan en correlación, se puede decir que el variograma es la varianza menos la covarianza dependiendo de la distancia entre los puntos (Calderón, Cisneros y Guerrero, 2004).

Se define como:

$$2\gamma(h) = \frac{1}{2} \text{var}[z(x+h) - z(x)] = \frac{1}{2} E[z(x+h) - z(x)] \quad (8-2)$$

Puede ser estimado a partir de:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (x_i - y_i)^2 \quad (9-2)$$

Donde:

$\gamma(h)$ es el variograma teórico

h distancia que separa los puntos muestreados

x_i posición de datos en donde $h = 0$

y_i es $x_i + h$

$n(h)$ es el número de puntos x_i, y_i separados una distancia $h = x_i - y_i$

Las coordenadas (x,y) representan un vector x de una variable aleatoria, denotada comúnmente en geoestadística como $Z(x)$ (Gutiérrez y López et al. 2011b).

Admite describir el comportamiento de la variación espacial de un fenómeno mediante un análisis exploratorio y con base en el modelado del variograma misma que da respuesta al proceso existente de puntos en el espacio se busca estructura en el comportamiento del fenómeno. Posteriormente, el modelado del variograma se ajusta a un modelo teórico que será utilizado en el modelo de interpolación de la variable, originariamente fue creada para el desarrollo del estudio de variables regionalizadas su aplicabilidad esta más centrada en el espacio temporal (Calderón, Cisneros y Guerrero, 2004).

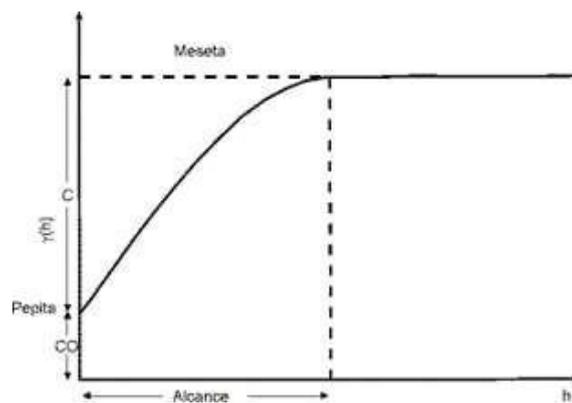


Ilustración 3–2: Parámetros básicos de un variograma

Fuente: (Giraldo, 2002)

Meseta: Este es el límite superior del semivariograma, también se puede definir como el límite del variograma a medida que la distancia h tiende a infinito, la meseta puede o no ser finita los semivariogramas con una meseta finita satisfacen la hipótesis de estacionariedad fuerte, se denota por $C1$ o $(C0 + C1)$ cuando el elemento es distinto de cero. Si una cláusula se interpreta como un error en medidas, esto explica por qué se supone que en un modelo que explica bien la realidad, la pepita no debe representar más del 50 % de la meseta solo un proceso con fuerte estacionariedad presentan un valor sill o meseta y esto corresponde a la asíntota $C(1)$ (Giraldo 2002) .

Nugget: Se denota C_0 también llamado efecto pepita y representa el punto de discontinuidad del semivariograma, esto puede deberse a errores de medición de la variable o escala, en algunos casos esto puede indicar qué parte de la estructura espacial se encuentra a distancias menores que las observadas (Cabrera Vargas 2017).

Rango: Es la distancia en la cual el variograma se estabiliza y muestra las relaciones espaciales representa la distancia dentro de la cual $C(h)$ algunos procesos alcanzan una correlación cero solo asintóticamente cuando h es infinito, mientras que otros tienen gama final hay modelos de semivariogramas en los que no existe una distancia finita para el cual las dos medidas son independientes, por lo que es necesario determinar el rango es decir, la distancia para la cual el variograma alcanza el 95 % de la varianza. Cuanto más pequeño es el rango, más cerca está del modelo de independencia espacial (Cabrera Vargas 2017).

2.1.14. *Semivariograma*

Describe la variabilidad espacial de un evento de interés, su pertinente aplicación ha sido ampliada al estudio de variables referentes a una serie de tiempo, esta denominada como semivariograma experimental o variograma en donde las muestras tomadas normalmente están a una distancia constante, la gráfica relacionada a intervalos de tiempo en su mayoría no representa todos los puntos posibles, debido a que únicamente representa los intervalos aproximados (Calderón, Cisneros y Guerrero, 2004).

$$\text{var}[Z(x+h)-Z(x)] = 2\gamma(h) \quad (10-2)$$

Donde:

A $2\gamma(h)$ se le llama variograma
 $\gamma(h)$, semivariograma.

Es común utiliza el término variograma o semivariograma para referirse a la función $\gamma(h)$, aunque estrictamente hablando la función $\gamma(h)$ es el semivariograma.

Se demuestra que la semivarianza de los incrementos $Z(x+h)-Z(x)$ se relaciona con la covarianza por medio de la siguiente relación:

$$Y(h) = \frac{1}{2}E [z(x+h) - z(x)]^2 = c(0) - c(h) \quad (11-2)$$

Siendo $c(0)$ y $c(h)$ las covarianzas para $h=0$ y h genérica

Para obtener un modelo de semivariograma a partir de los datos observados se realiza un semivariograma experimental definido como:

$$y(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (x_i - y_i)^2 \quad (12-2)$$

Donde:

h es el valor promedio de la diferencia al cuadrado de los valores de la propiedad en dos puntos separados por una distancia

y es independiente de la localización x_i

$y(h)$ depende del módulo y de la dirección del vector h

Puede ser ajustado para cuantificar el grado y variación temporal existen varios modelos, el ajuste de la misma permite extraer varios parámetros que son los que van a ser usados en la interpolación de krigeados los parámetros sumamente indispensables son el efecto pepita, rango, proporción de las varianzas, el rango indica la distancia que existe entre los puntos de una variable mostrando si son independientes el efecto pepita también se conoce como varianza error puesto que la varianza entre los distintos puntos separados nunca deben ser cero. La varianza umbral a menudo se expresa en porcentaje el ajuste de una función específica permite presentar ciertos parámetros que definen el grado y escala de variación espacial (Cabrero Ortega y García Pérez 2022).

2.1.15. Función umbral

El análisis umbral es una técnica estadística que identifica zonas o puntos específicos del problema a estudiar, es decir, es un valor que alcanzado por el variograma en el rango puede cubrir o sobrepasar la zona umbral de áreas geográficas que se han definido, el umbral se puede definir por los valores $(c_0 + c)$ denominado como la varianza muestral (Casco 2017)..

2.1.16. Error cuadrático medio

Se aplica especialmente para el ajuste de modelos, selección de parámetros óptimos de un modelo, ratificación de modelos, comparaciones de modelos entre varios modelos en competencia y valoraciones de pronóstico (Karunasingha 2022).

La raíz cuadrada de los errores de predicción es un valor promedio debe ser lo más pequeño posible, dicha medida manifiesta la media entre la diferencia cuadrática entre la predicción del modelo y el valor buscado esto determina la precisión de la predicción y el valor se aproxima a la desviación promedio de los valores medidos (Vieira et al. 2015).

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - f_i)^2} = \sqrt{\frac{SSE}{N}} \quad (13-2)$$

Donde:

y_i es el resultado actual en el tiempo t .

f_i es el pronóstico de valor en el tiempo t .

SSE es la suma del error cuadrático

Su raíz cuadrada toma valores pequeños si las predicciones son precisas (Karunasingha 2022).

2.1.17. Modelación de variogramas

La modelización de un grupo de variogramas es una tarea más compleja ya que se busca el ajuste correcto de variogramas, siendo importante el número de variables aleatorias analizadas. Normalmente, el ajuste de los variogramas se consigue usando múltiples componentes para cada estructura específica en una escala de variación espacial (Porras 2017).

Es un paso crítico para la mayoría de los métodos geoestadísticos. Sin embargo, la mayoría de las soluciones predominantes basadas en variogramas están diseñadas sin considerar suficientemente el efecto del proceso de interpolación en su aplicación (Zapata Bedoya, Walteros Acero y Mercado 2022).

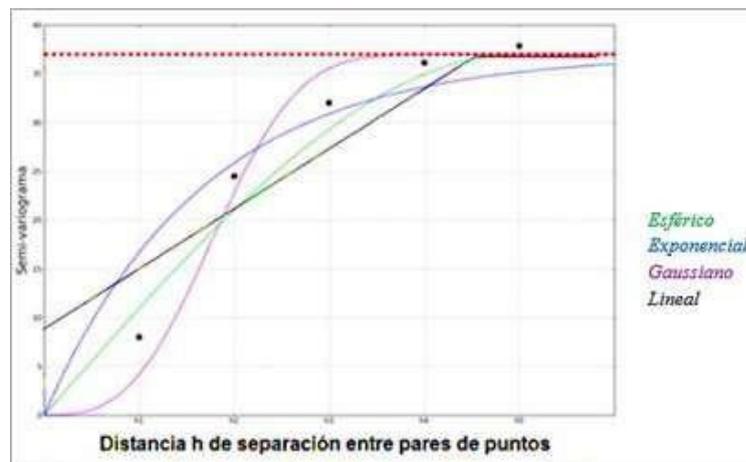


Ilustración 4-2: Principales modelos de variogramas espaciales

Fuente: (Sierra, 2014)

Optar por el modelo teórico más adecuado experimentalmente, ha llevado a la definición de varias funciones matemáticas seleccionadas según mayor o mejor similitud con la distribución de puntos del variograma (Sierra 2014)

En la ilustración 4-2 se muestran algunos de los modelos son el modelo esférico, el gaussiano, el exponencial y lineal con meseta como el grado de dependencia espacial puede no ser el mismo para todos, antes de que un modelo teórico pueda aceptarse como válido, debe tener en cuenta las direcciones espaciales corresponde a la variación máxima al examinar intervalos en las distancias h de la separación de los pares de puntos y ángulos por campos de estudio. La bondad de ajuste del modelo se determina mediante el método de validación cruzada. Esto permite compararlo con otros modelos de variograma para adoptar que coincida mejor con los datos. Se pueden usar diferentes estadísticas para comparar modelos, pero este es el caso el más común es utilizar la suma de los errores al cuadrado (RSS), su media (RSS/n), o la raíz cuadrada de la misma (RMSE) (Sierra 2014).

2.1.17.1. *Modelo lineal*

Representado por la ecuación de la siguiente forma:

$$y(h) = a(h) + b \quad (14-2)$$

Este modelo no alcanza el umbral por tal razón no requiere un rango definido, aumenta según la distancia de (h), no tiene un valor sill, se puede modelar hasta fenómenos que representan una capacidad infinita de dispersión los datos no están establecidos por eso no son independientes a medida que se incrementa la separación entre ellos, este modelo cambia a medida que se refleja una tendencia (Leiva González 2014).

2.1.17.2. *Modelo esférico*

Este dado por la ecuación:

$$y(h) = c \left[1 - \frac{3h}{2a} + \frac{h^3}{2a^3} \right], \quad h \leq a \quad (15-2)$$

Donde:

c es la meseta sill

a es el rango

Constituye una muestra en el crecimiento casi lineal hasta una distancia definida, representa una tendencia creciente curvilínea hasta llegar al umbral, el modelo alcanza un valor sill dependiendo la distancia del rango, a diferencia del lineal presenta un umbral (c) y un rango (a) lo que para estos valores significa la independencia entre si mismos la variación se estabiliza a medida que se cambia el valor h de separación (Leiva González 2014).

2.1.17.3. *Modelo exponencial*

Se muestra la siguiente ecuación:

$$y(h) = c \left(1 - \exp\left(-\frac{3h}{a}\right) \right) \quad (16-2)$$

Donde:

c es la meseta sill

a es el rango

Se observa que a partir del origen se levanta mas lentamente que el modelo esférico sin alcanzar completamente el umbral, sin embargo, para algunos otros conceptos se considera una línea recta tangente al origen, como es esférico muestra un crecimiento lineal al origen sin embargo puede crecer muy rápido a medida que definimos un rango se estabiliza paulatinamente (Leiva González 2014).

2.1.17.4. *Modelo gaussiano*

Dado por la ecuación:

$$y(h) = c(1 - e^{-(h/a)^2}) \quad (17-2)$$

Donde:

c es la meseta sill

a es el rango

Comúnmente muestra el comportamiento parabólico cerca del origen constituye un único modelo estacionario, se utiliza para generar modelos de fenómenos continuos, alcanza la meseta y se considera que el rango es la distancia en el que el variograma constituye un punto de curvatura (Leiva González 2014).

2.2. Referencias Conceptuales

2.2.1. Indicadores de morbilidad

Tienen el objetivo de medir la ocurrencia, enfermedad, lesión o discapacidad en la población, estos indicadores se pueden expresar al medir la frecuencia de pacientes dentro de un grupo específico, es importante confiar en la observación directa, otras encuestas, la notificación de eventos en el sistema de monitoreo y notificación de la enfermedad, sobre todo, pacientes ambulatorios, registros específicos, etc (OPS/OMS, 2018, p.1).

Entre los factores que pueden afectar la precisión de medición están:

La calidad y cantidad de datos

La eficacia de instrumentos de medición El

riesgo o gravedad de padecimiento

Confidencialidad de la información

Sistemas de registro

Los indicadores miden la frecuencia de problemas de salud específico u otros campos de estudio, las fuentes de datos suelen ser registros de las enfermedades esencialmente crónicas, desde su evolución a largo plazo, por tanto, requieren el monitoreo de las etapas clínicas, es preferible tener registros de enfermedades (OPS/OMS, 2018, p.1).

2.2.2. Sarc-cov-2 (COVID-19)

El SARC-COV-2 es el virus y COVID-19 la enfermedad conocido por la rápida propagación e infección humana, se menciona que este virus pudo haber surgido accidentalmente en un laboratorio aunque esta afirmación no tiene apoyo científico, pero confunde la opinión pública, la primera observación es su investigación estructural y bioquímica indica que SARS-CoV-2 tiene adaptación óptima para introducirse como receptor en el ser humano que participa en membranas celulares, otra noción indica que puede ser origen biológicamente de otras especies, la adaptación de los virus en los animales que están naturalmente infectados son los resultados de múltiples eventos genéticos a través de elecciones naturales, en los humanos se ha optimizado la capacidad de infección y propagación (Lopez, 2020, p.1).

2.2.3. Contagios

Se considera como la transmisión infecciosa por ciertos agentes biológicos, la incidencia de esta tiene un mecanismo calmoso que realiza su trabajo en provecho de las malas medidas de protección o de personas que están usualmente expuestas a la infección, estos problemas pueden ser temporales, permanentes o incluso de muerte, el impacto de la presencia de un virus en el cuerpo es arriesgado por los efectos de causa inmediato de una afección específica (Barroso, 2021).

2.2.4. Contagio de una enfermedad

Se produce si un individuo entra en contacto directo con una persona que padece la enfermedad o malestar en este caso por coronavirus puede causar problemas graves que van desde resfriados hasta neumonía severa, en casos más graves puede haber, insuficiencia renal y otras complicaciones, principalmente en pacientes mayores y pacientes que sufren algún otro problema de salud, si el problema se alarga esto puede llevarnos a una pandemia o epidemia, el distanciamiento social y el aislamiento de las poblaciones infectadas pueden contrarrestar una pandemia es necesario evitar una cumbre epidémica que pueda desbordar los servicios de salud, provocando muertes hasta por otras causas. ((Llor y Moragas, 2020, p.1).

2.2.5. Variables geográficas

Una variable hace referencia a una variación y puede ser modificada a partir de esto se puede desarrollar experimentos, formular un problema o al indicar la base teórica. Se llama una variable geográfica a todo lo que tiene sus propias características que lo distinguen del resto que están sujetos a cambios o modificaciones, y podemos estudiar, controlar o medir en el estudio desde un punto de vista estadístico matemático engloban variables cualitativas y cuantitativas discretas o continuas que permiten hacer uso de pruebas estadísticas apropiadas (Amiel Pérez 2007).

2.2.6. Latitud

Es la distancia angular entre un punto en la superficie de la tierra formando en arco meridiano desde Ecuador, es su línea base desde el punto donde se encuentra el observador se mide de 0° Ecuador hasta 90° N y 90°S respectivamente al polo norte y polo sur, es latitud sur si se muestra en el hemisferio sur y latitud norte si se muestra en el hemisferio norte, se mide de 0° a 90° indicados en grados sexagesimales (Ibañez, 2010).

2.2.7. Longitud

Es la distancia que existe entre un punto cualquiera marcan una posición estas líneas de longitud se conocen como meridianos de Greenwich, recorren de polo a polo cruzando el Ecuador en ángulos rectos se mide de 0° a 180° representado en grados sexagesimales la línea de longitud de 0 grados se conoce como el Primer Meridiano y divide el mundo en el hemisferio oriental y el hemisferio occidental, todos los puntos ubicados sobre el mismo meridiano tienen la misma longitud los polos norte y sur no tienen longitud (Ibañez, 2010).

A partir de las coordenadas geográficas se puede situar el punto exacto donde nos encontramos marcando la ubicación correspondiente a la superficie terrestre (Zapata Bedoya, Walteros Acero y Mercado, 2022).

2.2.8. Unidades espaciales

Son datos disponibles en los distintos sistemas de información hace referencia a ciertas unidades que permiten realizar un análisis geográfico de un cierto campo de estudio, en donde se consideran variables específicas al posterior análisis a partir de su elección es posible la visualización de indicadores, causas o resultantes en un mapa o base SIG (Barcellos, 2003).

2.2.9. Áreas geográficas en un (SIG)

Son indispensables para determinar los patrones de distribución geográfica para ciertos estudios como enfermedades, ciencias de la salud, entre otras, ha aumentado significativamente el modelado espacial en base a áreas geográficas se usa directamente para comprender las diferencias en la distribución espacial y su relación con ciertos factores hoy en día son muy utilizadas las herramienta SIG de suma importancia en la investigación espacial convirtiéndose en un complemento de estudios geográficos (Juárez y Montes Galbán, 2019).

El modelado espacial en QGIS se utiliza directamente para comprender las diferencias en la distribución espacial y su relación con ciertos factores es actualmente una importante herramienta en la investigación como la aplicación de indicadores locales de asociación espacial (LISA) en el SIG y otros procedimientos que son muy útiles, se puede describir y visualizar distribuciones espaciales, áreas geográficas (Andersen et al., 2020).

CAPÍTULO III

3. MARCO METODOLÓGICO

3.1. Enfoque de investigación

El presente trabajo de investigación es de tipo cuantitativo dado que en términos de operación de variables se analiza la cantidad o número de contagios de SARS-CoV- 2 (COVID-19), esto se aplica de acuerdo con el sujeto de investigación, según el objeto de estudio es aplicada porque se requiere modelar y comparar los estudios mediante dos técnicas, análisis estructural y modelamiento de variograma.

3.2. Nivel de Investigación

Según el nivel de profundización es exploratorio dado que se busca conocer el patrón espacial de morbilidad a través de la comparación de modelos de variogramas en función del umbral y el menor error cuadrático medio, según la inferencia es inductiva debido a que se quiere determinar el patrón espacial de morbilidad del total de casos de Covid-19 registrados en Chimborazo.

3.3. Diseño de investigación

El diseño de la investigación es de tipo no experimental, debido a que no se realiza ningún tipo de experimentos, aquí se fundamenta el estudio de las variables manejadas tal como se proporciona en la base de datos. La investigación que se va a aplicar es cuantitativa debido a que se va a analizar el patrón de casos Covid-19, este enfoque nos permitirá tener una mejor comprensión del problema de investigación y de analizar las zonas que más registran casos confirmados en la provincia de Chimborazo.

3.3.1. *Según la manipulación o no de la variable independiente*

Se utilizará el método de investigación cuantitativo, debido a que es una base de datos de una fuente ya establecida se denota que en cuanto a la manipulación de variables es no experimental dado que se tiene una base de datos descargada de la pagina web oficial del Ministerio de Salud Pública.

3.3.2. *Según las intervenciones en el trabajo de campo*

Según el periodo es temporal longitudinal puesto que se observa un evento de contagios SARS-CoV-2 (COVID-19), con los datos de la provincia de Chimborazo durante un tiempo definido.

3.4. Población de estudio

La población de estudio está conformada por los contagios SARS-CoV-2 (COVID-19) registrados entre los años 2020 -2021 en la provincia de Chimborazo, se trabajará con el total de casos registrados en la base de datos de la cuenta oficial del Ministerio de Salud Pública del Ecuador por lo cual no se aplicó ninguna muestra.

3.4.1. *Localización de estudio*

El proyecto de investigación se lo llevara a cabo en la provincia de Chimborazo.

3.4.2. *Técnicas de recolección de datos*

Los datos fueron recopilados a través de la página del ministerio de salud Pública del Ecuador, los cuales fueron almacenados y publicados en la página web de esta entidad.

3.5. Métodos, técnicas e instrumentos de investigación

Se aplicarán técnicas y modelos estadísticos de los datos SARS-CoV- 2 (COVID-19) en el período 2020-2021, respecto al número de contagios de la provincia de Chimborazo. Empleando técnicas espaciales en el análisis exploratorio, indicadores de autocorrelación (LISA), análisis estructural, modelización de variogramas, para de esta manera conocer el patrón de morbilidad en la provincia de Chimborazo.

3.6. Operacionalización de las variables

Tabla 1–3: Operacionalización de las variables.

Nombre de la variable	Descripción	Tipo de variable	Escala de medición	Categoría o intervalo
Casos COVID-19	Enfermedad infecciosa provocada por el virus SARS-CoV-2 (COVID-19)	Cualitativa	Nominal	Confirmado Descartado
Año notificación	Período temporal que tiene una duración de 12 meses, permitiendo dar mención de hechos suscitados.	Cualitativa	Nominal	(2020 ,2021)
Mes notificación	Es el segmento de la división de un año que representa uno de los 12 conocidos.	Cualitativa	Nominal	Enero Febrero Marzo Abril Mayo Junio Julio Agosto Septiembre Octubre Noviembre Diciembre

Continúa en la siguiente página.

Tabla 1–3 : *Continuación de la página anterior.*

Nombre de la variable	Descripción	Tipo de variable	Escala de medición	Categoría o intervalo
Cantón	Entidades de una provincia.	Cualitativa	Nominal	Alausí Chambo Chunchi Colta Cumanda Guamote Guano Pallatanga Penipe Riobamba
Sexo Paciente	El estado actual de las concepciones de género sobre hombres y mujeres.	Cualitativa	Nominal	Femenino Masculino
Edad Paciente	Tiempo que ha transcurrido desde el nacimiento.	Cuantitativa	Intervalo	[0 ,90[
Coordenada geográfica	Sistema de referencia que limita una ubicación específica.	Cualitativa	Nominal	Latitud Longitud
Condición final del paciente	Pacientes que murieron tras contagiarse de COVID-19.	Cualitativa	Nominal	Si, No

Fuente: Base de datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo periodo 2020 2021

Realizado por: Sayay Edison y Guzman Johana, 2023

CAPÍTULO IV

4. MARCO DE ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

anio_notifica	mes_notifica	dia_notificac	cod_provinci	provincia	canton	Latitud	Longitud	fecha_atenci	edad_pacien	tipo_edad	sexo_pacient	condicion_fir	clasificacion_final
2020	10	27	6	CHIMBORAZO	RIOBAMBA	9805916	759652	24/10/2020	63 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	6	5	6	CHIMBORAZO	RIOBAMBA	9805916	759652	6/4/2020	54 ANIOS	HOMBRE	VIVO	CONFIRMADO	
2020	9	17	6	CHIMBORAZO	RIOBAMBA	9805916	759652	17/9/2020	59 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	8	18	6	CHIMBORAZO	RIOBAMBA	9805916	759652	17/8/2020	39 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	7	6	6	CHIMBORAZO	RIOBAMBA	9805916	759652	7/5/2020	39 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	9	11	6	CHIMBORAZO	ALAUSSI	9750635	754493	9/4/2020	26 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	7	7	6	CHIMBORAZO	RIOBAMBA	9805916	759652	7/6/2020	30 ANIOS	HOMBRE	VIVO	CONFIRMADO	
2020	8	25	6	CHIMBORAZO	RIOBAMBA	9805916	759652	23/8/2020	18 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	8	25	6	CHIMBORAZO	RIOBAMBA	9805916	759652	23/8/2020	26 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	8	26	6	CHIMBORAZO	CHUNCHI	9742311	733923	24/8/2020	88 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	8	26	6	CHIMBORAZO	CHUNCHI	9742311	733923	24/8/2020	9 ANIOS	HOMBRE	VIVO	CONFIRMADO	
2020	8	2	6	CHIMBORAZO	ALAUSSI	9750635	754493	29/7/2020	29 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	7	3	6	CHIMBORAZO	GUAMOTE	9778398	764812	7/3/2020	31 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	8	14	6	CHIMBORAZO	RIOBAMBA	9805916	759652	14/8/2020	41 ANIOS	HOMBRE	VIVO	CONFIRMADO	
2020	7	5	6	CHIMBORAZO	RIOBAMBA	9805916	759652	7/5/2020	27 ANIOS	HOMBRE	VIVO	CONFIRMADO	
2020	10	28	6	CHIMBORAZO	RIOBAMBA	9805916	759652	21/10/2020	18 ANIOS	HOMBRE	VIVO	PROBABLE	
2020	6	18	6	CHIMBORAZO	RIOBAMBA	9805916	759652	17/6/2020	34 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	5	12	6	CHIMBORAZO	RIOBAMBA	9805916	759652	5/12/2020	37 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	10	30	6	CHIMBORAZO	COLTA	9796825	741471	7/2/2020	31 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	7	14	6	CHIMBORAZO	RIOBAMBA	9805916	759652	14/7/2020	27 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	8	1	6	CHIMBORAZO	CUMANDA	9754397	709154	29/7/2020	31 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	8	3	6	CHIMBORAZO	RIOBAMBA	9805916	759652	8/3/2020	48 ANIOS	HOMBRE	MUERTO	PROBABLE	
2020	10	6	6	CHIMBORAZO	COLTA	9796825	741471	10/5/2020	71 ANIOS	HOMBRE	VIVO	DESCARTADO	
2020	4	24	6	CHIMBORAZO	RIOBAMBA	9805916	759652	23/4/2020	44 ANIOS	HOMBRE	VIVO	CONFIRMADO	

Ilustración 1–4: Base de datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Fuente: (MINISTERIO DE SALUD PÚBLICA, 2022)

4.1. Análisis descriptivo

A continuación, se presenta el análisis descriptivo de las variables: Edad, Sexo, Condición final del paciente, Cantón donde se registró el caso y Fecha de atención; esto con la finalidad de presentar información preliminar sobre los casos de Covid-19 registrados en el período de análisis en el que se concentra el presente estudio, para obtener una visión sobre el comportamiento y distribución de los datos.

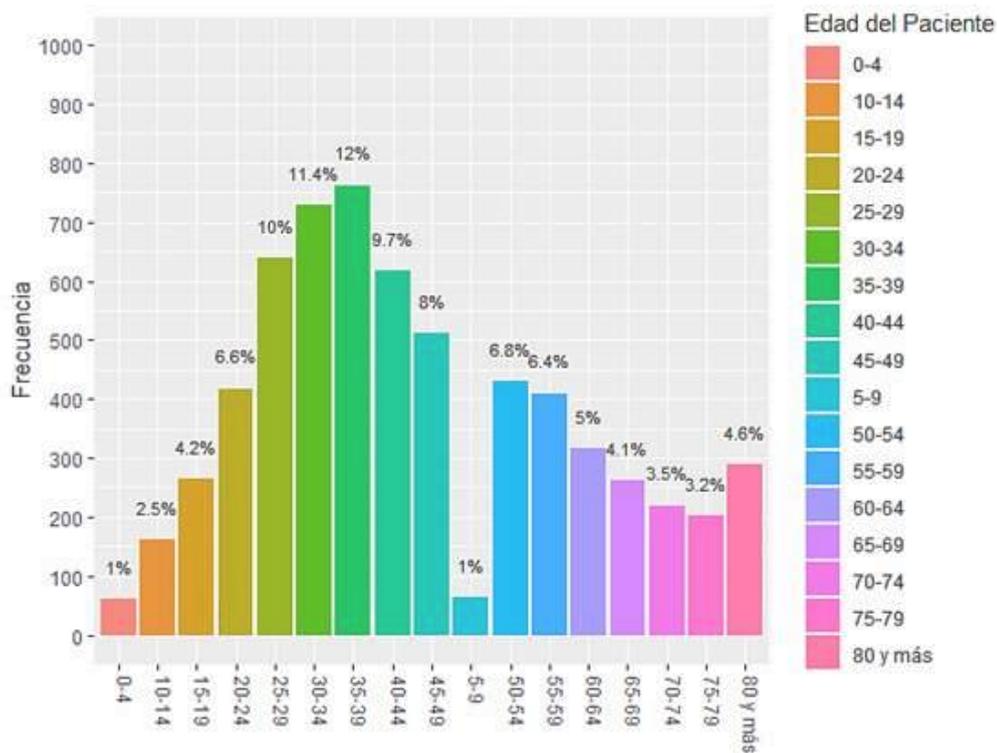


Ilustración 2-4: Diagrama de barras que presenta la edad de pacientes contagiados **Fuente:**

Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 2-4 muestra los resultados sobre la cantidad de pacientes Covid-19 registrados, de acuerdo con la edad de estos. El diagrama de barras organiza las frecuencias de la variable edad por grupos quinquenales, en donde es posible observar que el 12 % de pacientes Covid-19 fueron hombres y mujeres de entre 35 a 39 años registrados en la provincia de Chimborazo en el período 2020-2021, mientras que el 1 % son la menor representación en donde se denotan los niños de 0 a 9 años del total de pacientes registrados.



Ilustración 3-4: Diagrama de barras de las frecuencias absolutas de la variable sexo

Fuente: Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 3-4 presenta la distribución de la cantidad de pacientes Covid-19 registrados de acuerdo a su sexo; resultando de ello que, la mayoría de los pacientes fueron hombres con un total de 5849 casos, representando el 92 % del total de pacientes. Por otra parte, minoritariamente las mujeres representan apenas el 8% con 512 casos registrados.

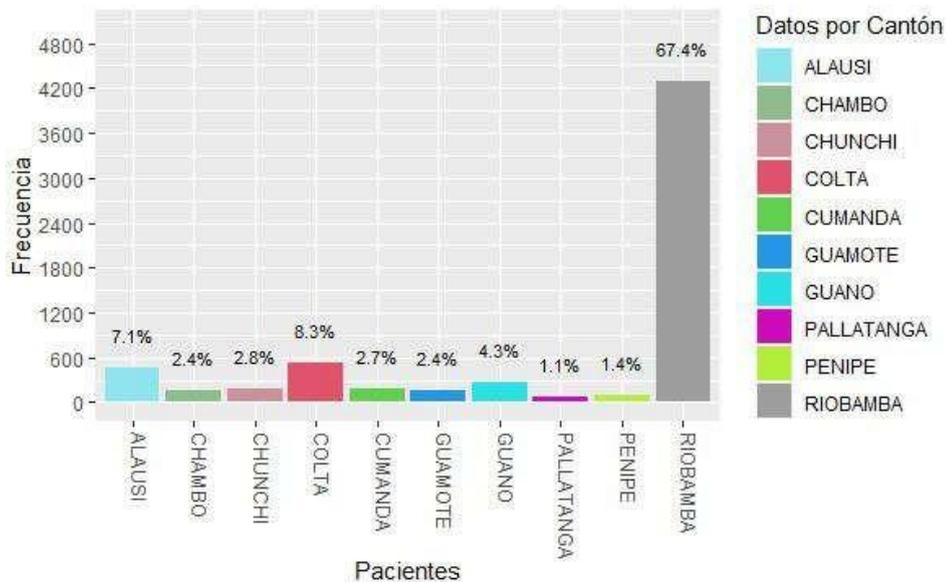


Ilustración 4-4: Diagrama de barras frecuencias absolutas de la variable Cantón

Fuente: Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 4-4 diagrama de barras de las frecuencias absolutas variable cantón representa los casos SARS-COV-2 (COVID-19) por cantones pertenecientes a la provincia de Chimborazo en el período 2020-2021, en el cual resalta Riobamba con la mayor representación de casos registrados con un 67.4 %, el cantón que menos casos registró fue Pallatanga con el 1.1 % a nivel provincial.

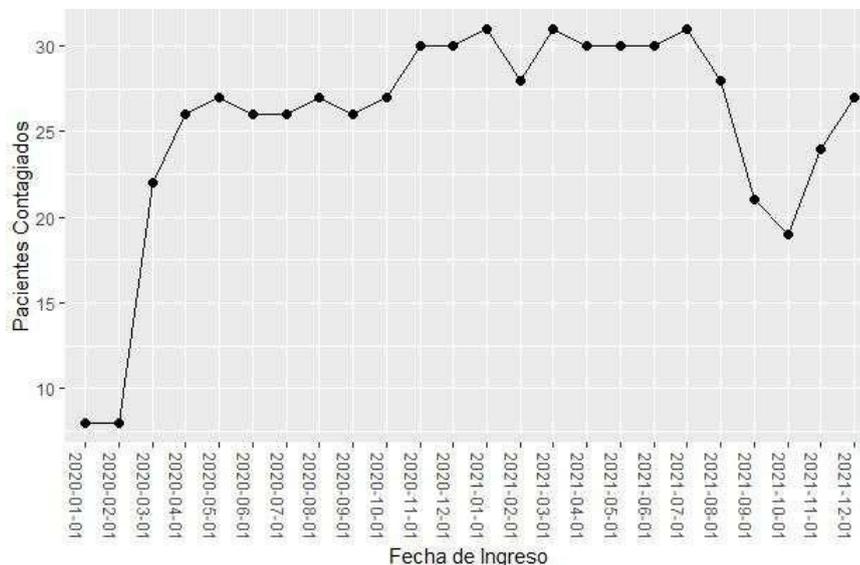


Ilustración 5-4: Series de tiempo mensual de la cantidad de contagios

Fuente: Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 5-4 series de tiempo mensual de la cantidad de contagios, representa los casos SARS-COV-2 (COVID-19) registrados en la provincia de Chimborazo mensualmente en el período 2020-2021. Se observó que en el mes de marzo del 2020 se experimenta un crecimiento abrupto de casos marcando un incremento general en los meses posteriores, es posible distinguir que los puntos más críticos en los que se registraron más casos se denotan en los meses de enero, marzo y julio del año 2021, en octubre disminuyeron los casos covid, pero para los meses de noviembre y diciembre la tendencia aumentó generando un crecimiento exponencial de los pacientes contagiados.

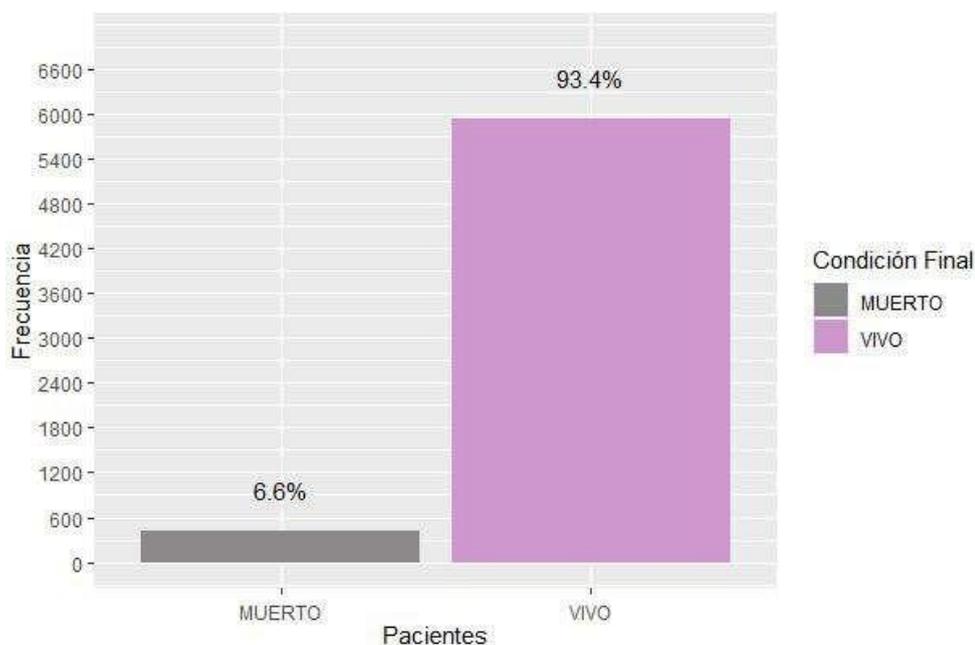


Ilustración 6–4: Diagrama de barras de las frecuencias absolutas de la variable condición final paciente

Fuente: Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 6–4 muestra el diagrama de barras de las frecuencias absolutas de la variable condición final del paciente, correspondiente al número de pacientes vivos y muertos de los casos confirmados de SARS-COV-2 (COVID-19) en la provincia de Chimborazo para el período 2020-2021, se puede observar que la tasa de mortalidad registrada es muy baja con un 6.6 % decesos mientras que un 93.4 % de pacientes contagiados no murieron.

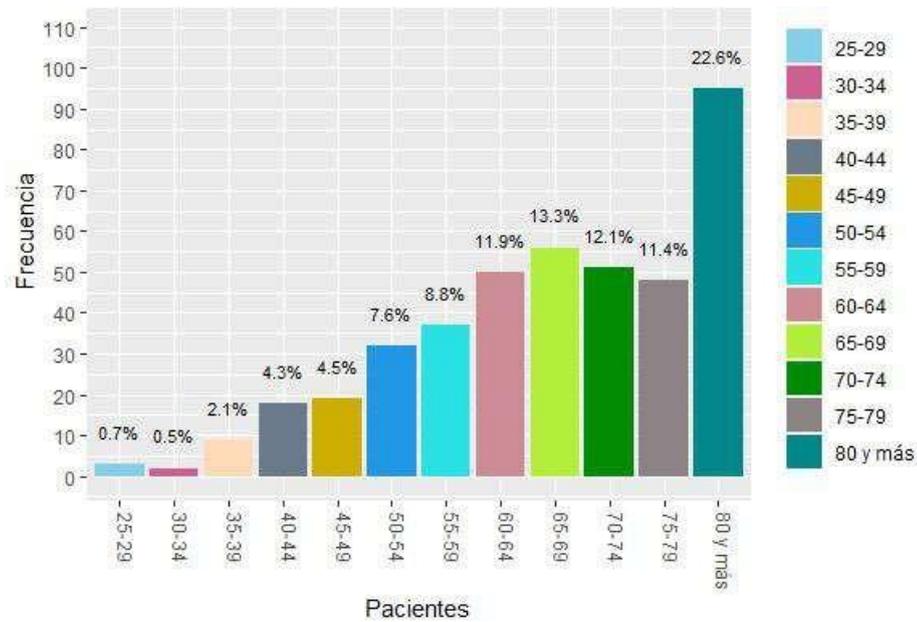


Ilustración 7-4: Diagramas de barras de pacientes muertos según edades

Fuente: Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 7-4 diagrama de barras de pacientes muertos según edades por casos de SARS-COV-2 (COVID-19) confirmados dentro del período 2020-2021 en la provincia de Chimborazo se identifica que el grupo de 80 años en adelante representa un 22,6 % de decesos, los pacientes de 25 a 34 años fueron menos vulnerables a este evento representando el 1,2 % del total de personas muertas.

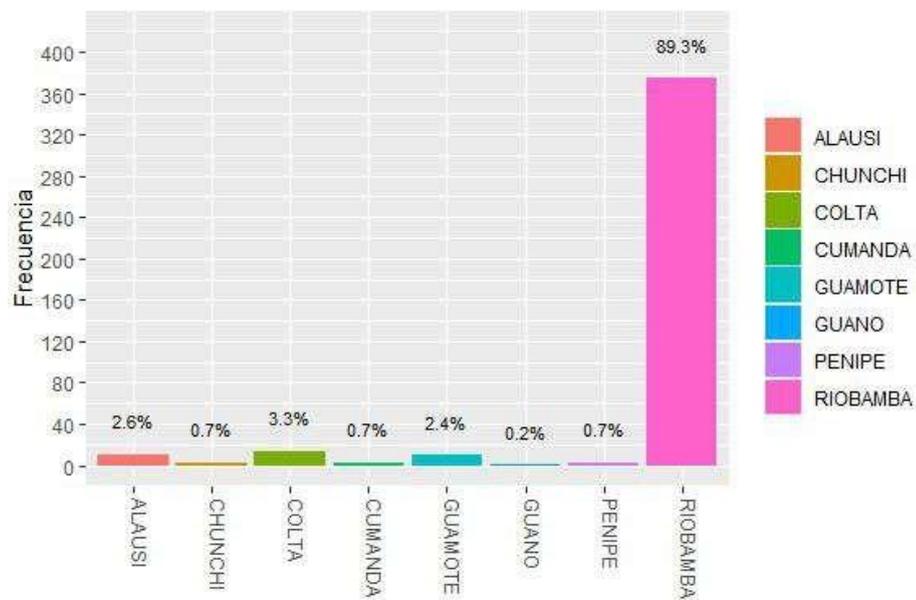


Ilustración 8-4: Diagrama de barras de pacientes muertos según cantones

Fuente: Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 8-4 diagrama de barras de pacientes muertos según cantones muestra la distribución del número de muertes registrados por Covid-19 confirmados dentro del período 2020-2021 en la provincia de Chimborazo, evidentemente Riobamba es el cantón que mayor cantidad de muertes registra con el 89,3 % a diferencia de Colta, Alausí y Guamote que apenas alcanzan la decena de muertes, Penipe, Guano, Cumandá y Chunchi son la menor representación de muertes del total de la provincia. Los cantones Chambo y Pallatanga no presentaron decesos durante este período.

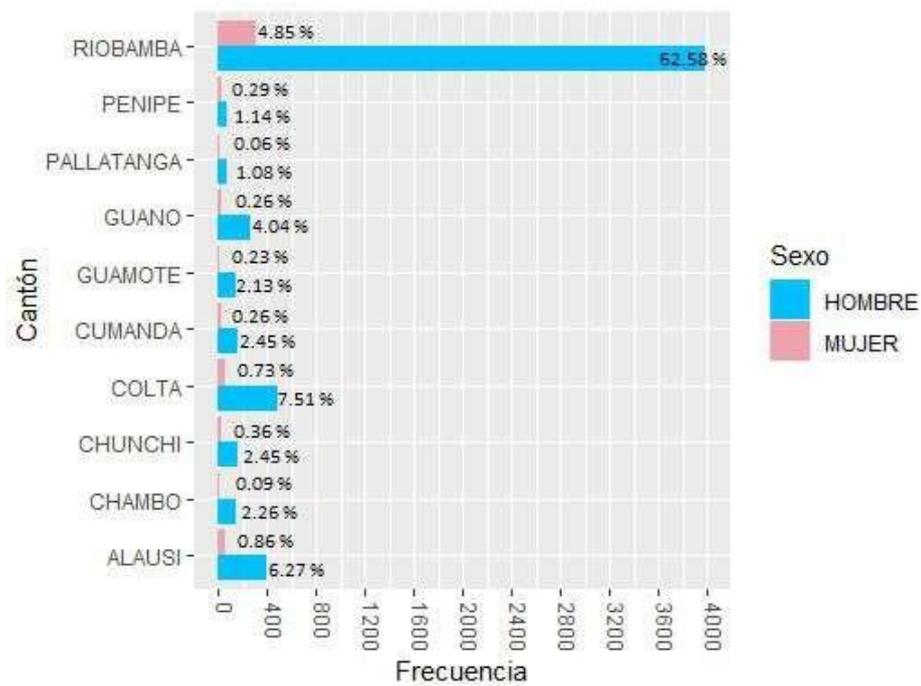


Ilustración 9–4: Diagrama de barras del número de contagios por sexo y cantón Fuente: Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 9-4 diagrama de barras del número de contagios por sexo y cantón muestra el número de pacientes contagiados por SARS-COV-2 (COVID-19) en cada cantón de la provincia de Chimborazo para el período 2020-2021 de acuerdo con el sexo de los pacientes es evidente que la mayoría de contagiados se concentran en el cantón Riobamba con un 62.58 % de hombres y un 4,85 % de mujeres, lo cual es congruente a los análisis anteriores. Otro aspecto relevante es que en todos los cantones de la provincia los casos son más frecuentes en hombres que en mujeres.

4.2. Análisis exploratorio espacial

Actualmente todo puede ser mapeado, gracias a los avances tecnológicos se ha registrado datos referentes al espacio que pueden ser leídos y analizados, desde nuestros propios movimientos hasta datos meteorológicos que son tomados mediante GPS o por otros aparatos digitales y a partir de estos datos se puede elaborar análisis espaciales con una perspectiva geográfica, en este caso utilizamos los datos registrados de SARS-COV-2 (COVID-19) de la provincia de Chimborazo.

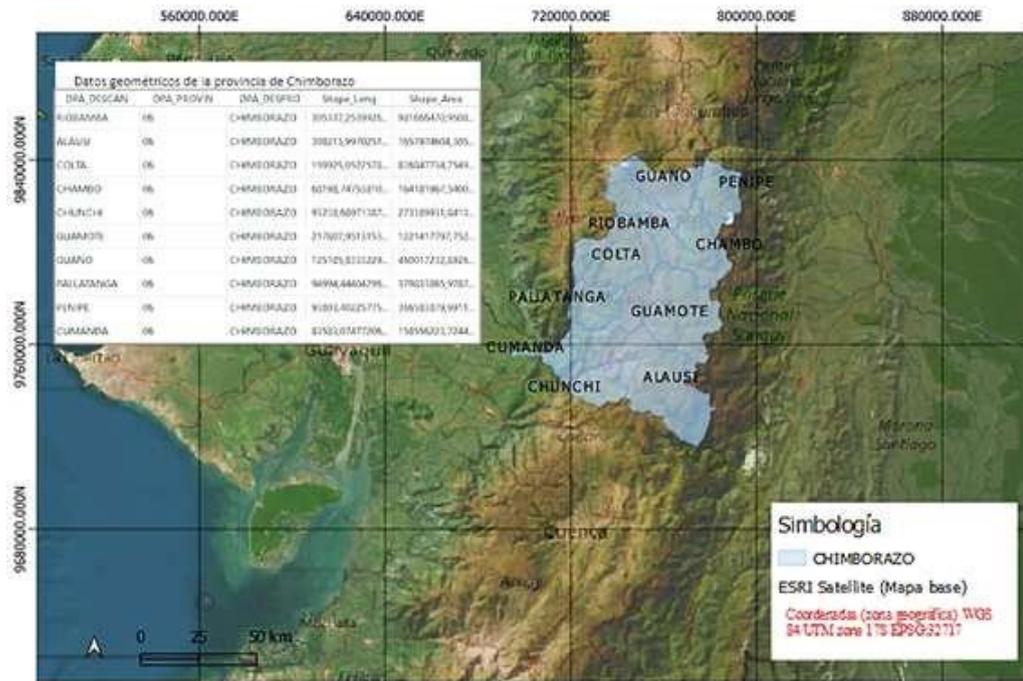


Ilustración 10-4: Representación espacial cantones de la Provincia de Chimborazo

Realizado por: Sayay Edison y Guzman Johana, 2023

Por medio del Qgis en la ilustración 10-4 se ha considerado la zona que se va analizar de manera espacial, en la ilustración podemos visualizar la información de los cantones de la provincia de Chimborazo, cargados mediante un archivo shape con extensión (.shp) en donde se almacena la información geométrica en formato vectorial de localizaciones geográficas.

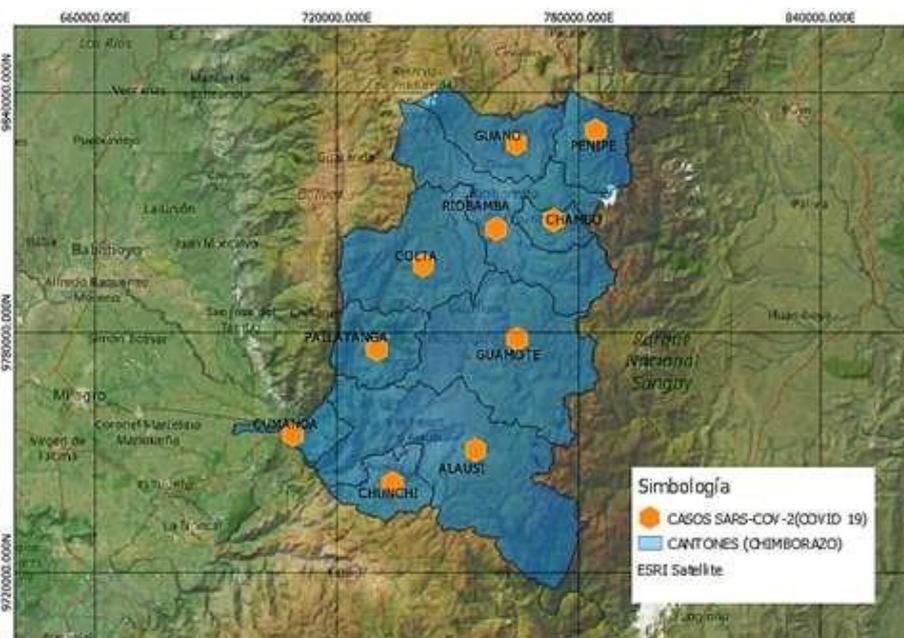


Ilustración 11-4: Representación de la capa vectorial de datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 11–4 se construyó la base vectorial a partir los datos sars cov 2 obtenidos de los registrados del ministerio de salud, la variable ubicación fue esencial para definir los valores de latitud y longitud, que permiten generar la capa vectorial de geometrías puntuales y de esta manera conocer el comportamiento espacial o aglomeración de distribución de los datos, se puede observar que tenemos 10 puntos de localización dado que contamos únicamente con las referencias de ubicación por cantón de cada paciente, por lo tanto, para este caso aún no se puede definir las zonas más aglomeradas.

4.2.1. *Mapa temático de cuantiles*

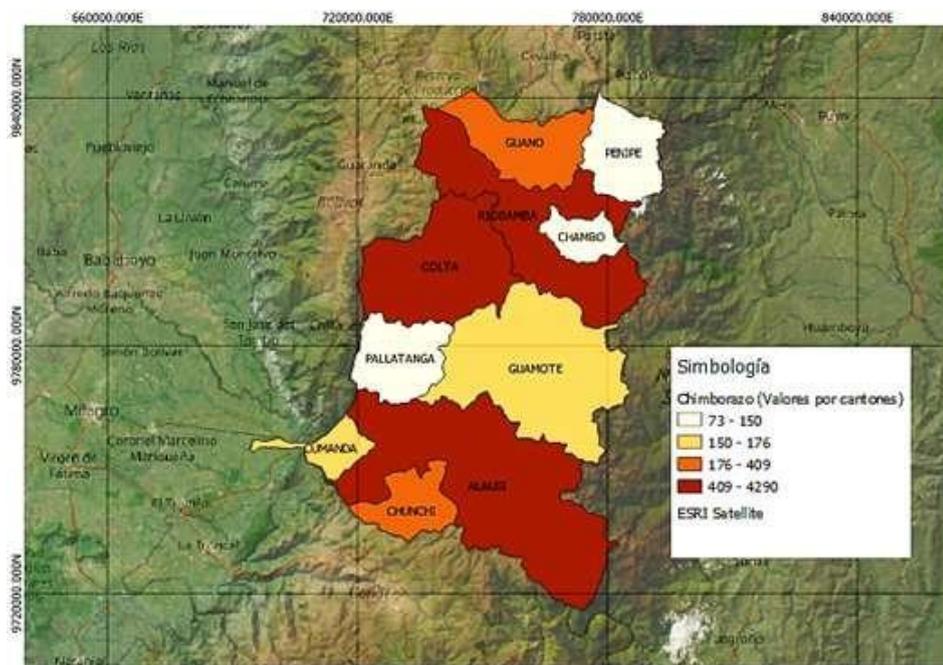


Ilustración 12–4: Mapa de cuantiles de los casos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 12–4 se puede apreciar la información de pacientes contagiados por covid en categorías, cada bloque es una categoría puesto que tiene un intervalo de valores definidos, los valores de estos grupos pueden ser manipulados dependiendo el número de intervalos que se quiera observar, se realiza esta representación con la intención de que los datos se distribuyan de manera uniforme, en este caso los intervalos se desglosan de la siguiente manera el bloque 1 (73 – 150), bloque 2 (150 – 176), bloque 3(176 – 409), bloque 4(409- 4290) de pacientes contagiados por covid 19,por lo tanto, se puede determinar cuál es el bloque de cantones que tienen los valores más altos de contagios por SARS-COV-2 (COVID-19) debido a que están representados por un color azul con diferente nivel de intensidad para cada cantón de la provincia de Chimborazo.

Cada bloque se diferencia por el sombreado y valor de intervalo, organizado de menor a mayor en donde el grupo más bajo está representado por el color blanco, el cual indica que los pacientes contagiados por covid pertenecen a los cantones Penipe, Chambo, Pallatanga dado que están en un intervalo de (73 a 150) pacientes contagiados por covid ya se puede distinguir entre el bloque más grande, mismo que se encuentra en un intervalo de (409 a 4290) pacientes contagiados por covid se identifica por su predominio color intenso en los cantones Colta, Riobamba y Alausí.

4.2.2. Medidas centrográficas

Se procede a calcular las medidas de tendencia central esenciales de los casos SARS-CoV-2 (COVID-19) y mediante el software Qgis se ejecuta las herramientas necesarias, para saber si estos valores se encuentran dispersos o presentan aglomeración en zonas específicas de la provincia de Chimborazo.

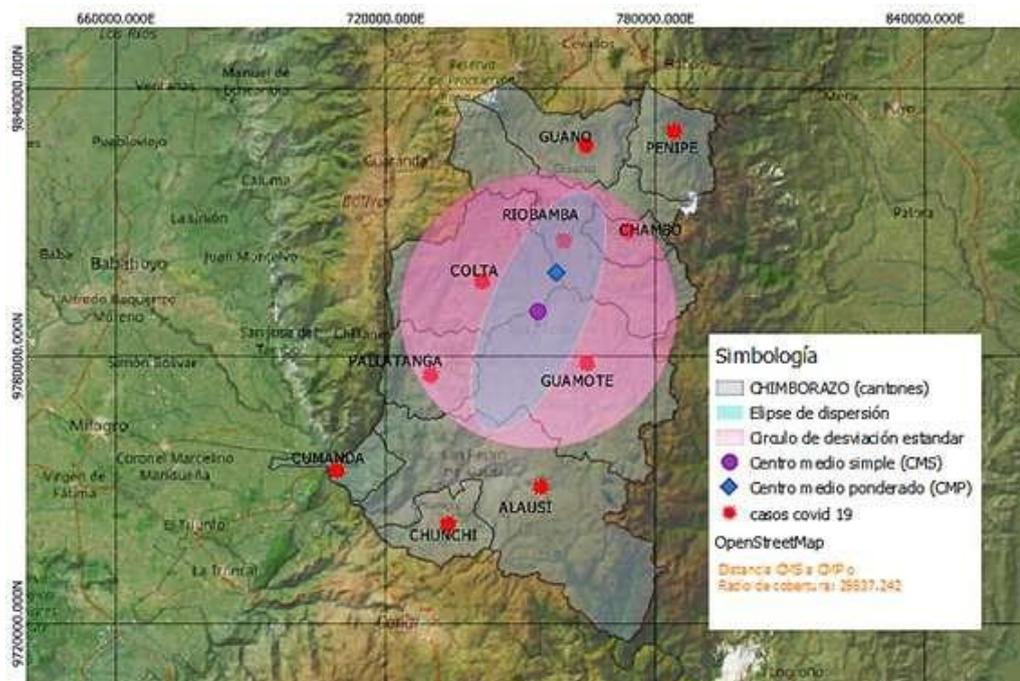


Ilustración 13–4: Centro medio simple, ponderado y desviación estándar de los datos

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 13–4 evaluamos la centralidad espacial de conjunto de valores de contagios SARS-CoV-2 (COVID-19) con el complemento Spatial Point Pattern Analysis en Qgis puesto que genera una nueva capa vectorial de geometrías de puntos, se representa el centro medio simple denotado por un punto de color morado ubicado en el sector del cantón Colta puesto que el resultado obtenido indica el centro de gravedad de los diez cantones de la provincia de Chimborazo.

Para el cálculo del centro medio ponderado se utilizó el número total de contagios en cada localización, a partir de este dato se calculó el círculo de desviación estándar, generado en una capa vectorial de geometría poligonal este valor representa el radio de cobertura medido desde el centro medio simple a la circunferencia se obtiene un valor de 28536,050 metros de distancia y desde el centro medio simple al centro medio ponderado con un valor de 7884,744m.

4.2.3. *Mapa de intensidad*



Ilustración 14-4: Mapa de intensidad o de calor de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 14-4 se representan las zonas con un mayor nivel de concentración se les denomina puntos calientes, efectivamente se ve reflejado en el cantón Riobamba, el área focal estimula un nivel de densidad notoria en referencia a la presencia de contagios SARS-COV-2 (COVID-19), en los demás cantones de la provincia de Chimborazo, por lo tanto, se puede denotar con un color amarillo de baja tonalidad, las zonas con una menor cantidad de presencia de contagios, en los cantones Alausí, Colta, seguidos los siete cantones restantes con un tono mas bajo indicando que existe una aglomeración menor en comparación a los demás cantones.

4.3. Análisis estructural espacial

En esta etapa se estimó la existencia de autocorrelación, variabilidad espacial con la finalidad de encontrar un buen ajuste de variograma, a través de la modelación de semivariograma ejecutado en el software Qgis y Geoda, analizando un modelo geoestadístico de la variable regionalizada más adecuada acerca de su variabilidad, a través de un análisis del patrón de localizaciones, es posible medir el grado de similitud entre vecinos; por ende, es indispensable conocer la distribución geográfica de los datos para continuar con esta etapa de estudio.

4.3.1. Análisis del Índice de Moran

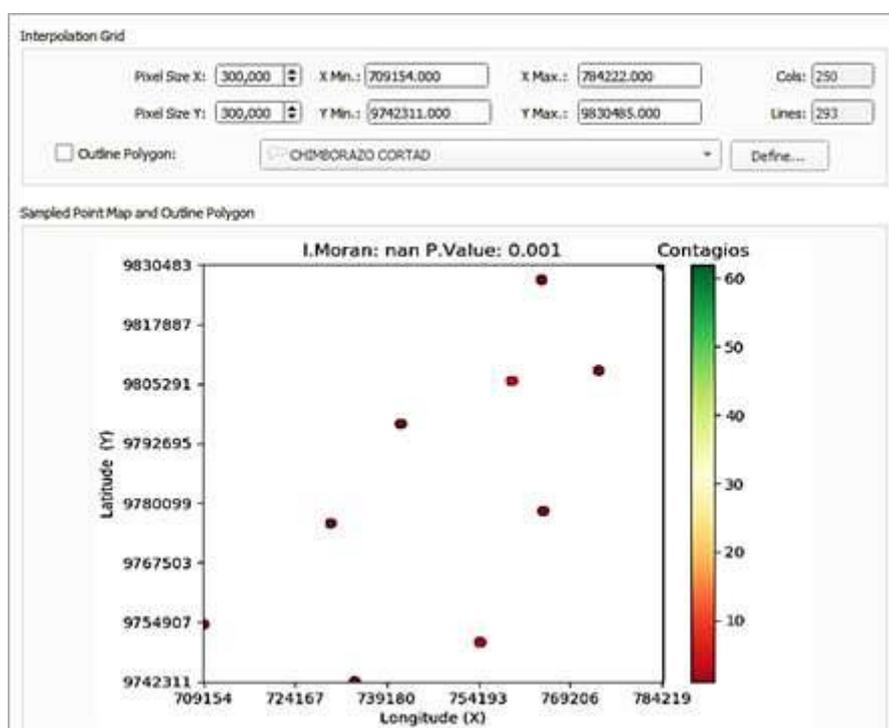


Ilustración 15–4: Resultado I de Moran de casos SARS-COV-2 (COVID-19)

Realizado por: : Sayay Edison y Guzman Johana, 2023

En la ilustración 15–4 El índice de moran calculado con el complemento Smart-map en Qgis, esta medido en dos dimensiones x,y, con tamaño de pixel por defecto 10, se ha ampliado el tamaño a 300 pixeles para obtener una mejor resolución, de esta manera se observa mucho mejor los casos de pacientes contagiados, el valor-p obtenido es una medida que detecta la existencia de autocorrelación ya sea positiva o negativa, si el valor-p es cero no hay autocorrelación espacial, se puede decir que este resultado es parecido al coeficiente de correlación pero en este caso mide el grado similar de las características de los cantones entre sus vecinos con relación a todas las unidades espaciales.

Se obtiene un valor-p 0.001, representa un resultado significativo que permite afirmar la existencia de autocorrelación espacial de la variable contagios por SARS-COV-2 (COVID-19) en los cantones de la provincia de Chimborazo, dado que los episodios de transmisión de contagios covid en los años 2020, 2021 fueron progresivos en la mayoría de las zonas, esta medida nos indica la existencia de un patrón espacial que asegura la presencia de autocorrelación.

4.3.2. Análisis de autocorrelación espacial (LISA)

Mediante este análisis se requiere conocer la presencia de autocorrelación LISA que identifica las subáreas de un espacio geográfico donde se muestran los patrones de distribución espacial similares, de manera que es indispensable la aplicación de un I de Moran, complementado con el análisis de autocorrelación LISA que permite identificar las características similares entre los cantones de la provincia de Chimborazo, explicando el nivel de significancia de los agrupamientos espaciales y como cada unidad espacial se asemeja a las demás unidades espaciales vecinas, puesto que representa las localizaciones mediante agrupamiento o clúster.

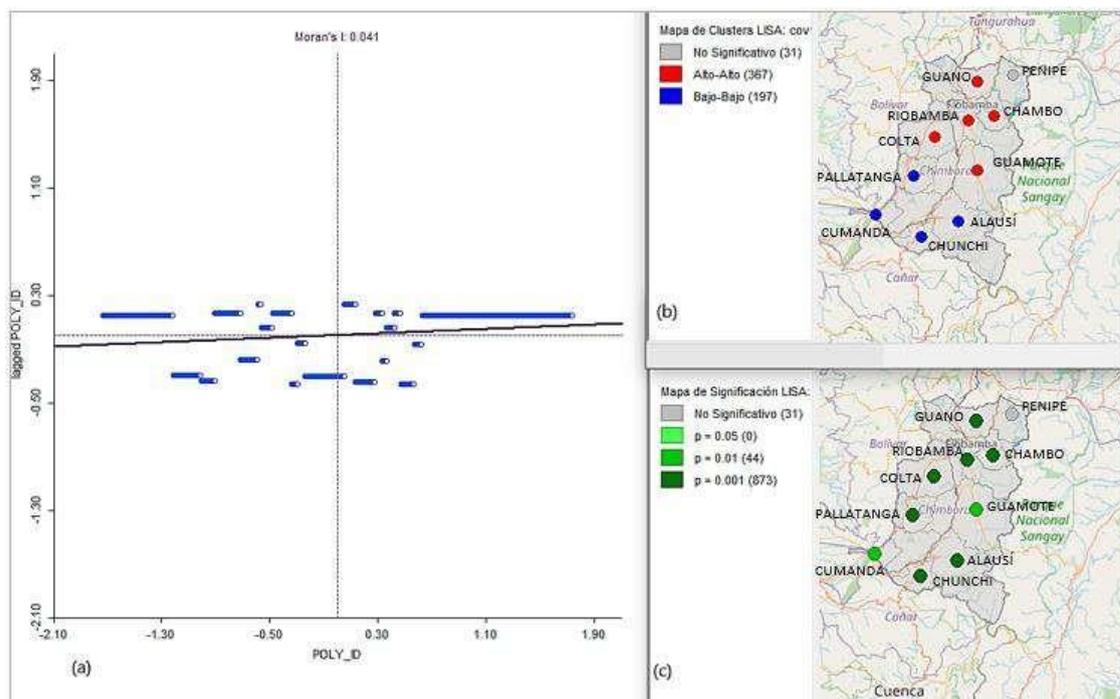


Ilustración 16-4: Mapa de clusters, mapa de significancia (LISA), valor - p, e índice de morán

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 16-4 aplicado el análisis de autocorrelación espacial LISA se determina que existe una relación espacial entre los diferentes cantones de la provincia de Chimborazo, procesado los cálculos mediante el software Geoda se deduce el mapa de agrupaciones clúster, mapa de significancia y el I de Moran en la gráfica (a) en el cual presenta un valor de 0.041 cercano a

cero indicando que si existe autocorrelación entre los grupos cantonales, cabe recalcar que se complementa con el análisis anterior dado un p-valor significativo.

El mapa de significación LISA (c) se manifiesta en distintas tonalidades, el área verde más oscura representa a los sectores con valores significativos altos, denotado 7 zonas entre estos los cantones, Riobamba, Guano, Chambo, Colta, Pallatanga, Alausí, Chunchi, conformadas por 873 unidades espaciales que hacen referencia al valor más alto de pacientes contagiados por SARS-COV-2 (COVID-19), seguidos con un nivel de significancia 0.01, dos las zonas de la provincia de Chimborazo, Guamote y Cumandá; no presentan valores significativos un total de 31 unidades espaciales de estudio correspondientes al cantón Penipe, dado los resultados se puede determinar la existencia de autocorrelación.

A través del mapa de clústeres LISA (b) se identifica la agrupación espacial entre las unidades de cada una de las zonas cantonales, resultando la formación de clúster de 3 agrupaciones en el nivel (alto – alto) representado por el color rojo, que contienen 387 observaciones con una alta similitud entre vecinos de los cantones Colta, Guamote, Riobamba, Chambo, Guano, las zonas con valores (bajo – bajo) representado por el color azul con un total de 197 observaciones, se identifican en los cantones Pallatanga, Alausí, Chunchi, Cumandá, existe un grupo sin significancia estadística correspondiente al cantón Penipe con un total de 31 observaciones, dados los resultados podemos ya tener una noción de las zonas calientes y frías posibles que permitan identificar un patrón específico de cierta agrupación de las cantones de Chimborazo.

4.3.3. *Análisis de vecinos más próximos*

Observed mean distance: 0.0000000000
Expected mean distance: 1321.18468206948
Nearest neighbour index: 0.0000000000
Number of points: 948
Z-Score: -58.90267951615

Ilustración 17–4: Resultados del análisis de vecinos más próximos generado en Qgis

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 17–4 se puede apreciar una buena capacidad de aproximación dado que, los resultados estadísticos indican la distancia entre todos los puntos respecto a sus vecinos más próximos, en este caso la distancia observada con un valor 0 es sumamente inferior a la distancia media esperada con un valor de 13231.185 lo cual nos indica que corresponde a un patrón clusterizado puesto que presenta agrupaciones, el valor z-score bajo indica que es poco probable que el procedimiento sea espacialmente aleatorio hasta este punto de investigación se afirma que los puntos se distribuyen espacialmente de forma agrupada.

4.3.4. Ejecución modelación de variogramas

Para modelar un variograma se considera los siguientes parámetros; la discontinuidad en el origen o efecto nugget, el valor máximo de variabilidad (sill) también mencionada meseta de la variable regionalizada, el área de correlación llamada alcance, con el fin de conocer el comportamiento de un variograma y dar un buen ajuste al modelo a seleccionar como óptimo de estimación de varianza.

A través del software Qgis, se procede a ejecutar el completo Smart Map (Mapa inteligente completo) que permite realizar la representación de modelos de variogramas, ya cargados los datos de contagios SARS-CoV-2 (COVID-19); determinamos la máxima distancia y lag(h) estos valores son primordiales para iniciar con la ejecución del modelo.

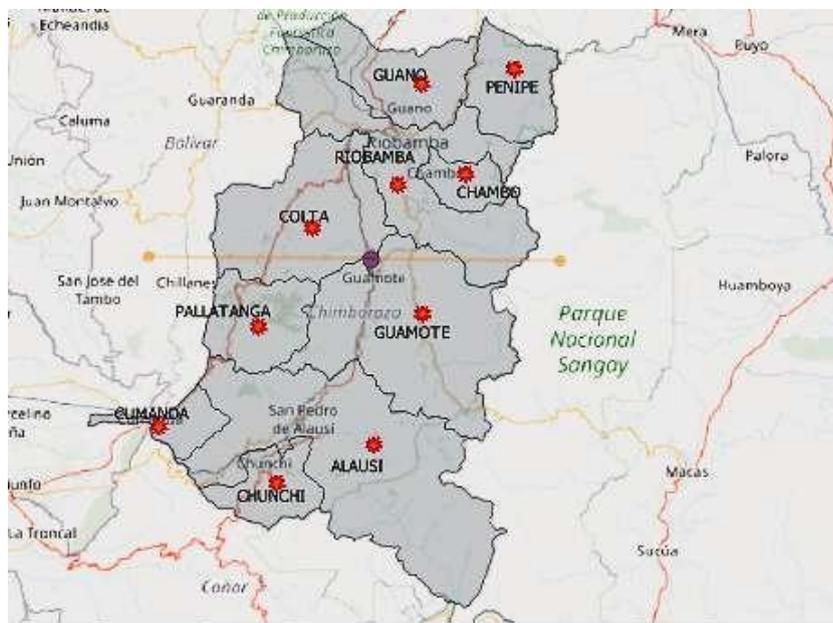


Ilustración 18–4: Máxima distancia casos SARS-COV-2 (COVID-19) en Chimborazo

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 18-4 se observa el segmento del área total que cubre la zona de puntos de pacientes contagiados por covid de la provincia de Chimborazo, con un valor de 81614.175 m, este valor es necesario para el calcular un variograma además, se observa el punto de coordenadas medias calculado anteriormente que cae justo en la mitad del mapa siendo un punto guía de división que asegura la obtención de un valor más exacto desde el centro hacia la cobertura que encierra los cantones.

Se calcula el lag(h) previo a ser utilizado en los variogramas se obtiene desde el punto denominado centro medio simple que indica el centro de gravedad de la distribución espacial de los casos covid, hasta el centro medio ponderado que representa a un conjunto de puntos ajustados a cada unidad espacial por lo tanto se puede generalizar un determinado valor de dicha distancia h entre dos puntos cualquiera con un valor h de 7884,744m, para establecer este valor en un variograma se procede a probar el modelo si se puede observar o no una buena capacidad de distribución de datos, caso contrario el valor puede aumentar o disminuir.

4.3.4.1. Modelación lineal

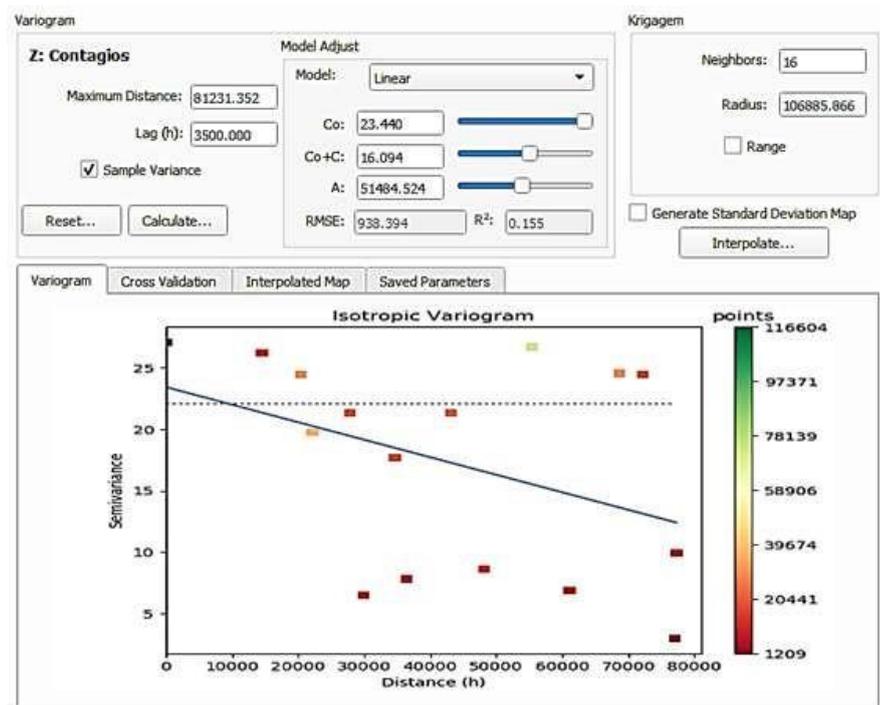


Ilustración 19-4: Modelo lineal de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 19-4 se observa el modelo lineal dado que la línea de tendencia no tiene estacionalidad, cruza el umbral, pero no es muy significativo el valor resultante de R^2 es muy

bajo con un valor de 0. 155 además en el mapa de calor se observa que los puntos se encuentran entre los valores de de 0 a 8 que indica el número de contagios, con una distancia máxima de de 81231.352 y un valor lag(h) de 382.823 que se ha calculado anteriormente ya se puede describir un modelo lineal con un efecto nugget de 23.440 (co) un umbral de $(co-c) = 16.094$ en donde se observa que que la línea de tendencia desde el rango $a = 51484.524$ tiene un corte minúsculo con la varianza del umbral por lo cual no optamos por este modelo, además se observa un mapa de calor que indica la existencia de correlación entre los puntos amarillos mismos que indican las posibles predicciones de los pacientes contagiados.

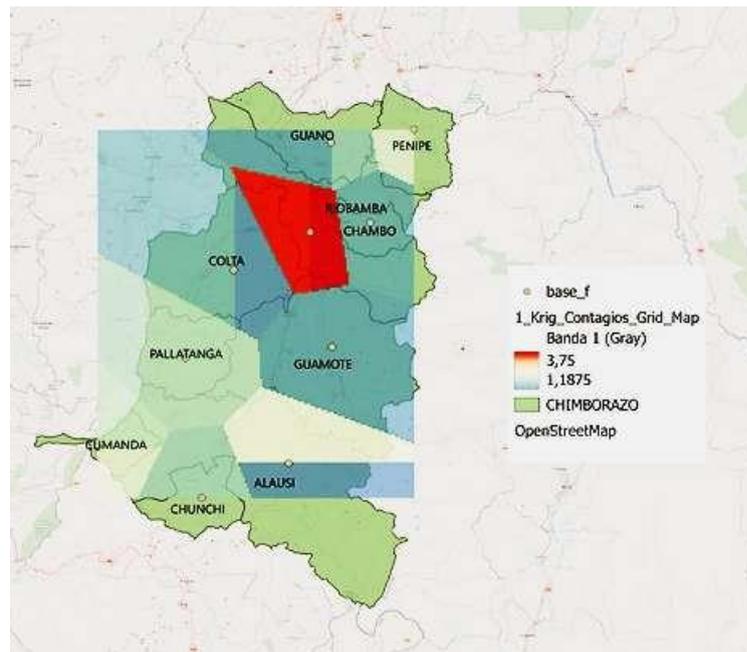


Ilustración 20–4: Krigging modelo lineal de los datos

SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 20–4 se observa que la mayor presencia de pacientes contagiados están en el cantón Riobamba, es la zona más caliente con el mayor número de registros, notoriamente se puede observar que las zonas frías corresponden a los cantones Chambo, Guano, Guamote, Colta, se indican con el color azul intenso, las zonas verdes reflejan la presencia de contagios como un valor intermedio en los cantones de Pallatanga, Cumanda, Penipe, Chunchi, esto nos permite identificar que existen casos covid 19 en todos los cantones, en alguno con un valor mínimo, pero hubo presencia de contagios en toda la zona.

4.3.4.2. Modelación lineal to still

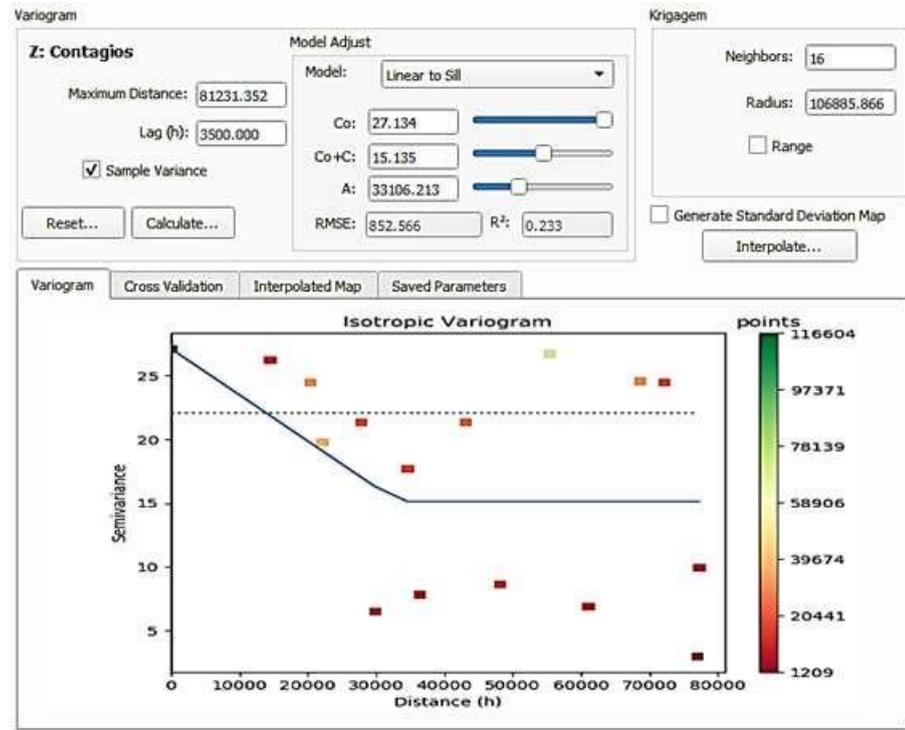


Ilustración 21–4: Modelo lineal to still de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 21–4 se observa que en el mapa de calor algunos puntos se encuentran entre los valores de 0 a 25 que indican los contagios en los cantones con una distancia máxima de 81231.352 y una distancia lag(h) de 3500.00 en este caso la barra de colores representa al mapa de calor donde existen menos contagios representado el color rojo y donde hay más contagios representado por el color verde, además la distancia máxima dentro de un variograma representa la estimación de la varianza y qué influye en el modelo las distancias máximas es una parte muy importante para representar h dentro del variograma podemos ver que existe estacionalidad además con un valor $co = 27.134$ que indica el efecto nugget no se observan valores fuera del patrón general descartando la existencia de outliers, $co+c=15.135$ con un rango de 33106.213 alcanza el umbral con un R^2 de 0.233 cabe destacar que presenta estacionalidad la cual no es muy favorable para la representación del modelo.

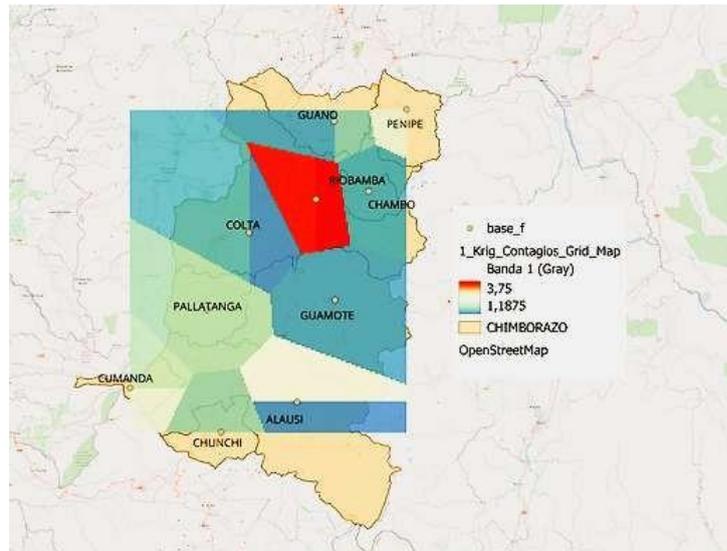


Ilustración 22–4: Krigging modelo lineal to still de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 22–4 se observa que la mayor presencia de pacientes contagiados están en el cantón Riobamba, la interpolación es similar en relación al modelo anterior.

4.3.4.3. Modelación exponencial

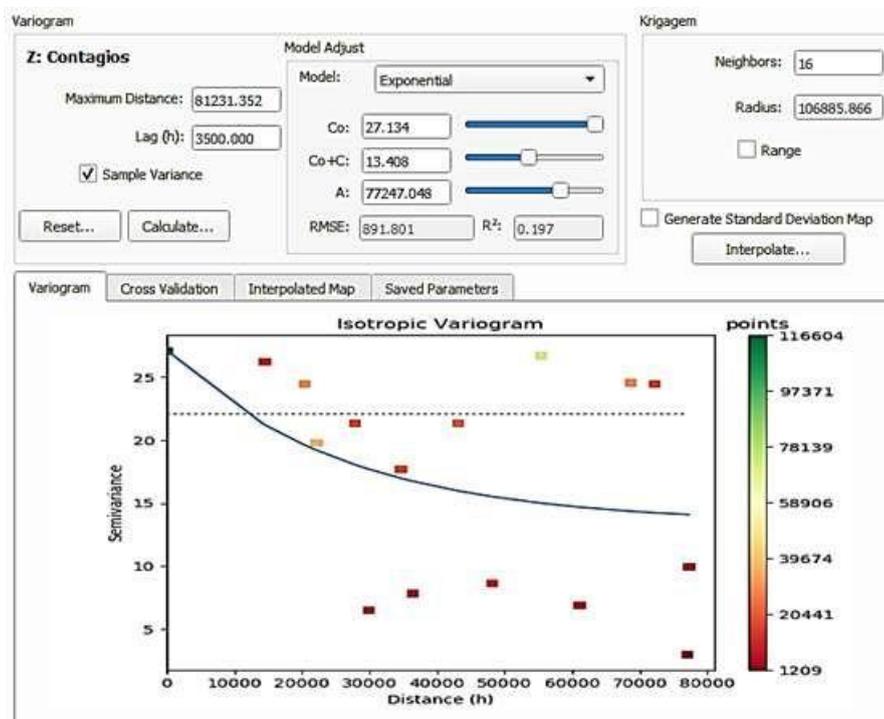


Ilustración 23–4: Modelo exponencial de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 23–4 se observa el modelo exponencial se puede denotar los puntos que representan a los pacientes contagiados mismos que se encuentran entre los valores de 0 a 25 que indican los contagios en los cantones con una distancia máxima de 81231.352 y una distancia lag(h) de 3500.00 la distancia máxima es una parte muy importante para representar h está considerada en base al valor encontrado del índice de Moran que cubre la cobertura del mapa estudiado dentro del variograma podemos ver que existe estacionalidad además con un valor $c_0 = 27.134$ que indica el efecto nugget no se observan valores fuera del patrón general descartando la existencia de outliers, $c_0+c_1=13.408$ con un rango de 77247.048 alcanza el umbral con un R^2 de 0.197 denota que es una buena representación.

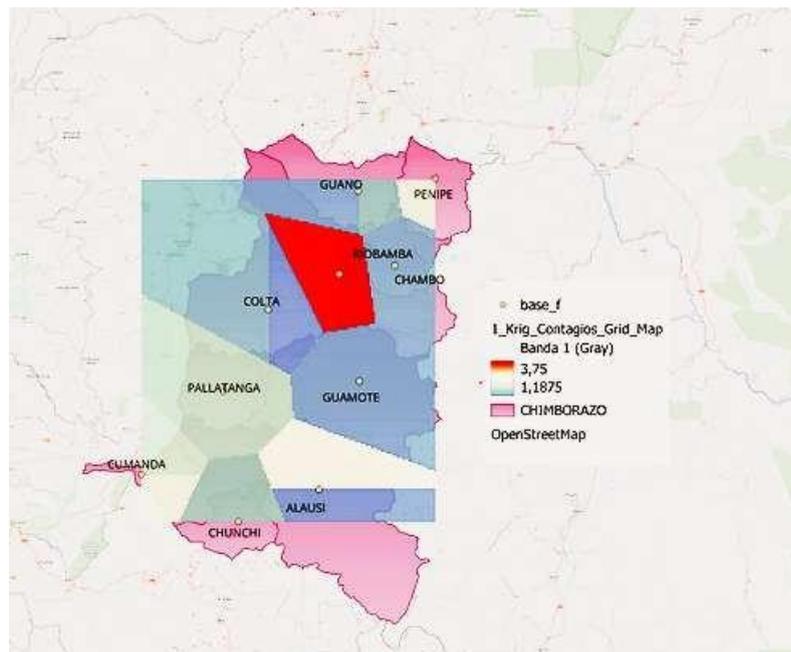


Ilustración 24–4: Krigging modelo exponencial de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 24–4 se observa que al aplicar la interpolación krigging no hay relevancia en la variación de intensidad de colores que indiquen cambios en el modelo, los resultados corresponden a una misma predicción en relación al anterior modelo, puesto que también representa la mayor presencia de pacientes contagiados en el cantón Riobamba, por ser la zona más caliente, las zonas frías corresponden a los cantones Chambo, Guano, Guamote, Colta, se indican con el color azul intenso, como un valor intermedio en los cantones de Pallatanga, Cumanda, Penipe, Chunchi, con un color verde esto nos permite identificar que absolutamente todos los cantones fueron afectados.

4.3.4.4. Modelación esférico

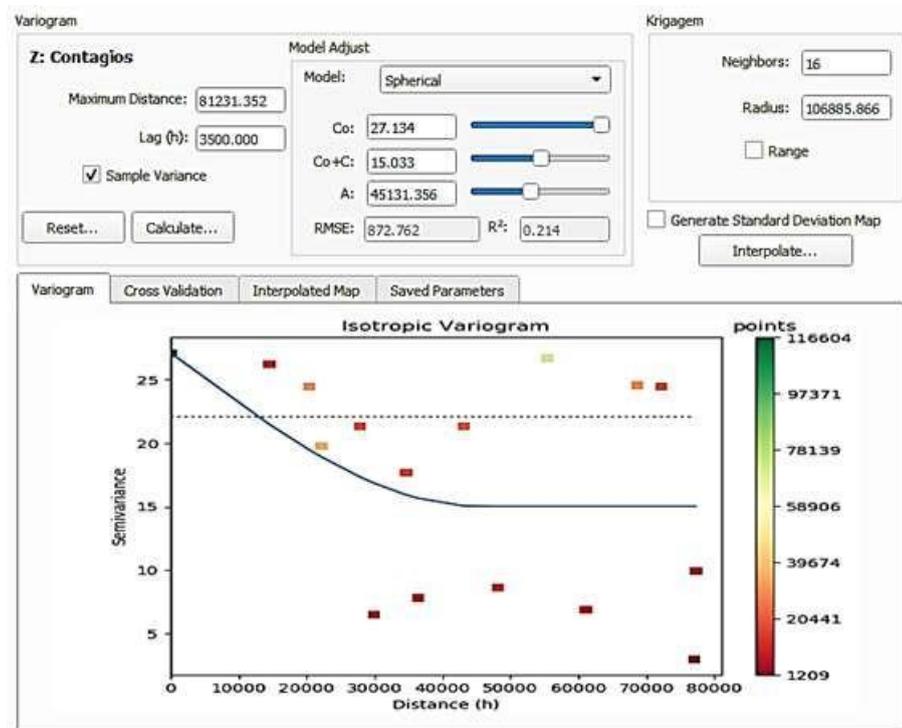


Ilustración 25–4: Modelo esférico de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 25–4 se puede observar que en el modelo esférico existe una mínima estacionalidad con una distancia máxima de 81231.352 y un valor lag(h) de 3500.00 se puede notar que con un efecto nugget de 27.134 con valor un umbral de 15.033 se observa que la línea de tendencia tiene corte con la varianza es decir corta el umbral efectuado, en el cálculo nugget no se identifica la existencia de un atípico y además el valor R^2 es mucho más alto que los valores anteriores por lo cual optamos por este modelo considerando que es el mas apto hasta el momento, de igual forma se observa un mapa de calor el cual indica que existen peronas contagiadas en gran parte de las zonas con un nivel mucho mas intenso además se puede afirmar que indican posibles predicciones favorables de los pacientes contagiados.

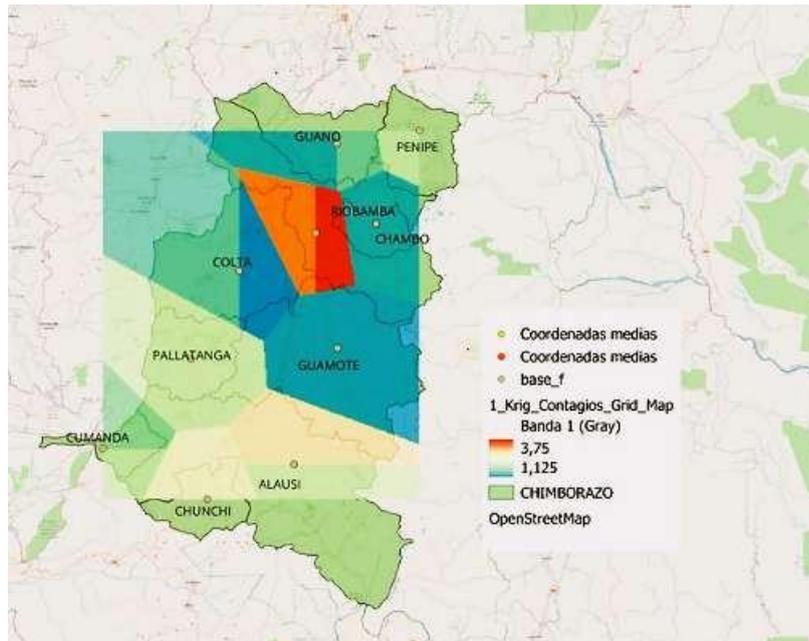


Ilustración 26–4: Krigging modelo esferico de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 26–4 se observa que existe un ligero cambio en una de las zonas donde se indica que los cantones con mayor número de contagios después de Riobamba corresponden a Pallantanga, Alausí, Cumandá, posiblemente Penipe, aun asi sigue siendo Riobamba la zona más caliente con el mayor número de registros, notoriamente se puede observar que las zonas frias corresponden a los cantones Chambo, parte de Guano, Guamote, Colta, se indican con el color azul intenso, las zonas amarillas reflejan la presencia de contagios como un valor notorio en los cantones de Chunchi y parte de Alausí esto nos permite identificar que existen casos covid 19 en todos los cantones, en alguno con un valor mínimo pero hubo presencia de contagios en toda la zona.

4.3.4.5. Modelación gaussiana

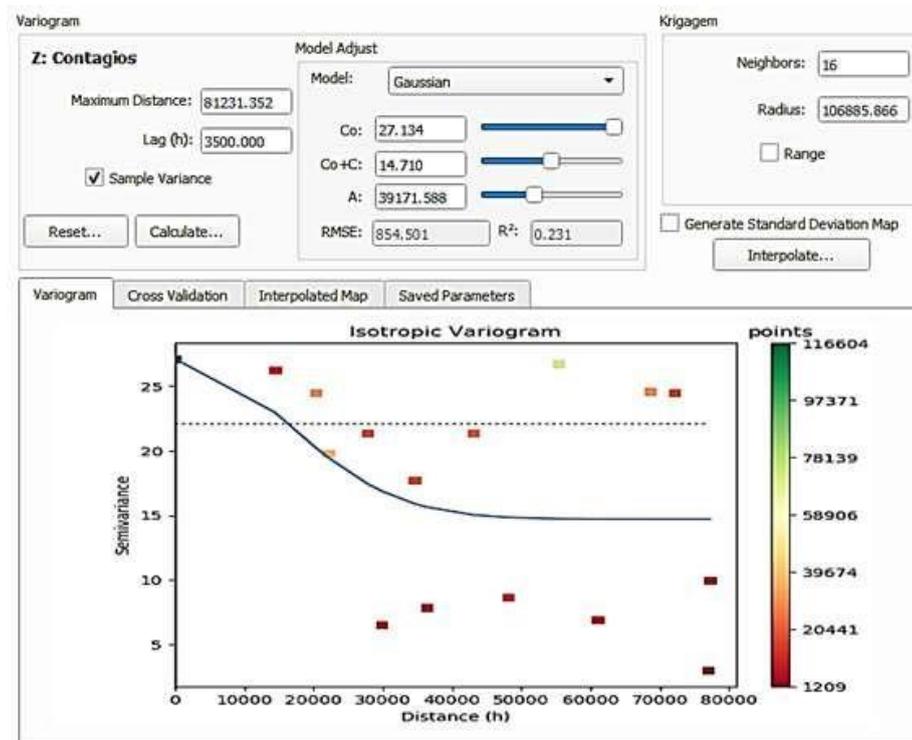


Ilustración 27-4: Modelo gaussiano de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 27-4 se observa el modelo gaussiano se puede denotar los puntos que representan a los pacientes contagiados mismos que se encuentran entre los valores de 0 a 25 que indican los contagios en los cantones con una distancia máxima de 81231.352 y una distancia lag(h) de 3500.00 la distancia máxima es una parte muy importante para representar h está considerada en base al valor encontrado del índice de Moran que cubre la cobertura del mapa estudiado dentro del variograma podemos ver que existe estacionalidad además con un valor $co = 27.134$ que indica el efecto nugget se observa un valor fuera del patrón general indicando la inexistencia de outliers, con un valor de $co+c=14.710$ con un rango de 39171.588 alcanza el umbral con un R^2 de 0.231 este valor nos puede indicar que es un buen modelo pero presenta estacionalidad sobre el umbral dado que se afirma que el modelo planteado no es muy representativo.

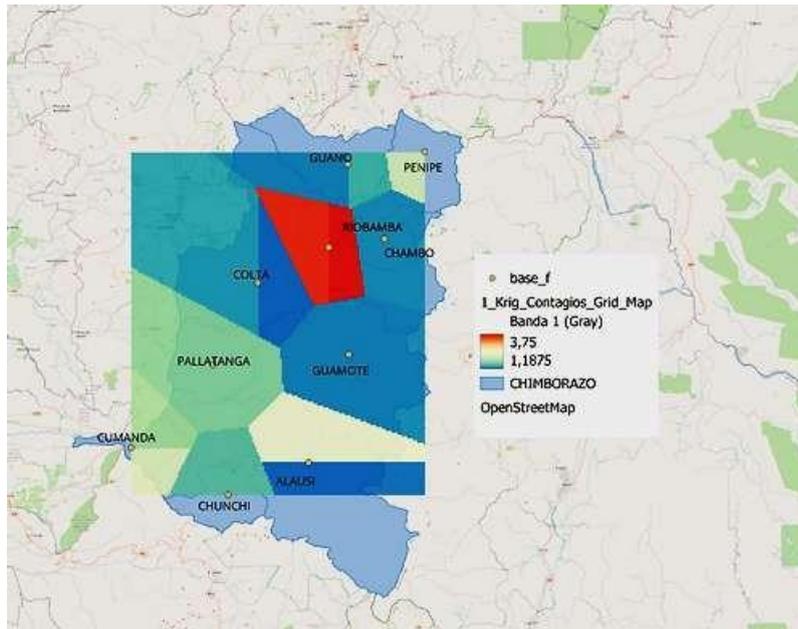


Ilustración 28–4: Krigging modelo gaussiano de los datos SARS-COV-2 (COVID-19)

Realizado por: Sayay Edison y Guzman Johana, 2023

En la ilustración 28–4 se observa que la mayor presencia de pacientes contagiados están en el cantón Riobamba como en los anteriores modelos a excepción del modelo esférico, en este tampoco se ve cambios relevantes de igual manera se identifica la zona más caliente con el mayor número de registros, notoriamente se puede observar que las zonas frías corresponden a los cantones Chambo, Guano, Guamote, Colta, se indican con el color azul intenso, las zonas verdes reflejan la presencia de contagios como un valor intermedio en los cantones de Pallatanga, Cumanda, Penipe, Chunchi, esto nos permite identificar que existen casos covid 19 en todos los cantones, en alguno con un valor mínimo pero hubo presencia de contagios en toda la zona, no existen otras particularidades que lo diferencien, obteniendo como resultado un comportamiento similar en casi todos los modelos.

4.4. Comparación de los modelos de variogramas, en función del umbral y el menor error cuadrático medio

Tabla 1-4: Valores aplicados en la modelación de variogramas, respecto a la meseta, rango y efecto pepita.

Modelo	Meseta	Rango(m)	Efecto pepita (Varianza)
Lineal	16.094	51484.524	23.440
Lineal con meseta	15.135	33106.213	27.134
Exponencial	13.408	77247.048	27.134
Esférico	15.033	45131.356	27.134
Gaussiano	14.710	39171.588	27.134

Fuente: Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

Los valores corresponden a los resultados de los modelos de variogramas, en todos los casos se trabajó con un valor de distancia definida, por lo tanto, se evidenció un valor de varianza poco similar en referencia a cada modelo, con un rango cambiante en cada uno, se obtuvo un mismo valor de umbral y error cuadrático medio, a excepción del modelo lineal presenta un valor efecto pepita más bajo, en los demás modelos se identifica valores similares, existe una diferencia entre los valores del rango estos aumentan o disminuyen dependiendo la variabilidad del modelo, el que mejor ajuste representa es el esférico, ejecutado el modelo se identificó graficamente que tiene un mejor corte en el umbral con un valor de varianza de 15.033 muestra el comportamiento de los datos con mejor precisión de representación de morbilidad de los casos covid-19.

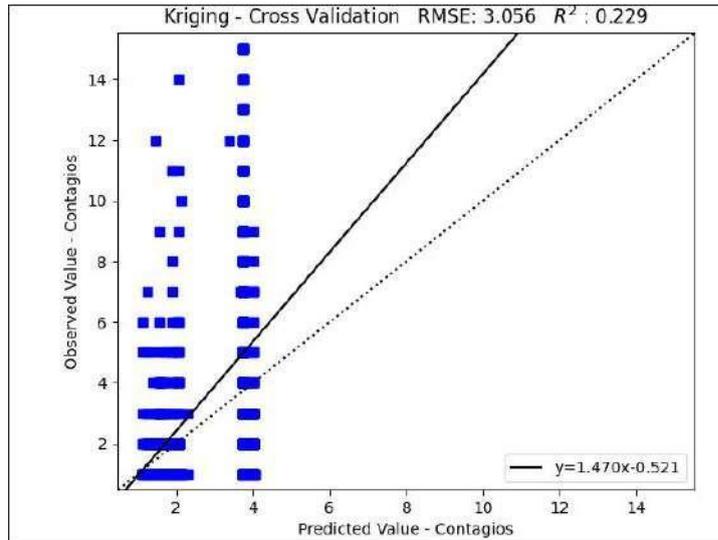


Ilustración 29-4: Validación cruzada del modelo de variograma esférico)

Realizado por: Sayay Edison y Guzman Johana, 2023

Realizado el proceso de validación, cruzada permitió evaluar la calidad y el rendimiento de un modelo de variograma, en la ilustración 30-4 se representa el valor predictivo del variograma esférico con un buen ajuste, esto es útil para estimar valores en lugares donde no se tienen datos, se representa de manera efectiva la estructura de correlación espacial con un RMSE de 3.056 lo cual implica que las diferencias entre los valores observados y predichos son pequeñas, con un R^2 de 0.229 cercano a 1 también nos indica un buen ajuste.

Tabla 2-4: Resultados de los modelos de variogramas, en función del umbral y el menor error cuadrático medio

Modelo	RMSE	R^2
Lineal	3.056	0.229
Lineal con meseta	3.056	0.229
Exponencial	3.056	0.229
Esférico	3.056	0.229
Gaussiano	3.056	0.229

Fuente: Datos de los casos SARS-COV-2 (COVID-19) de la provincia de Chimborazo en el período 2020-2021

Realizado por: Sayay Edison y Guzman Johana, 2023

Los resultados obtenidos del error cuadrático medio de los modelos aplicada la validación cruzada en cada uno, son un indicador para seleccionar un buen modelo de variograma, pero estudiadas las gráficas en este caso no podemos afirmar que el mejor modelo es el que tiene un menor valor de error cuadrático medio, realizado los ajuste correspondientes se selecciona el modelo esférico, que representa de una manera más objetiva el patrón de morbilidad de los casos SARS-COV-2 (COVID-19), dado que el RMSE esta en el rango de valoración aproximadamente bajo en todos los modelos, pese a que debería resultar al menos una mínima variación en alguno de ellos, por esa razón se procedió hacer un estudio gráfico en donde se nota la gran diferencia existente entre los modelos gaussiano y lineal con meseta, que también representan valores bajos al igual que el modelo esférico pero no tiene una buena precisión de varianza ni umbral, en el sentido de estacionalidad no presenta ninguna variación, por lo que descartamos la selección del modelo con menor error cuadrático medio.

El modelo esférico se observa en la ilustración 25-4 es el que mejor se acopla a los datos, puesto que alcanza el umbral, muestra un patrón característico en el que la semivarianza aumenta a medida que las distancias entre los puntos disminuyen, presentando una correlación negativa, pero se estabiliza cuando la distancia aumenta, los valores de morbilidad tienden a tener una correlación negativa esto puede deberse a los picos altos de contagios en el año 2020 para el año 2021 se redujo el número de contagios, lo cual significa que las diferencias entre ellos se reducen, hay una mayor similitud entre los puntos a distancias mayores en relación a las ubicaciones geográficas.

CONCLUSIONES

- A partir del análisis de los indicadores de autocorrelación espacial LISA se determinó que con un valor-p 0.001 calculado mediante el I de Moran resultó ser significativo, por lo tanto, existe autocorrelación espacial positiva, en donde se puede distinguir el predominio de los sectores con valores altos, agrupados mediante clúster, entre estos grupos están los cantones Colta, Guamote, Riobamba, Chambo, Guano con mayor número de pacientes covid 19, los resultados están ligados a una asociación espacial que indica similitud entre las unidades espaciales de estudio para cualquier subgrupo de la población, denotando relaciones de proximidad.
- Mediante el mapeo de geometrías puntuales se identificaron las zonas específicas de ubicaciones pacientes SARS-COV-2 (COVID-19), representada mediante una capa vectorial de puntos, lo cual permitió proceder a realizar el análisis exploratorio espacial de datos, en donde se ha identificado que existe mayor nivel de concentración de pacientes contagiados en ciertas zonas, aplicando los procedimientos necesarios se identificaron las zonas integradas por varios cantones en el siguiente orden zona 1, (Riobamba, Colta, Alausí), zona 2 (Guano, Chunchi), zona 3 (Guamote, Cumanda), zona 4 (Penipe, Chambo, Pallatanga) distribuidos desde los valores de picos de contagios más altos a los más bajos consecutivamente por lo tanto se concluye que resulta que los datos provienen de una distribución espacial agrupada para los casos de morbilidad que se dieron en los años 2020, 2021 en la provincia de Chimborazo.
- En el desarrollo del análisis estructural evidentemente se pudo observar que los datos están autocorrelacionados, presentando un patrón de localizaciones puntuales, efectuados a partir del análisis de autocorrelación espacial lisa en donde se indica que con un valor de índice de Moran de 0.041 y un p-valor estadísticamente significativo, podemos reiterar que los datos siguen un patrón de distribución agrupada dado que la distancia entre individuos vecinos se reduce al valor mínimo; en relación al modelamiento de variogramas se estima un buen ajuste del comportamiento gráfico, puesto que casi todos los modelos tienen el mismo comportamiento dada la estructura de variación espacial como la zona de influencia tanto de alcance como meseta presenta una variabilidad mínima entre los modelos, representando el comportamiento de los casos SARS-COV-2 (COVID-19) con una tendencia descendente esencialmente se puede identificar un componente de datos agrupado a través de nugget (efecto pepita) que detalla un mismo valor en todos los modelos.

- Comparados los modelos en función al umbral y el menor error cuadrático medio de variogramas, resultó que el modelo mas aceptable no precisamente es el que tiene un valor mas bajo en referencia al RMSE, en este caso se obtuvo el mismo valor predictivo para todos los modelos, con un valor de 3.056 y R^2 de 0.229 por lo tanto, no se consideró esta opción, gracias al estudio de las gráficas, correspondientes a los diferentes modelos y analizados los parámetros de meseta, alcance y nugget, se visualiza que el modelo mas aceptable es el esférico, dado que el rango alcanza la varianza y tiene una tendencia regionalizada descendente, evidentemente tiene buen corte en el umbral, resultando ser un modelo poco mas adecuado en base al estudio gráfico, para la representación de los datos, no tanto representativo dado que los valores de RMSE Y R^2 no fueron muy efectivos.

RECOMENDACIONES

- Es conveniente que en futuras investigaciones se pueda trabajar con un mejor registro de datos, en donde se tome muy en cuenta variables para estudios espaciales, primordialmente la ubicación de puntos específicos, dado que en este estudio se generaliza la ubicación por cantón, sería mucho mejor implementar el análisis referente a localizaciones mas precisas para obtener resultados mucho más efectivos.
- Se recomienda analizar los diferentes softwares que permiten emplear análisis espaciales, dado que existen diferentes herramientas para realizar el modelamiento y el ajuste de variogramas, utilice colores y tamaños de puntos según los atributos relevantes, esto permitirá visualizar patrones de distribución o áreas de alta concentración de casos.
- Utilice las herramientas de análisis de variogramas en QGIS que prueba los diferentes modelos de variogramas como esférico, exponencial y gaussiano para conseguir un buen ajuste de los parámetros en cada modelo que permita comparar sus gráficas visualmente, es importante explorar los diferentes modelos y sus características para poder seleccionar el más representativo en relación a los datos de estudio.
- Es indispensable la aplicación de análisis estadísticos espaciales se podría implementar para todo tipo de estudios y mas aun si es referente al conocimiento ciudadano es mucho mas dinámico y accesible al receptor en relación a la transmision de resultados.

BIBLIOGRAFÍA

ALAVA, Juan José., & GUEVARA, Angel. “A critical narrative of Ecuador’s preparedness and response to the COVID-19 pandemic”. *Public Health in Practice* [en línea], 2021, (Colombia), vol.2, pp. 9- 12. [consulta: 10 enero 2023]. ISSN 2666-5352. Disponible en: <https://www.sciencedirect.com/science/article/pii/S2666535221000525>.

ALVARADO, Welman. “Modelo geoestadístico espacio-temporal del crimen en el salvador: análisis estructural y predictivo”. *Revista de Matemática: Teoría y Aplicaciones* [en línea], 2011, (Salvador) vol. 18, p. 325-342. [consulta: 21 diciembre 2022]. Disponible en: <https://www.scielo.sa.cr/pdf/rmta/v18n2/a07v18n2>

AMIEL PÉREZ, José. “Las variables en el método científico”. *Revista de la Sociedad Química del Perú* [en línea], 2007, vol. 73, p. 171-177. [consulta: 3 julio 2023]. ISSN 1810-634X. Disponible en: http://www.scielo.org.pe/scielo.php?pid=S1810-634X2007000300007&script=sci_arttext&tlng=en

ANAYA COVARRUBIAS, Julio; et al. COVID-19 in Latin America and the caribbean region: Symptoms and morbidities in the epidemiology of infection. *Current Opinion in Pharmacology* [en línea], 2022, (Caribe), vol. 63, p. 1023. [consulta: 8 enero 2023]. ISSN 1471-4892. Disponible en: <https://www.sciencedirect.com/science/article/pii/S1471489222000303>.

BARCELLOS, Christovam. Unidades y escalas en los análisis espaciales en salud. *Revista Cubana de Salud Pública* [en línea] 2003, vol. 2, [consulta: 3 julio 2023]. Disponible en: http://scielo.sld.cu/scielo.php?pid=S0864-34662003000400003&script=sci_arttext

BARROSO, María. El contagio por Sars-CoV-2 en el medio laboral. Entre el accidente de trabajo y la enfermedad profesional. *Semergen* [en línea], 2021, vol. 47, p. 69. [consulta: 2 julio 2023]. ISSN 1138-3593. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7980089/>.

BAXENDALE, Claudia., & BUZAI, Gustavo. Análisis Espacial con Sistemas de Información Geográfica [en línea]. Aportes de la Geografía para la elaboración del Diagnóstico en el Ordenamiento Territorial. 2011. [consulta: 23 mayo 2023]. Disponible en: <https://ri.conicet.gov.ar/handle/11336/193310>

BOHÓRQUEZ, Ingrid; et al. “Algunos conceptos de la econometría espacial y el análisis exploratorio de datos espaciales”. Ecos de Economía: A Latin American journal of applied economics [en línea], 2008, vol. 12, p. 9-2. [consulta: 8 mayo 2023]. ISSN 2462-8107. Disponible en: <https://publicaciones.eafit.edu.co>.

BOLAÑOS, Paulina; et al. “Análisis espacial de contagios de Covid-19 en la Provincia de Chimborazo, a través de medidas de centralidad y dispersión”. Polo del Conocimiento [en línea], 2022, (Ecuador), vol. 7, p. 1035-1053. [consulta: 10 enero 2023]. ISSN 2550-682X. Disponible en: <https://polodelconocimiento.com/ojs/index.php/es/article/view/4272>.

BRAVO LÓPEZ, P. Revisión de literatura [en línea]. Autocorrelación Espacial e Índices para determinar su existencia en datos geográficos Universidad-Verdad 2021, [consulta: 22 mayo 2023]. ISSN 2600-5786, 1390-2849. Disponible en: <http://universidadverdad.uazuay.edu.ec/article/view/351>.

BUCHELI, Grace Estefanía. “Uso del Índice de Moran y LISA para explicar el ausentismo electoral rural en Ecuador”. Revista Geográfica [en línea], 2019, p. 91-108. [consulta: 4 junio 2023]. ISSN 2663-399X. Disponible en: <https://revistasipgh.org/index.php/regeo/article/view/746>.

BUFFALO, Luciana, et al. Dinámica territorial de pandemia COVID-19 en la provincia de Córdoba Argentina. [en línea]. Boletín de la Asociación de Geógrafos Españoles, 2021. [consulta: 16 abril 2023]. Disponible en: <https://bage.age-geografia.es/ojs/index.php/bage/article/view/3149>

CABRERA VARGAS, Patricio. Uso del variograma generalizado para inferencia de estructuras espaciales no-estacionarias [en línea]. 2017, [consulta: 20 julio 2023]. Disponible en: <https://repositorio.uchile.cl/handle/2250/145386>.

CABRERO ORTEGA, Yolanda; & GARCÍA PÉREZ, Alfonso. Análisis Estadístico de datos espaciales con QGIS Y R. Editorial UNED [en línea], 2022. S.l.: UNED - Universidad Nacional de Educación a Distancia. [consulta: 20 julio 2023]. ISBN 978-84-362-7839-2. Disponible en: <https://elibro.net/es/ereader/epoch/218566>.

CALDERÓN CISNEROS, Juan; et al. Análisis Espacial de la Distribución de la Delincuencia en Guayaquil. Tesis Doctoral. ESPOL. FCNM. [en línea]. Ecuador 2004. [consulta: 18 junio 2023]. Disponible en: <http://www.dspace.espol.edu.ec/handle/123456789/56659>.

CASCO, Samuel. Análisis de umbral como identificador de patrones delictuales en varios niveles geográficos del cantón Cuenca, Ecuador año 2013 al 2016. Memorias y Boletines de la Universidad del Azuay [en línea], 2017, p. 375-380, [consulta: 22 junio 2023]. Disponible en: <https://50.uazuay.edu.ec/index.php/memorias/article/view/84>

CELEMÍN, Juan. “Autocorrelación espacial e indicadores locales de asociación espacial: Importancia, estructura y aplicación”. Revista Universitaria de Geografía [en línea], 2009, vol. 18, no 1, p. 11-31. [consulta: 19 mayo 2023]. Disponible en: <http://www.scielo.org.ar/scielo/php>

CLAROS GIL, Adrián. Uso de datos estadísticos y herramienta QGIS [en línea]. Estudio estadístico de la criminalidad en las fiestas de la Magdalena. 2018. [consulta: 16 junio 2023]. Disponible en: <https://repositori.uji.es/xmlui/handle/10234/177072>

DABANCH, Jeannette. “ Emergencia de SARS-CoV-2, aspectos básicos sobre su origen, epidemiología, estructura y patogenia para clínicos”. Revista Médica Clínica Las Condes [en línea], 2021, vol. 32, p. 14-19. vol. 32, [consulta: 8 mayo 2023]. ISSN 0716-8640. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0716864020300924>

DEY, Lopamudra; et al. Machine learning techniques for sequence-based prediction of viral–host interactions between SARS-CoV-2 and human proteins. Biomedical journal, [en línea], 2020, vol. 43, p. 438-450. [consulta: 4 mayo 2023]. ISSN 2319-4170. Disponible en: <https://www.sciencedirect.com/science/article/pii/S2319417020301360>

ESCOBAR RIVERA, Valentina. Análisis espacial y temporal del Covid-19 en Chile. 2022. [en línea], 2022, (Chile), pp. 16-24. [consulta: 8 mayo 2023]. Disponible en: <http://repositoriobibliotecas.uv.cl/handle/uvsc1/3957>

GIRALDO, Ramon. Introducción a la geoestadística. [en línea], 2002. [consulta: 14 mayo 2023]. Disponible en: <https://geoinnova.org/wp-content/uploads/2021/08/LIBRO-GEOESTADISTICAA-R-Giraldo>

GRIFFITH, Daniel A. Spatial autocorrelation. Encyclopedia of social measurement. En: A. KOBAYASHI (ed.), International Encyclopedia of Human Geography (Second Edition) [en línea], Oxford: Elsevier, 2005, vol. 3, pp. 355-366. [consulta: 20 diciembre 2022]. ISBN 978-0-08-102296-2. Disponible en: <https://www.sciencedirect.com/science/article/pii/B9780081022955105967>

GUTIÉRREZ LÓPEZ, Alfonso; et al. “El variograma y el correlograma, dos estimadores de la variabilidad de mediciones hidrológicas”. Revista Facultad de Ingeniería Universidad de Antioquia, [en línea], 2011, p. 193-202. [consulta: 6 de enero 2023]. ISSN 0120-6230. Disponible en: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-62302011000300018

IBAÑEZ ASENSIO, Sara; et al. Coordenadas geográficas. 2010. [consulta: 18 enero 2023]. Disponible en: <https://riunet.upv.es/bitstream/handle/10251/8931/Coordenadas>

JUÁREZ, M.; & MONTES GALBÁN, E. Sistemas de Información Geográfica en Salud. [en línea], 2019. pp. 59-74. [consulta: 6 julio 2023]. ISBN 978-987-1548-95-8. Disponible en: https://www.researchgate.net/publication/333641990_Sistemasdeinformaciongeografica_en_Salud

KARUNASINGHA, Dulakshi. Root mean square error or mean absolute error use their ratio as well. Information Sciences [en línea], 2022, vol. 585, p. 609-629. [consulta: 20 diciembre 2022]. ISSN 0020-0255. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0020025521011567>

LEIVA GONZÁLEZ, César. Determinación de modelos de predicción espacial de la variable ondulación geoidal, para la zona urbana del cantón Quito y la zona rural del cantón Guayaquil, utilizando técnicas geoestadísticas. [en línea], (Quito), 2014. [consulta: 22 julio 2023]. Disponible en: <http://bibdigital.epn.edu.ec/handle/15000/8703>

LLOR, Carles., & MORAGAS, Ana. Coronavirus y atención primaria. Atención primaria [en línea], 2020, vol. 52, p. 294. [consulta: 2 julio 2023]. ISSN 0212-6567. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7135685/>

LÓPEZ, Javier. "What is the origin of SARS-CoV2?". Revista Médica del Instituto Mexicano del Seguro Social [en línea], 2020, vol. 58, p. 1-2 [consulta: 2 julio 2023]. Disponible en: <https://www.medigraphic.com/cgi-bin/new/resumenI.cgi?IDARTICULO=92771>.

LORENTE, Alvaro; et al. "El uso del lenguaje R para el procesamiento de datos y la generación de mapas sobre COVID-19". Revista Cubana de Transformación Digital [en línea], 2020, vol. 1, p. 37-50. [consulta: 10 junio 2023]. ISSN 2708-3411. Disponible en: <https://rctd.uic.cu/rctd/article/view/84>

LUCAY, Freddy; et al. Aplicación de Redes Neuronales Artificiales de Base Radial y Geoestadística para la Interpolación/Reconstrucción de Base de Datos de Leyes de Cobre [en línea]. Medio Ambiente y Minería, 2018. [consulta: 12 mayo 2023]. Disponible en: <https://bibliotecadigital.udea.edu.co/handle/10495/9413>

MARRUGO PETRO, Horacio; & MELO MAYORGA, Jorge. Propuesta análisis exploratorio de datos georreferenciables y temporales [en línea]. Caso de estudio: producción nacional del oro 2012 a 2020. 2020. [consulta: 2 agosto 2023]. Disponible en: <https://repository.libertadores.edu.co/handle/11371/3596>

MONCADA FUENTES, Celina. Análisis espacio-temporal de los accidentes de tránsito en Bogotá utilizando patrones puntuales [en línea]. Departamento de Estadística, 2018. [consulta: 4 junio 2023]. Disponible en: <https://repository.unal.edu.co/handle/unal/69640>.

MORENO LÓPEZ, Almudena. Producción BCN200: desarrollo del proceso de obtención de la base cartográfica nacional a escala 1: 200.000: detección, edición y resolución de conflictos cartográficos en la representación a escala. 2015. [en línea]. [consulta: 4 junio 2023]. Disponible en: <https://oa.upm.es/37322/>

MSP. Indicadores de salud Aspectos conceptuales y operativos. Ministerio de Salud Pública. [blog]. 2018. [consulta: 17 marzo 2023]. Disponible en: https://www.academia.edu/90841112/Covid19_y_la_pandemia_econ%C3%B3mica_global

NIÑO, Larry. Interpolación espacial de la abundancia larval de *Aedes aegypti* para localizar focos de infestación. Revista Panamericana de Salud Pública [en línea], 2011, vol. 29, p. 416-422. [consulta: 18 junio 2023]. ISSN 1020-4989. Disponible en: <https://www.scielosp.org/pdf/rpsp/v29n6/06.pdf>

OMS. Indicadores de salud Aspectos conceptuales y operativos [blog]. 2018. [consulta: 23 enero 2023]. Disponible en: https://www3.paho.org/hq/index.php?option=com_content&view=article&id=14405:health-indic

SANTANA JUÁREZ, M.; & MONTES GALBÁN, E. Sistemas de Información Geográfica en Salud [en línea]. Teoría y métodos de la Geografía Cuantitativa, 2019. [consulta: 21 enero 2023]. Disponible en: https://www.researchgate.net/publication/333641990_Sistemas_de_Informacion_Geografica_Salud

SAASTAMOINEN, U; et al. The limits of spatial data? Sense-making within the development and different uses of Finnish urban-rural classification. Land Use Policy [en línea], 2022, vol. 120, [consulta: 10 enero 2023]. ISSN 0264-8377. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0264837722002587>

SANTANA JUÁREZ.; & MONTES GALBÁN, E. Sistemas de Información Geográfica en Salud. GD Buzai, L. Humacata, SL Lanzelotti, E. Montes Galbán, y N. Principi, Teoría y métodos de la Geografía Cuantitativa, 2019, vol. 2, p. 59-74.

SIABATO, Willington; & GUZMÁN MANRIQUE, Jhon. “La autocorrelación espacial y el desarrollo de la geografía cuantitativa”. Cuadernos de Geografía: Revista Colombiana de Geografía [en línea], 2019, vol. 28, no 1, p. 1-22. [consulta: 23 julio 2023]. ISSN 0121-215X. Disponible en: http://www.scielo.org.co/scielo.php?script=sci_abstract&pid=S0121-215X2019000100001

SICILIA, Giuseppe.; & RIVERA, Maribel. Métodos gráficos de análisis exploratorio de datos espaciales con variables espacialmente distribuidas. Cuadernos latinoamericanos de administración [en línea], 2017, vol. 13, p. 92-104. [consulta: 18 agosto 2023]. ISSN 2248-6011. Disponible en: <https://revistacolombianadeenfermeria.unbosque.edu.co/index.php/cuaderlam/article/view/2417>

SIERRA, A. Herramientas geoestadísticas en la construcción de Modelos Digitales de Elevación [en línea]. (MDE) Ing. Agrim. Eduardo Sierra setiembre de PDF Descargar libre. 2014. [consulta: 22 julio 2023]. Disponible en: <https://docplayer.es/86731043-Herramientas-geoestadisticas-en-la-construccion-de-modelos>

SOHRABI, Catrin; et al. Impact of the coronavirus (COVID-19) pandemic on scientific research and implications for clinical academic training—a review. International Journal of Surgery, [en línea], 2021, vol. 86, p. 57-63. [consulta: 9 enero 2023]. ISSN 1743-9191. Disponible en: <https://www.sciencedirect.com/science/article/pii/S1743919121000029>.

SOTELA, Barrantes; & MAYORGA, Solano. Comportamiento espacial de la pandemia COVID-19 en Costa Rica durante los meses de marzo y abril de 2020 mediante un análisis de autocorrelación espacial. [en línea], 2020, vol. 3, p. 1-17. [consulta: 16 diciembre 2022]. Disponible en: https://www.researchgate.net/profile/Omar-Barrantes-Sotela/publication/341552796_espacial

STEIN, Alfred. The development of the journal Spatial Statistics: The first 10 years. Spatial Statistics [en línea], 2022, vol. 50, pp. 324 - 576. [consulta: 20 diciembre 2022]. ISSN 2211-6753. Disponible en: <https://www.sciencedirect.com/science/article/pii/S2211675321000750>

TECHNOLOGIES, R. Análisis de patrones espaciales con QGIS. Geoinnova [en línea]. 2020, [consulta: 3 mayo 2023]. Disponible en: <https://geoinnova.org/blog-territorio/analisis-de-patrones-espaciales-con-qgis/>

VELÁZQUEZ, Alberto. Diplomado en análisis de información geoespacial. Centro de Investigación en Geografía y Geomática [en línea]. 2017, p. 1-14. [consulta: 20 diciembre 2022]. Disponible en: <https://centrogeo.repositorioinstitucional.mx/jspui/bitstream/1012/160/1/16-M%C3%A9todo%20>

VIEIRA, Eliane; et al. Variabilidad espacial de la concentración de nitratos en el embalse de nova ponte, Minas Gerais, Brasil, por medio de la geoestadística y los sistemas de información geográfica. Aqua-LAC [en línea], 2015, vol. 7, p. 49-55. [consulta: 22 julio 2023]. ISSN 1688-2873. Disponible en: <https://aqua-lac.org/index.php/Aqua-LAC/article/view/135>

VILCA, Carla., & MOYANO, Emma. Análisis exploratorio de datos espaciales de pacientes con rinofaringitis, asma y bronquitis en consulta externa del Hospital Santa Rosa [en línea]. 2019, [consulta: 18 junio 2023]. Disponible en: <https://revistasinvestigacion.unmsm.edu.pe/index.php/matema/article/download/16127/14018/56385>

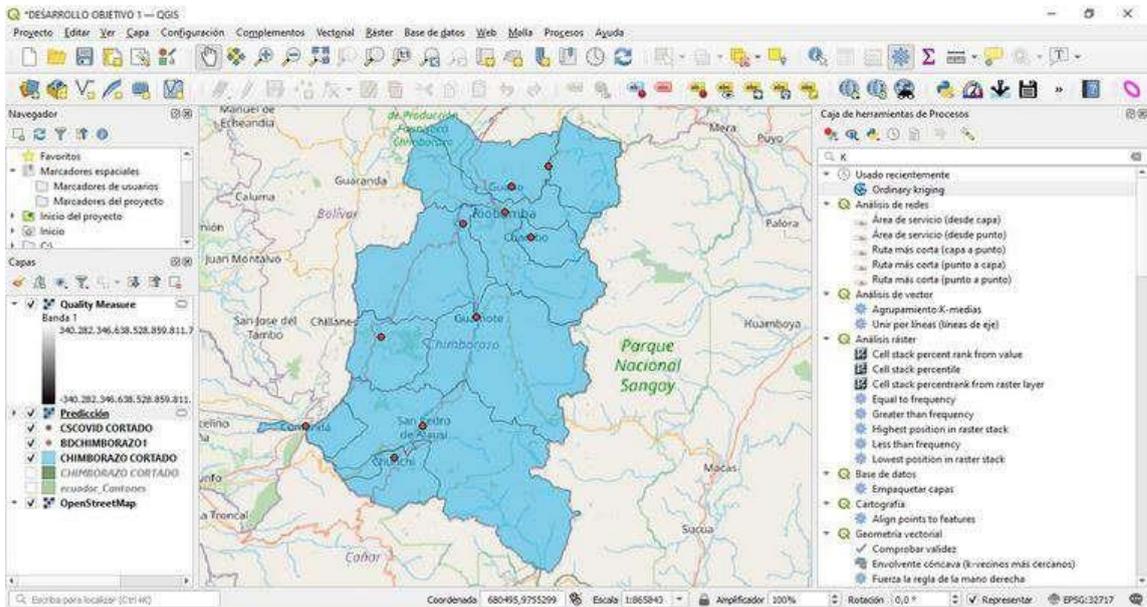
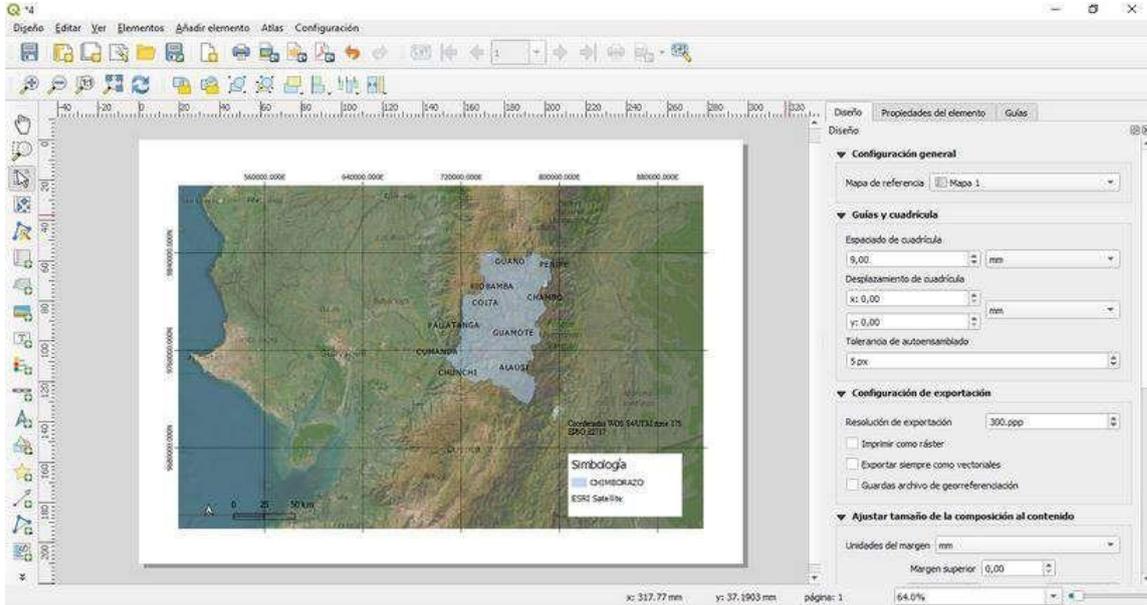
VIRAMONTES, José., & HERNÁNDEZ, Miriam. “Análisis exploratorio espacial de las transferencias en México”. Horizontes Empresariales [en línea], 2021, vol. 20, p. 49-66. [consulta: 18 junio 2023]. ISSN 0719-0875. Disponible en: <https://revistas.ubiobio.cl/index.php/HHEE/article/view/4824>.

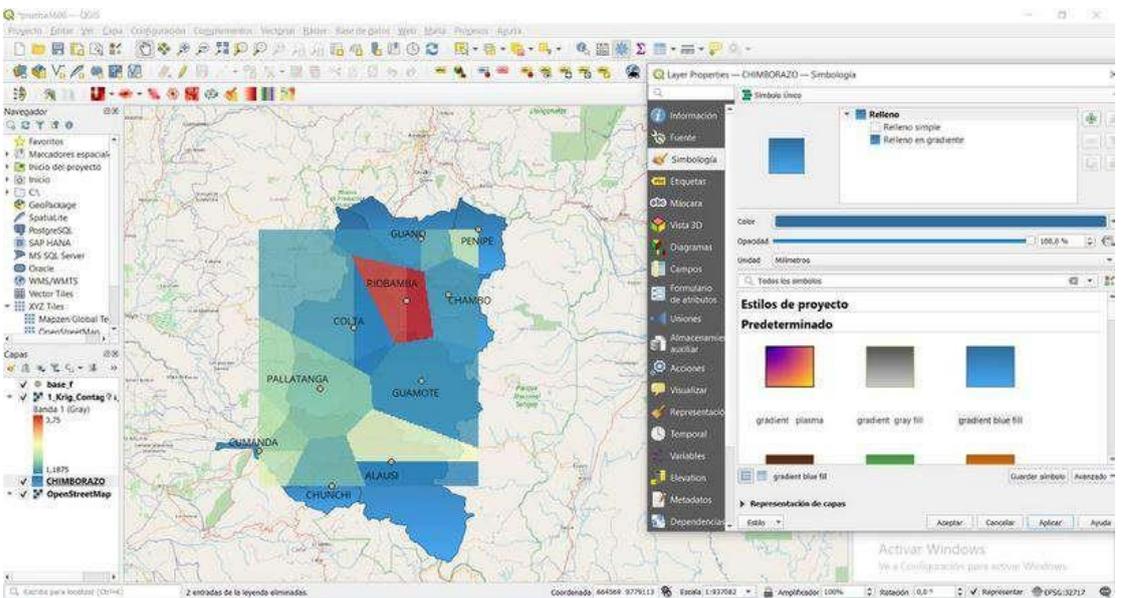
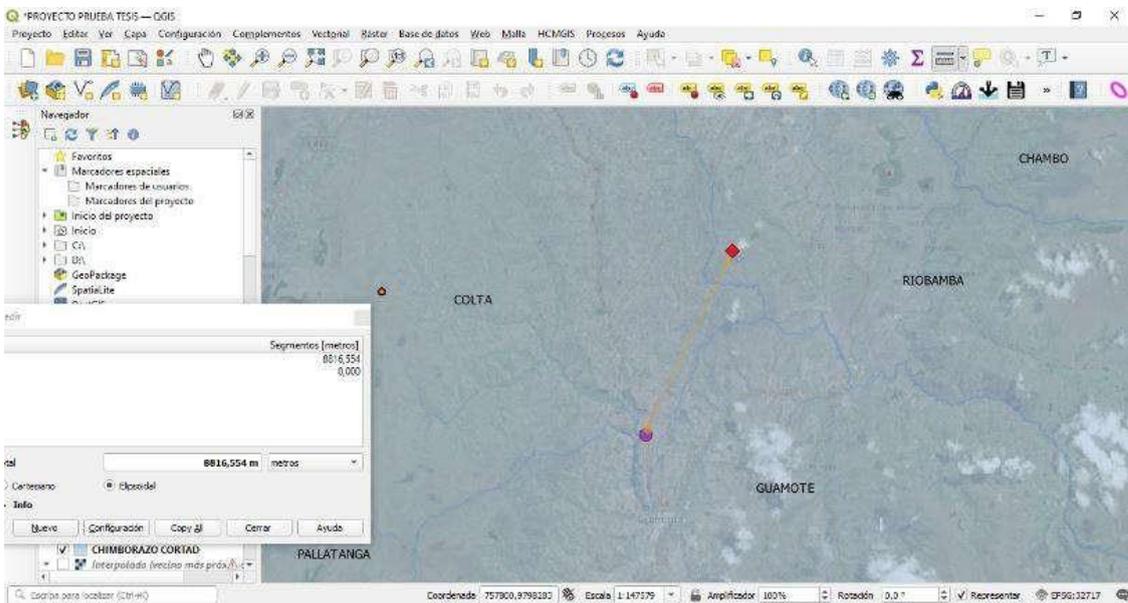
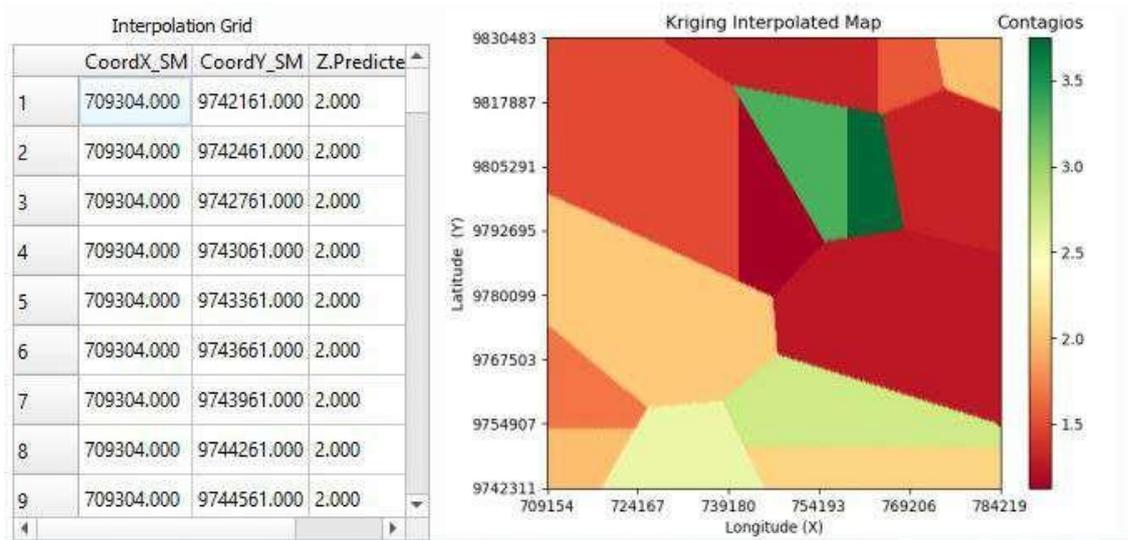
ZAPATA BEDOYA, Silvana; et al. Modelos geoespaciales para control de brotes de SARS-CoV-2 en Cartagena y Barranquilla, Colombia, 2020. Revista Panamericana de Salud Pública [en línea], 2023, vol. 46, p. e26. [consulta: 20 diciembre 2022]. ISSN 1020-4989, 1680-5348. Disponible en: <https://iris.paho.org/handle/10665.2/55889>



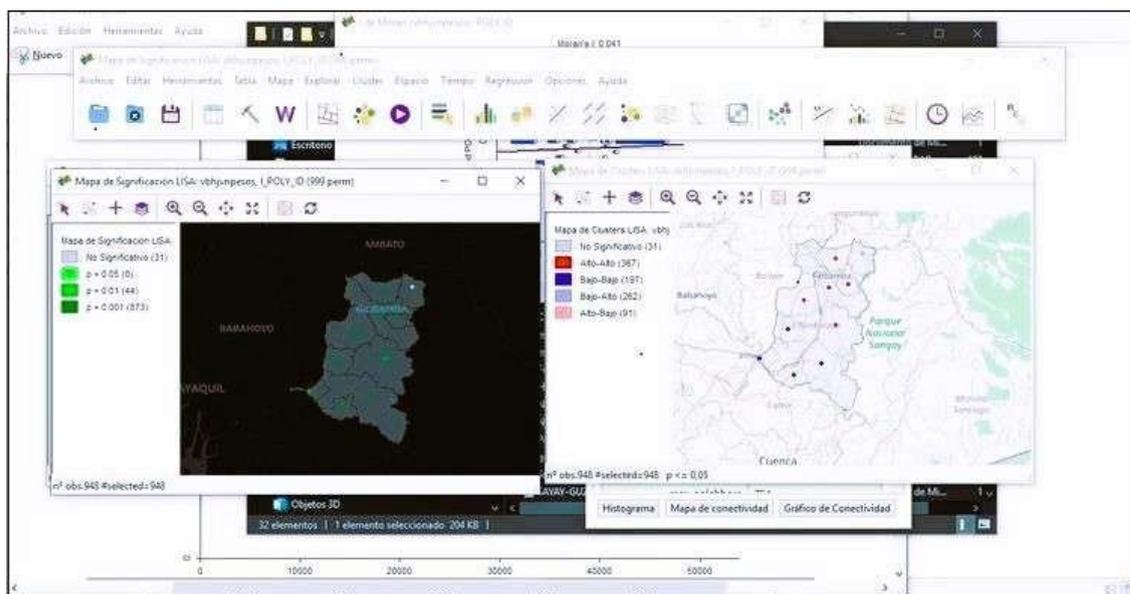
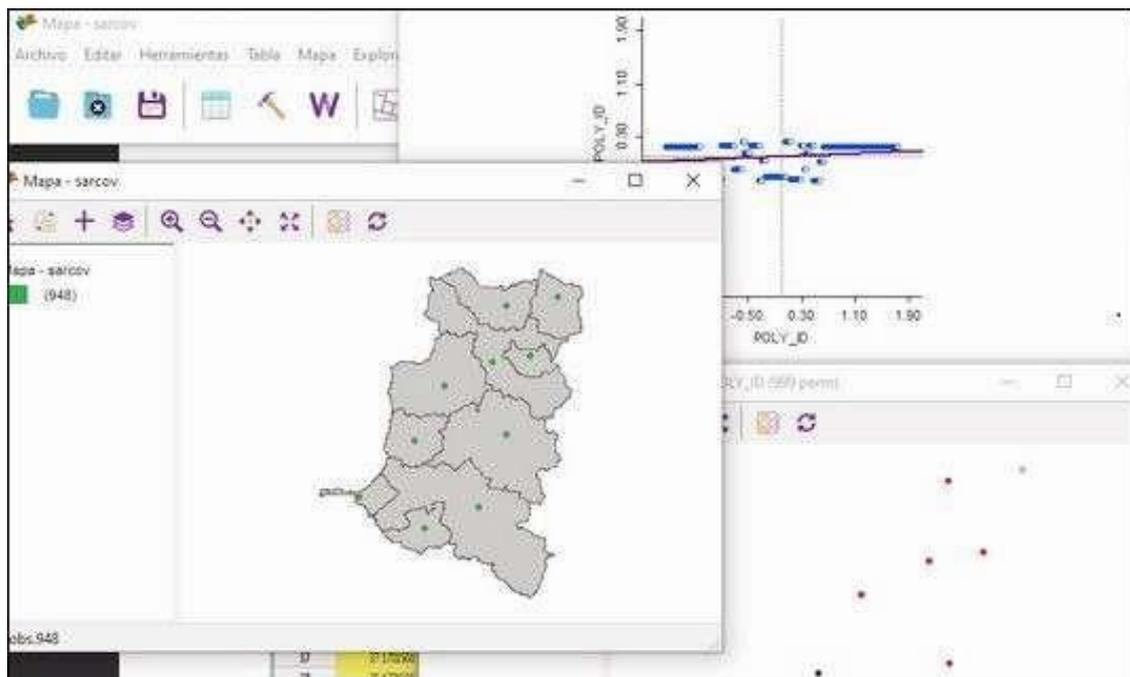
ANEXOS

ANEXO A: PROYECTO QGIS

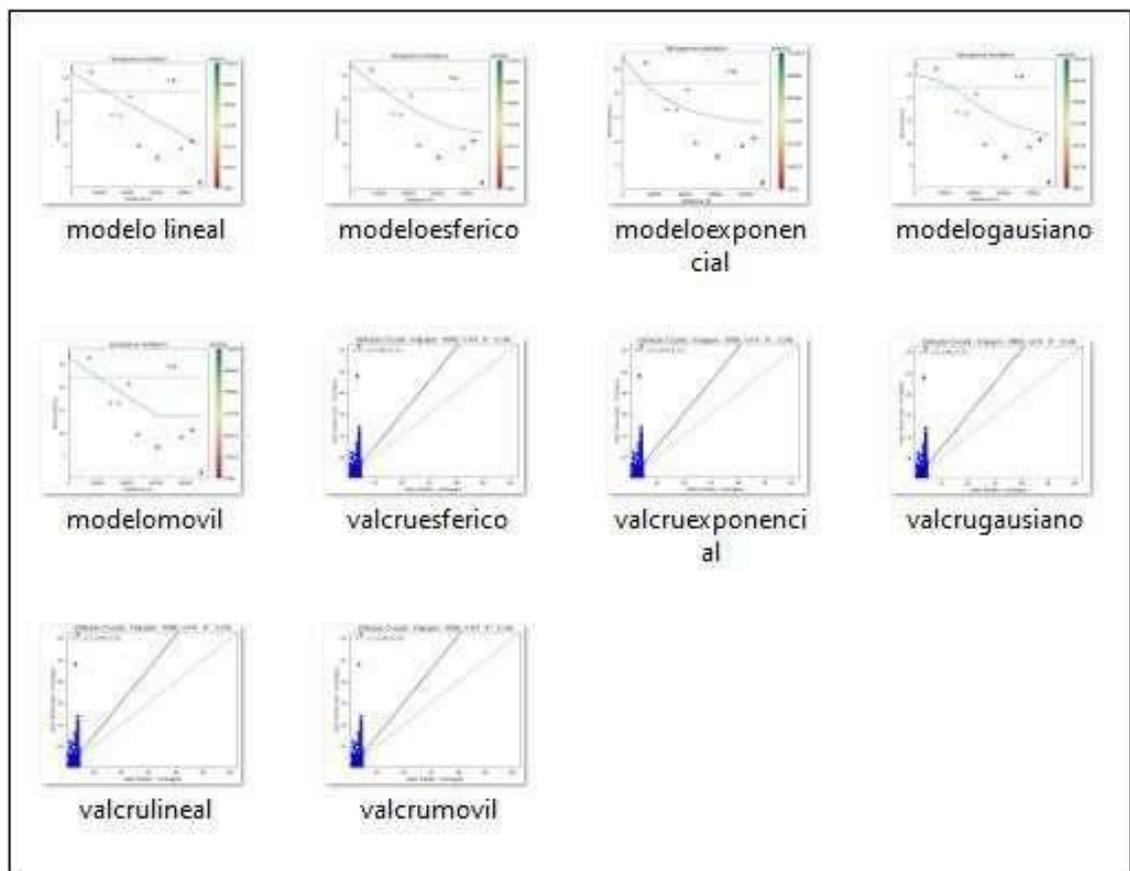
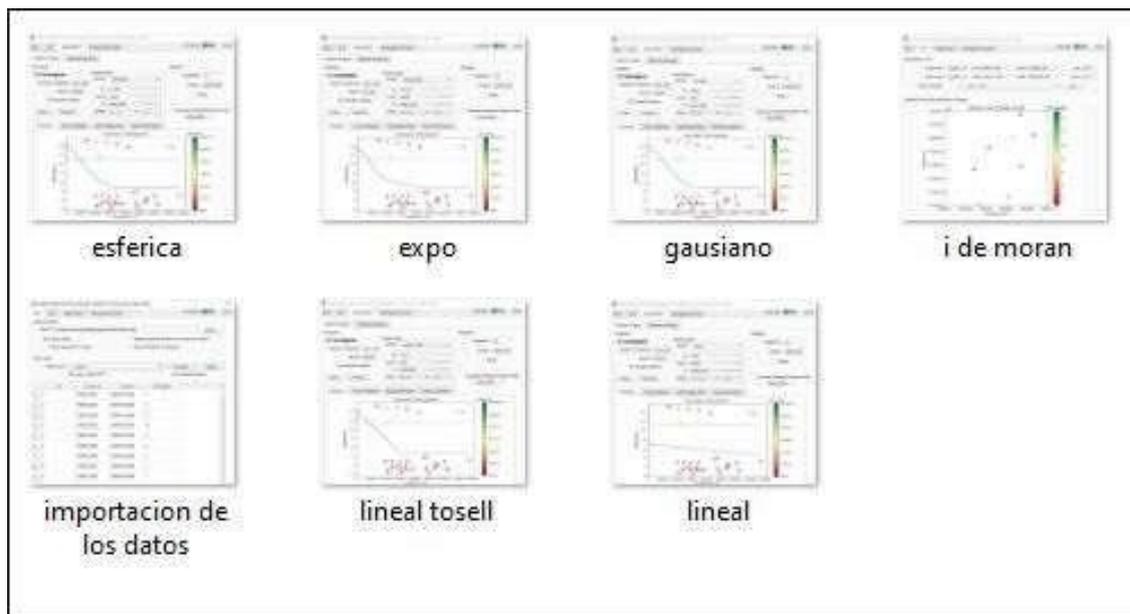




ANEXO B: PROYECTO GEODA



ANEXO C: PRUEBAS DE VARIOGRAMAS



ANEXO D: CÓDIGO DE R

```
1 #####
2 ##### ANaLISIS DESRIPTIVO
3 -----
4
5 datos <- read.xlsx("basedatos.xlsx",na.strings = T,rowNames = F)
6 View(datos)
7 ##Variable Condición Final del Paciente
8 datos<-filter(datos,clasificacion_final=="CONFIRMADO")
9 Tabla <- datos %>% group_by(condicion_final) %>%
10   summarise(Total=n()) %>%
11   dplyr::mutate(Porcentaje = round(Total/sum(Total)*100, 1))
12 ggplot(Tabla, aes(x = condicion_final, y=Total,
13   fill=condicion_final) ) +
14   geom_bar(width = 0.9, stat="identity",
15   position = position_dodge())+
16   #ylim(c(0,7500))+
17   labs(x="Pacientes", y= "Frecuencia")+
18   labs(fill = "")+
19   geom_text(aes(label=paste0(Porcentaje,"%", "", "\n", "", "", "")),
20   vjust=0.05,
21   color="black",
22   hjust=0.5,
23   position = position_dodge(0.9),
24   angle=0,
25   size=4.0) +
26   scale_y_continuous(limits=c(0,7000),
27   breaks=seq(0,7000,600))+
28   scale_fill_discrete(name = "Condición Final del Paciente",
29   labels =
30   c("Muerto", "Vivo"))+ theme(axis.text.x =
31   element_text(angle = 0, vjust = 0,
32   hjust=0.5)) +
33   #theme_linedraw() +
34   scale_fill_manual(name="Condición Final del Paciente",
35   values=c("snow4", "plum3"))
36
37 #values=c("turquoise3", "palegreen2"))
38
```

```

39 #####
40 ##Variable Sexo
41 Tabla <- datos %>% group_by(sexo_paciente) %>%
42   summarise(Total=n()) %>%
43   dplyr::mutate(Porcentaje = round(Total/sum(Total)*100, 1))
44 ggplot(Tabla, aes(x = sexo_paciente, y=Total,
45                 fill=sexo_paciente) ) +
46   geom_bar(width = 0.9, stat="identity",
47           position = position_dodge()+
48   #ylim(c(0,7500))+
49   labs(x="Pacientes", y= "Frecuencia") +
50   labs(fill = "")+
51   geom_text(aes(label=paste0(Porcentaje,"%", "", "\n", "", "")),
52             vjust=0.05,
53             color="black",
54             hjust=0.5,
55             position = position_dodge(0.9),
56             angle=0,
57             size=4.0) +
58   scale_y_continuous(limits=c(0,7000),breaks=seq(0,7000,600))+
59   scale_fill_discrete(name = "Sexo del Paciente",
60                      labels = c("Mujer", "Hombre"))+
61   theme(axis.text.x = element_text(angle = 0, vjust = 0,
62                                   hjust=0.5)) +
63   #theme_linedraw() +
64   #facet_grid(~"Diagrama de barras para la variable Sexo")+
65   scale_fill_manual(values=c("deepskyblue3", "pink"))
66
67 #####
68 ### Variable Cant n Contagio
69 Tabla <- filtrado %>% group_by(canton) %>% summarise(Total=n()) %>%
70   dplyr::mutate(Porcentaje = round(Total/sum(Total)*100, 1))
71 ggplot(Tabla, aes(x = canton, y=Total, fill=canton) ) +
72   geom_bar(width = 0.9, stat="identity",
73           position = position_dodge()+
74   #ylim(c(0,5000))+
75   labs(x="Pacientes", y= "Frecuencia") + labs(fill = "")+
76   geom_text(aes(label=paste0(Porcentaje,"%", "", "\n", "", "", "")),
77

```

```

78         vjust=0.01,
79         color="black",
80         hjust=0.5,
81         position = position_dodge(0.9),
82         angle=0,
83         size=3.0) +
84 #theme_linedraw() +
85 scale_y_continuous(limits=c(0,5000),breaks=seq(0,5000,600))+
86 theme(axis.text.x = element_text(angle = -90, vjust = 0,
87                                     hjust=0))+
88 scale_fill_manual(name="Datos por Cant n ",
89                   values=c("cadetblue2", "darkseagreen",
90                             "pink3", "2", "3", "4", "5", "6", "olivedrab2", "8"))
91
92
93
94 #####
95 ##Variable Nmero de Contagios en Funci n del Tiempo
96 fechas1<-table(filtrado$fecha_atencion)
97 fechas<-as.data.frame(fechas1)
98 fechas$Var1<-as.Date(fechas$Var1)
99 f <- cut(fechas$Var1, "month")
100 a o<-table(f)
101 a os<-as.data.frame(a o)
102 ggplot(a os, aes(x = f, y = Freq, group=1)) +
103   geom_line() +
104   geom_point(size=2)+
105   theme(axis.text.x = element_text(angle = -90, vjust = 1,
106                                       hjust=-0.8))+
107   labs(x="Fecha de Ingreso", y= "Pacientes Contagiados")
108 #facet_grid(~"Serie de Tiempo de Pacientes Contagiados")+
109
110 #####
111 ##Variable Edad
112 Tabla <- filtrado %>% group_by(edad_quinquenal) %>%
113   summarise(Total=n()) %>%
114   dplyr::mutate(Porcentaje = round(Total/sum(Total)*100, 1))
115 ggplot(Tabla, aes(x = edad_quinquenal, y=Total,
116                   fill=edad_quinquenal) ) +

```

```

117 geom_bar(width = 0.9, stat="identity",
118           position = position_dodge()+
119 #ylim(c(0,1000))+
120 labs(x="", y= "Frecuencia") +
121 labs(fill = ""))+
122 geom_text(aes(label=paste0(Porcentaje,"%", "", "\n", "", "")),
123           vjust=0.01,
124           color="black",
125           hjust=0.5,
126           position = position_dodge(0.9),
127           angle=0,
128           size=3.0) +
129 scale_y_continuous(limits=c(0,1000),breaks=seq(0,1000,100))+
130 scale_fill_discrete(name = "Edad del Paciente")+
131 theme(axis.text.x = element_text(angle = -90, vjust = 0,
132                                   hjust=0))
133 #facet_grid(~"Diagrama de barras para la variable Edad" ) }
134
135 #####
136 ##Variable Edad Sexo y Canton
137 data<-table(filtrado$sexo_paciente, filtrado$canton) data<-as.data.frame(data)
138 datos <- data.frame(nombre=nombre, tipo=tipo, valor=valor) ggplot() +
139   geom_bar(data=data,aes(x=Var2, y=data$Freq,fill=Var1), stat='identity', position='dodge')
140   +
141   coord_flip()+
142   scale_fill_manual(values=c("deepskyblue1", "lightpink2"))+ labs(x="Cant n", y=
143   "Frecuencia")+ guides(fill=guide_legend(title = "Sexo"))
144
145 ##### #
146 FRECUENCIA DE FALLECIDOS DE ACUERDO A LA EDAD #
147 Tabla <- datos %>% group_by(edad_quinquenal) %>% summarise(Total=n()) %>%
148   dplyr::mutate(Porcentaje = round(Total/sum(Total)*100, 1)) ggplot(Tabla, aes(x =
149 edad_quinquenal, y=Total,
150                               fill=edad_quinquenal) ) +
151   geom_bar(width = 0.9, stat="identity",
152
153
154
155

```

```

156         position = position_dodge()+
157 #ylim(c(0,100))+
158 labs(x="Edad del fallecido", y= "Frecuencia ") +
159 labs(x="Pacientes", y= "Frecuencia")+
160 labs(fill = "")+
161 geom_text(aes(label=paste0(Porcentaje,"% ", "", "\n", "", "", "")),
162           vjust=0.01,
163           color="black",
164           hjust=0.5,
165           position = position_dodge(0.9),
166           angle=0,
167           size=3.0) +
168 labs(fill = "")+
169 scale_fill_discrete(name = "Edades")+
170 theme(axis.text.x = element_text(angle = -90, vjust = 0,
171                                   hjust=0))+
172 scale_y_continuous(limits=c(0,120),breaks=seq(0,120,10))+
173 scale_fill_manual(values=c("dodgerblue4", "hotpink3",
174                             "peachpuff1", "slategray4", "gold3", "4",
175                             "5", "lightpink3", "olivedrab2", "green4",
176                             "lavenderblush4", "turquoise4"))
177
178 #####
179 # FRECUENCIA DE FALLECIDOS DE ACUERDO AL CANTON DE REGISTRO DEL CASO #
180
181 Tabla <- datos %>% group_by(canton) %>% summarise(Total=n())
182 %>%
183 dplyr::mutate(Porcentaje = round(Total/sum(Total)*100, 1)) ggplot(Tabla, aes(x = canton,
184 y=Total, fill=canton) ) +
185 geom_bar(width = 0.9, stat="identity", position =
186           position_dodge()+
187 #ylim(c(0,400))+
188 labs(x=" ", y= "Frecuencia") +
189 labs(fill = "")+
190 geom_text(aes(label=paste0(Porcentaje,"% ", "", "\n", "", "", "")),
191           vjust=0.01,
192           color="black",
193           hjust=0.5,
194           position = position_dodge(0.9),

```

```

195         angle=0,
196         size=3.0) +
197     theme(axis.text.x = element_text(angle = -90, vjust = 0,
198                                     hjust=0))+
199     scale_y_continuous(limits=c(0,420),breaks=seq(0,420,40))
200
201 #####
202 # FRECUENCIA DE FALLECIDOS DE ACUERDO AL SEXO #
203
204 Tabla <- datos %>% group_by(sexo_paciente) %>%
205     summarise(Total=n()) %>%
206     dplyr::mutate(Porcentaje = round(Total/sum(Total)*100, 1))
207 ggplot(Tabla, aes(x = sexo_paciente, y=Total,
208                 fill=sexo_paciente) ) +
209     geom_bar(width = 0.9, stat="identity",
210             position = position_dodge()+
211             ylim(c(0,500))+
212             labs(x=" ", y= "Frecuencia") +
213             labs(fill = ""))+
214     geom_text(aes(label=paste0( Porcentaje, "%")),
215             vjust=0.01,
216             color="black",
217             hjust=0.5,
218             position = position_dodge(0.9),
219             angle=0,
220             size=4.0) +
221     scale_fill_discrete(name = "Sexo",
222                        labels = c("Hombre", "Mujer"))+
223     theme(axis.text.x = element_text(angle = 0,
224                                     vjust = 0.5, hjust=0.5))+
225     scale_y_continuous(breaks=seq(0, 400, 40))

```



esPOCH

**Dirección de Bibliotecas y
Recursos del Aprendizaje**

**UNIDAD DE PROCESOS TÉCNICOS Y ANÁLISIS BIBLIOGRÁFICO Y
DOCUMENTAL**

REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y BIBLIOGRAFÍA

Fecha de entrega: 12/ 01 / 2024

INFORMACIÓN DEL AUTOR/A (S)
Nombres – Apellidos: JOHANA ELIZABETH GUZMAN LEON EDISON RIGOBERTO SAYAY TENELEMA
INFORMACIÓN INSTITUCIONAL
Facultad: CIENCIAS
Carrera: ESTADÍSTICA
Título a optar: INGENIERA ESTADÍSTICA INGENIERO ESTADÍSTICO
f. Analista de Biblioteca responsable: Ing. Rafael Inty Salto Hidalgo

2238-DBRA-UPT-2023

