



**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

**Facultad de Ciencias Naturales y Matemáticas**

“Proyecciones espacio-temporal del virus SARS-CoV-2, en las principales provincias del Ecuador, a través de modelos Geoestadísticos y series de Tiempo”

**PROYECTO INTEGRADOR**

Previo la obtención del Título de:

**INGENIERO EN ESTADÍSTICA**

Presentado por:

Melissa Cando Velasco

Rubén Gavilanes Gavilanes

**GUAYAQUIL - ECUADOR**

Año: 2022

## **DEDICATORIA**

En primer lugar, al Creador que nos dio la oportunidad de tener las facultades intelectuales para desarrollarnos. A nuestras familias y amigos que nos acompañaron en este lapso de vida Universitaria.

## **AGRADECIMIENTOS**

Al Creador por su guía y luz que nos a llenado de conocimiento durante en cada etapa de nuestras vidas, desde que estuvimos pequeños hasta esta instancia tan importante de nuestra vida.

A nuestros padres y nuestros hermanos que nos han servido de apoyo y consolación en tiempos difíciles.

A nuestra amada alma mater ESPOL que nos ha proporcionado de abundante conocimiento durante cada ciclo, adicional a conocer a nuestros amigos y maestros, en especial a nuestro tutor Omar Ruiz, el cual nos ha brindado su apoyo incondicional durante el desarrollo de este proyecto.

## DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, nos corresponde conforme al reglamento de propiedad intelectual de la institución; Melissa Alejandra Cando Velasco y Ruben Dario Gavilanes Gavilanes damos nuestro consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”

  
Melissa Alejandra Cando V.

  
Ruben Dario Gavilanes G.

# EVALUADORES

---

**PhD. Sandra García Bustos**

PROFESOR DE LA MATERIA

---

**PhD. Omar Ruiz Barzola**

PROFESOR TUTOR

## RESUMEN

En el Ecuador, como en todo el mundo, ha sido afectado por la COVID-19, lo que puso en emergencia los sistemas de salud y a sus gobernantes. Ante este escenario, obtener predicciones de los casos de contagios se convirtió en un factor muy importante para tomar decisiones. Para predecir el nivel de contagios se utilizaron, series de tiempo, entre ellas los modelos VAR, además de modelos geoestadísticos de espacio- tiempo en los que se incluye el nivel de contagio en las principales provincias del Ecuador. Se utilizaron datos proporcionados por las infografías del Instituto Nacional de Estadísticas y Censos de Ecuador, en el período de abril del 2020 a abril del 2022. El procesamiento de los datos se realizó con el Software R-studio, versión: 5.1.0. Los modelos Geoestadísticos son satisfactorios gracias a la proporción de los datos geoespaciales. Por tal motivo, se aplicó la metodología Kriging Simple, una técnica Geoestadística, para la modelización y la estimación de los pronósticos de contagios a nivel provincial, utilizando el lenguaje de programación R, aplicado en las principales provincias del Ecuador.

**Palabras claves:** COVID-19; pandemia; modelos VAR; Geoestadística.

## ABSTRACT

In Ecuador, as in the whole world, it has been affected by COVID-19, which put the health systems and their rulers in an emergency. Given this scenario, obtaining predictions of contagion cases became a very important factor in making decisions. To predict the level of contagion, time series, geostatistical space-time models were used, which include the level of contagion in the main provinces of Ecuador. data provided by the infographic of the National Institute of Statistics and Census of Ecuador are used, in the period from April 2020 to April 2022. The data processing was carried out with the R studio Software, version: 5.1.0. Geostatistical models are satisfactory thanks to the proportion of geospatial data. For this reason, the Kriging Simple methodology was applied, a Geostatistical technique, to model and estimate contagion forecasts at the provincial level, using the R programming language, applied in the main provinces of Ecuador.

**Key words:** COVID-19; pandemic; VAR model; Geostatistical.

## ÍNDICE GENERAL

RESUMEN .....	6
ABSTRACT .....	7
ÍNDICE GENERAL .....	8
ABREVIATURAS.....	10
SIMBOLOGÍA.....	11
ÍNDICE DE FIGURAS .....	12
ÍNDICE DE TABLAS .....	13
CAPÍTULO 1 .....	14
1. Introducción.....	14
1.1 Descripción del problema.....	15
1.2 Justificación del problema.....	17
1.3 Objetivos .....	17
1.3.1 Objetivo General .....	17
1.3.2 Objetivos Específicos.....	17
1.4 Marco Teórico .....	18
CAPÍTULO 2 .....	24
2. METODOLOGÍA.....	24
2.1 Selección de variables de estudio.....	24
CAPÍTULO 3 .....	27
3. Resultados Y ANÁLISIS .....	27
CAPÍTULO 4 .....	35
4. CONCLUSIONES Y RECOMENDACIONES.....	35
4.1 Conclusiones.....	35



4.2	Recomendaciones.....	36
	BIBLIOGRAFÍA .....	37

## ABREVIATURAS

OPS	Organización Panamericana de la Salud
OMS	Organización Mundial de la Salud
INEC	Instituto Nacional de Estadísticas y Censos
ARIMA	Autoregressive Integrated Moving Average
VAR	Vectores Auto-Regresivos
SARS	Síndrome Respiratorio Agudo Severo
ICTV	Comité Internacional de Taxonomía de Virus
ETL	Extrac Transform Load

## SIMBOLOGÍA

$X$	Matriz de datos
$\gamma$	Semivarianza
$\hat{Z}$	Interpolación de kriging
$C_0$	Nugget o pepita
$C_1$	Psill o la meseta
$a$	Rango y el ángulo de tolerancia
$\phi$	Ángulo de tolerancia

## ÍNDICE DE FIGURAS

<b>Gráfico 3. 1 Mapeo y nube de puntos del semivariograma dentro Ecuador.....</b>	<b>27</b>
<b>Gráfico 3. 2 Nube de puntos del semivariograma y dirección.....</b>	<b>28</b>
<b>Gráfico 3. 3 Ajuste de los semivariograma direccionales.....</b>	<b>28</b>
<b>Gráfico 3. 4 Interpolación de Kriging para los contagios semanales .....</b>	<b>30</b>
<b>Gráfico 3. 5 Correlación entre valores observados y las estimaciones.....</b>	<b>31</b>
<b>Gráfico 3. 6 Series temporales de parámetros de los semivariogramas. ....</b>	<b>32</b>
<b>Gráfico 3. 7 Pronósticos del modelo VAR de cada uno de los parámetros. ....</b>	<b>32</b>
<b>Gráfico 3. 8 Correlación resultante de nuevos ajustes Kriging.....</b>	<b>33</b>

## ÍNDICE DE TABLAS

<b>Tabla 3. 1 Parámetros de cada modelo ajustado a los semivariogramas semanales. ....</b>	<b>29</b>
<b>Tabla 3. 2 Parámetros Pronosticados para 5 semanas.....</b>	<b>33</b>
<b>Tabla 3. 3 Contagios pronosticados dado el nuevo ajuste kriging .....</b>	<b>34</b>

# CAPÍTULO 1

## 1. INTRODUCCIÓN

En marzo del 2020 todo cambió radicalmente en el Ecuador y el mundo, el COVID-19 que afecta al ser humano cuando se infecta por el síndrome respiratorio agudo severo (SARS-CoV-2) de la familia Coronavirus, empezó a propagarse por el mundo entero, y más tarde fue declarado pandemia por la Organización Mundial de la Salud (OMS). El número de nuevos casos por COVID-19 en todo el mundo alcanzó los 5.5 millones de casos y generó más de 9000 muertes.

La tendencia de casos y muertes se mantiene en los niveles más altos desde el comienzo de la pandemia; disminuyó en las regiones de Europa y el Mediterráneo Oriental, mientras que en la Región de Asia Sudoriental se incrementó en un 6% debido a las cuatro variantes de SARS-CoV-2, según el reporte del 11 de mayo del 2021 (WHO, 2021). En Ecuador la situación fue de igual manera alarmante, debido a la desinformación que circulaba en esos días. La población especulaba sobre acontecimientos que muchas veces no eran ciertos, pero que llegaban a la sociedad a través de redes sociales.

Un punto que desconcertó a la gran mayoría de la población fue el número de fallecidos que en ese entonces las redes registraban diariamente. Los habitantes del país no tenían acceso a datos geoestadísticos que les permitieran conocer con exactitud la cantidad de fallecidos en una zona determinada y debido al confinamiento muchas personas no salieron de sus hogares. Esta situación generó ansiedad y temor en los ciudadanos.

Con el pasar del tiempo en América Latina se activó la circulación de nuevas variantes de COVID-19 que aumentaron los casos y las hospitalizaciones, mientras se continúan administrando las vacunas disponibles. La pandemia, además, ha acentuado las desigualdades sociales como consecuencia de la disminución del ingreso, ha descubierto un sistema de salud fallido y desatendido, principalmente en los países en vías de desarrollo, como es el caso de Ecuador.

El grupo etario que muestra un mayor número de contagios involucra a las personas de 20 a 49 años que corresponde al 60 % de la población. En las ciudades de mayor población, el crecimiento del nivel de contagios ha obedecido a un ajuste exponencial y se relaciona con eventos masivos, la densidad poblacional, patrones de dispersión y la lenta respuesta del Ministerio de Salud (Fernández et al., 2021), no así para el caso de los fallecidos (Cruz, 2021). Por lo antes mencionado, considerar introducir modelos geoestadísticos de fácil uso a la población ecuatoriana, es una herramienta que sin duda facilitará el acceso a la información del SARS-CoV-2 que esta y estará presente en nuestras vidas.

## **1.1 Descripción del problema**

Posterior a confirmarse la sospecha de la existencia de un nuevo coronavirus en China, a finales de enero del 2020, La Organización Mundial de la Salud (OMS) declaró una emergencia en la salud pública con repercusión internacional. Dicha emergencia se evidenciaba porque las personas sufrían un Síndrome Respiratorio Agudo Severo o SARS-CoV-2, denominado así por el Comité Internacional de Taxonomía de Virus (ICTV) en febrero del 2020. Este virus se diseminó a una velocidad vertiginosa por lo que en marzo fue declarada pandemia (WHO,2020), marcando un hito en el mundo al producirse una cuarentena a escala global.

Transcurrieron pocos días hasta que, en Ecuador, se registró la primera persona fallecida a causa del virus SARS-CoV-2 (MSP,2020). Consecuentemente se observó que las principales provincias de Ecuador se vieron afectadas en gran medida por un aumento considerable entre casos confirmados y decesos ocasionados por el COVID; siendo Guayaquil el epicentro de la pandemia. Al mismo tiempo, la incertidumbre sobre la exactitud de los datos registrados aumentaba y se incrementaba al no conocer el número de personas asintomáticas o no registradas con COVID; imposibilitando una eficaz implementación de medidas de control (Aguirre-Duarte,2020). Por lo tanto, la incerteza se transmitió a los modelos estadísticos y proyecciones de personas contagiadas con COVID realizadas por los investigadores a inicios de la pandemia.

Por otro lado, a medida que los datos comenzaron a incrementarse se logró visualizar de forma más precisa la dimensión del virus en cada una de las provincias y, por ende, las proyecciones sobre nuevos contagios eran más apegados a la realidad.

Bajo lo indicado anteriormente se ha considerado oportuno implementar modelos estadísticos que nos permitan realizar proyecciones espacio-temporales del virus SARS-CoV-2 a nivel nacional. Con el objetivo de alcanzar esto, se introducirá un modelo geoestadístico que nos permitirá mapear el aumento del número de personas contagiadas dentro de las áreas provinciales, adicional a esto se introducirá un modelo de series de tiempo que ayudará a visualizar la evolución de los casos de contagio o fallecidos a fin de determinar algún tipo de tendencia o patrón; Como alternativa a dicho enfoque la modelización mediante los modelos VAR evita las imposiciones derivadas de la estimación e identificación de un modelo de predicción que refleje lo más fielmente posible las regularidades empíricas e interacciones entre las variables objeto de análisis.



## **1.2 Justificación del problema**

Puesto que la sociedad continúa en alerta a causa de las nuevas variantes del COVID y dado que se cuenta con un volumen considerable de registros de casos confirmados posterior a la pandemia, este trabajo pretende aplicar series de tiempo sobre los registros a partir de la semana 46 del año 2021. Esta metodología permite seguir un proceso de identificación, diagnóstico y predicción de componentes presentes dentro de un modelo de serie temporal, mediante un modelo VAR en la que el tiempo es una variable clave durante el proceso (Ortiz Cardona,2021).

Por otro lado, se usa la metodología de Kriging sobre el registro de casos confirmados en distintos puntos del país. Al trabajar con coordenadas de las distintas provincias, se mapea el avance y expansión del COVID sobre el territorio ecuatoriano, mediante un método de interpolación geoestadística se estiman los valores desconocidos a través del uso de variogramas (Bravo,2020).

## **1.3 Objetivos**

### **1.3.1 Objetivo General**

Establecer modelos estadísticos que permitan la visualización de la evolución de los indicadores relacionados al virus SARS-CoV-2 a través de mapas de estimaciones.

### **1.3.2 Objetivos Específicos**

1. Pronosticar el impacto de un futuro rebrote del virus SARS-CoV-2.
2. Utilizar métodos geoestadísticos para identificar las zonas del país que registran aumento o disminución de contagios.
3. Implementar series de tiempo para visualizar pronósticos que se relacionen al aumento o disminución de contagios, generados a partir de la semana 46 del año 2021.

## **1.4 Marco Teórico**

### **1.4.1 Evolución de los contagios a causa de SARS-COV-2**

Al revisar las estadísticas concernientes al año 2021, se puede observar que la tasa de pacientes contagiados por el coronavirus a mediados de septiembre fue de 16.74 por cada 100.000 habitantes (INEC,2021) y esta cifra evolucionó al 31 diciembre, puesto que se registró que el 19,4% de los ecuatorianos tenían síntomas relacionados con el SARS-COV-2. Es decir, existían 3.473.887 personas contagiadas.

### **1.4.2 Proyecciones de casos confirmados mediante modelos matemáticos**

Al tener presente los cambios en las tasas de contagios, además de observar la evolución de estos dentro de un lapso determinado o dentro de una geografía, es necesario considerar el uso de una estructura matemática o un modelo que permita esbozar el comportamiento de estos registros constituidos como variables. Posterior a la construcción del modelo, se puede informar sobre un potencial incremento en el futuro, todo esto mediante un análisis de los pronósticos o las proyecciones que se obtengan (Prades y Marín,2020).

### **1.4.3 Proyecciones temporales mediante modelos de series de tiempo**

Al introducirse en el campo de las series de tiempo, la variable de estudio es analizada en distintos intervalos de tiempo, así pues, tanto el número de personas contagiadas o el número de personas fallecidas será estudiada de forma semanal. De esta forma, el enfoque del modelo de series de tiempo se concentrará sobre aquellas transiciones

epidemiológicas críticas en intervalos de tiempo semanales. A causa de esto, se espera observar tendencias o patrones indispensables en el pronóstico de valores futuros (Sotomayor et al., 2021).

#### 1.4.4 Modelo de exponencial doble o Método lineal Holt

Charles Holt en 1957 postuló un modelo aplicado a tendencias lineales que varían en conjunto dentro de una serie de tiempo y permiten producir pronósticos, se conoce a dicho modelo como suavización exponencial doble.

Este es empleado en series de tiempo que presentan tendencia lineal, pero carecen de estacionalidad (Izar, 2019). Dentro del modelo se consideran una constante de suavizamiento  $\alpha$  y una de tendencia  $\beta$ . Así pues, se constituye la ecuación 1.1 de los pronósticos para el modelo lineal Holt.

$$\hat{Z}_t = l_t - b_t h \quad (1.1)$$

Donde:

$\hat{Z}_t$ : Indica el valor pronosticado al periodo t.

$l_t$ : Indica la estimación del nivel de la serie al tiempo t.

$b_t$ : Indica la estimación de la pendiente (tendencia) de la serie al tiempo t.

$h$ : Periodos por pronosticar en el futuro.

Además, se puede desglosar la ecuación 1.2 de pronósticos y obtener la ecuación de la estimación del nivel de la serie al tiempo t.

$$l_t = \alpha Z_t + (1 - \alpha)(l_{t-1} + b_{t-1}) \quad (1.2)$$

Por último, se puede describir la ecuación 1.3 de la tendencia o pendiente al tiempo t.

$$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \quad (1.3)$$

### 1.4.5 Modelo de media móvil integrado regresivo automático (ARIMA)

Este modelo paramétrico, propuesto por Yule y Lutsch en la década de los 20, los retrasos presentes en los pronósticos de las series estacionarias se denominan términos autorregresivos, además se considera a los retrasos de los errores en los pronósticos como media móvil. Por otro lado, muchos de los modelos de series de tiempo a analizar no son estacionarias, debido a esto al ajustar una serie no estacionaria se elimina la variación no estacionaria. Por ende, se puede enfatizar que los modelos ARMA son el origen de los modelos ARIMA, luego de se aplique procesos de diferenciación a los modelos ARMA a fin de el modelo sea estacionario en la media (Diaz,2020).

Se define el modelo ARIMA  $(p, d, q)$ , cuyos parámetros se establecen como:

$p$ : Indica el número de términos autorregresivos

$d$ : indica el número de diferencias necesarias para la estacionariedad.

$q$ : Indica el número de errores de pronóstico.

Así pues, se constituye la siguiente ecuación 1.4:

$$\phi(B)X_t = \phi(B)\nabla^d X_t = \theta_0 + \theta(B)w_t b_{t-1} \quad (1.4)$$

Donde:

$\phi(B)$ : es el operador autorregresivo estacionario.

$\phi(B) = \phi(B)\nabla^d$ : es el operador autorregresivo generalizado, no estacionario con  $d$  diferencias.

### 1.4.6 Modelo vector autorregresivo (VAR)

Los modelos ARIMA estudian de forma univariante o de forma separada a una serie de tiempo, pero en casos donde existen conjuntos de series de tiempo, es necesario utilizar otro enfoque o modelo en el que tome en consideración la independencia entre cada una de las series del conjunto, donde dichas interrelaciones son expresadas mediante una matriz de intersección compuesta bajo un constructo de pesos (Rodriguez,2021).

Previo a la construcción del modelo es necesario establecer ciertas instancias:

- Clasificar las variables en: endógenas y exógenas
- Asignación de restricciones o pesos a fin de llegar a identificarlos.

Se puede indicar que, en la construcción del modelo, se puede considerar un vector de tipo columna de  $k$  variables,  $Y_t = (Y_{1t}, Y_{2t}, \dots, Y_{kt})^T$ , el cual hace referencia a valores pasados, la ecuación 1.5 denota un vector autorregresivo de  $k$  variables y orden  $p$ .

$$Y_t = m + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t \quad (1.5)$$

Donde:

$A_i$ : para  $i = 1, 2, \dots, p$ , son matrices de coeficientes, con orden  $K \times K$

$m$ : vector de constantes con orden  $K \times 1$

$\varepsilon_t$ : vector en representación al ruido blanco, con orden  $K \times 1$

### 1.4.7 Semivariogramas

En la implementación geoestadística a un conjunto de datos que son mapeados dentro de área o espacio, se puede considerar el uso de una técnica que permite analizar el comportamiento espacial de la variable de interés en el área definida y estudiar su influencia sobre las observaciones que están lo más cercanos, denotando un agrupamiento espacial de

conjuntos de observaciones similares. Todo esto es posible a causa del uso de semivariogramas, este estadístico trabaja en función de distancias entre cada uno de las observaciones circundantes y usa las diferencias de las observaciones elevadas al cuadrado. Este proceso de agrupamiento denotado bajo la ecuación 1.6, forma parte del proceso de análisis estructural experimental. (Bedoya et al.,2022)

$$\gamma_{(h)} = \frac{1}{2} N_{(h)} \sum [Z_{(x)} - Z_{(x+h)}]^2 \quad (1.6)$$

Donde:

$\gamma_{(h)}$ : Semivarianza total

$N(h)$ : denota el número total de casos separados bajo una distancia h

$Z_{(x)}$ : Número de casos ubicados en X

$Z_{(x+h)}$ : Número de casos a una distancia h de X

#### 1.4.8 Kriging Ordinario

Al implementar un modelo geoestadístico, se debe tener en consideración el efecto de la autocorrelación o las relaciones estadísticas entre los puntos en cuestión, este es el supuesto fundamental sobre el cual se basa la construcción e implementación del kriging, puesto que, al tener una correlación espacial, entonces se puede tratar de explicar la variación presente en la superficie de estudio. Posterior a esto, se puede pretender realizar las estimaciones de la variable en cuestión a fin de obtener un conjunto de estimaciones. (Masero,2020)

$$\hat{Z}_{(s_0)} = \sum_{i=1}^N \lambda_i Z_{(s_i)} \quad (1.7)$$

Donde:

$\hat{Z}_{(s_0)}$ : estimación realizada en la ubicación  $n_i$

$\lambda_i$ : peso desconocido para la estimación realizada en la ubicación  $n_i$   
 $S_0$ : Ubicación de la estimación

# CAPÍTULO 2

## 2. METODOLOGÍA

En este capítulo se presentan los procedimientos utilizados para la construcción de las proyecciones espacio-temporal del virus SARS-CoV-2.

Para la elaboración del siguiente estudio se consideró realizar un mapeo sobre cada una de las provincias del Ecuador, así pues, se consideró utilizar la infografía diaria proporcionada por el MSP, en la cual se observó las diferentes provincias en las que se produjeron mayor número de contagios.

### 2.1 Selección de variables de estudio

Previo al semivariograma se consideró delimitar las semanas de estudio, así pues, se definió que la semana 46 del año 2021 como fecha de inicio del estudio y la semana 16 del año 2022 como la fecha fin del estudio. A fin de iniciar con la realización del presente estudio, se procedió a seleccionar las variables indispensables, dado que nuestro objetivo es implementar modelos geoestadísticos en las provincias del Ecuador durante las últimas semanas en las que se han registrado casos de personas contagiadas con COVID, se describe la presente tabla donde constan las variables del estudio.

**Tabla 2. 1 Variables de estudio**

**Fuente: Elaboración propia.**

<b>Variable</b>	<b>Soporte</b>	<b>Tipo</b>
Confirmados	[1, 868.285]	Cuantitativa
Longitud	[-4.8628, 1.0602]	Cuantitativa
Latitud	[-70,977, -80.9]	Cuantitativa
Provincias	Desde Azuay Hasta Zamora Chinchipe	Categorica



Posterior a la selección de las variables indispensables para el estudio, se consideró agrupar este conjunto de datos dentro de una matriz de datos  $X$ , de la cual usaremos los valores de casos confirmados, dichos valores son estudiados a fin de determinar la distribución y en caso de no existir una tendencia normal de los datos visualizado a través de un histograma de frecuencias, entonces se aplicará una transformación lineal de tipo logarítmica que permita obtener dicha tendencia. Adicional a esto, se usarán las coordenadas de latitud y longitud para la construcción e implementación del semivariograma.

## 2.2 Cálculo de los semivariograma semanales

Para realizar un mapeo de la distribución espacial de los casos de personas contagiadas con COVID en las distintas provincias del Ecuador, es necesario implementar el cálculo del semivariograma  $\gamma(h)$  que es una herramienta eficaz para describir el comportamiento de procesos aleatorios espaciales. En un inicio se construye y se bosqueja un semivariograma, el cual recibe parámetros iniciales tales como nugget o pepita  $C_0$ , el psill o la meseta  $C_1$ , el rango  $\alpha$  y el ángulo de tolerancia  $\phi$ .

Mediante el uso de cada uno de los parámetros iniciales, se procede a la construcción de cada uno de los semivariogramas semanales, lo cual nos lleva a realizar un ajuste a la nube de puntos proyectada, donde se realiza un ajuste a los puntos mediante una curva que interprete su comportamiento.

## 2.3 Interpolación del Kriging en semivariogramas

Luego de implementar los semivariogramas, se procede a realizar la interpolación de kriging  $\hat{Z}$ , en dicha interpolación se utiliza cada uno de los semivariogramas ajustados a modelos esféricos, de cada una de esta estimación se realiza un mapa de estimación sobre el mapa de Ecuador

continental, donde se logra visualizar las observaciones predichas en cada una de las semanas. A fin de mejorar las estimaciones de casos confirmados con COVID se hace uso de un modelo de series de tiempo que estudie el conjunto de parámetros de los semivariogramas ajustados.

#### **2.4 Modelo VAR aplicado en parámetros de los semivariogramas.**

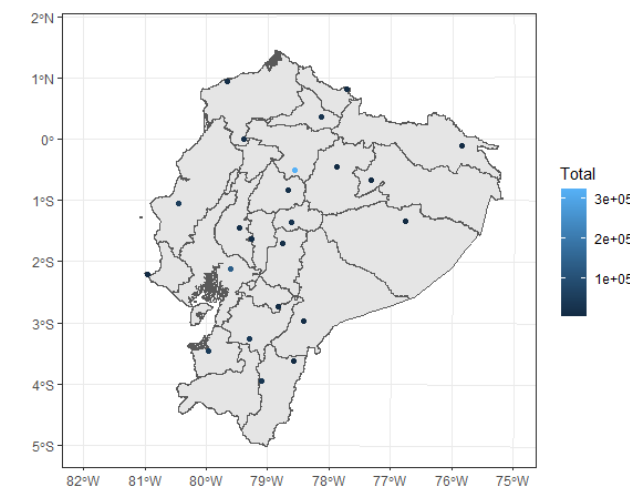
Al obtener un conjunto de parámetros de los semivariogramas semanales, se considera realizar un análisis de series de tiempo para estos parámetros y puesto que es un conjunto de series se puede realizar la implementación del modelo VAR, el cual requiere como supuesto inicial que cada una de las series de tiempo puedan cumplir con la estacionariedad, una vez satisfecho dicho supuesto se puede considerar el estudiar las series de tiempo en conjunto y realizar las proyecciones de los parámetros semanales, donde se puede obtener un nuevo subconjunto de parámetros los cuales son introducidos en el modelo de ajuste de cada uno de los semivariogramas o en el último semivariograma de la última semana y obtener así un mejor ajuste, posterior al cual se pueda realizar una nueva interpolación a fin de realizar nuevas estimaciones de los casos de contagio de COVID.

# CAPÍTULO 3

## 3. RESULTADOS Y ANÁLISIS

Este capítulo tiene como finalidad presentar los resultados obtenidos durante cada una de las etapas de este proyecto integrador, para alcanzar el cumplimiento de cada uno de los objetivos planteados en el Capítulo 1. En relación a estos, en un inicio se realizó un procedimiento Extracción, Transformación y Carga de Datos (ETL por sus siglas en inglés) que sirvió como guía en la extracción de un volumen considerable de datos de las últimas semanas relacionados con la cantidad de personas contagiadas a nivel provincial; no obstante, se estableció excluir a la provincia de las Galápagos y posterior a esto, mediante la transformación de los datos, se aplicó la función de logaritmos a la variable cantidad de contagios, para así usar dichos datos depurados en la carga de los modelos destinados para el proyecto.

En el Gráfico 3.1, se puede observar las coordenadas y el total de personas contagiadas en cada una de las provincias del estudio.

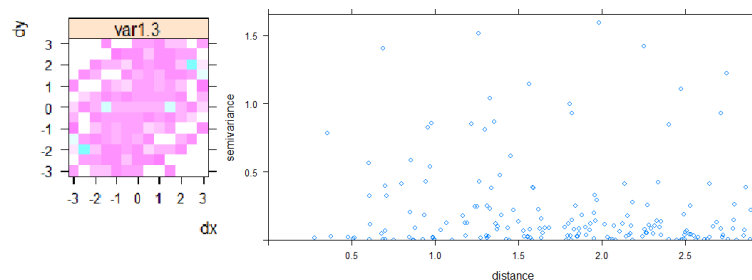


**Gráfico 3. 1 Mapeo y nube de puntos del semivariograma dentro Ecuador.**

**Fuente: Elaboración propia.**

### 3.1 Construcción y ajuste de semivariogramas

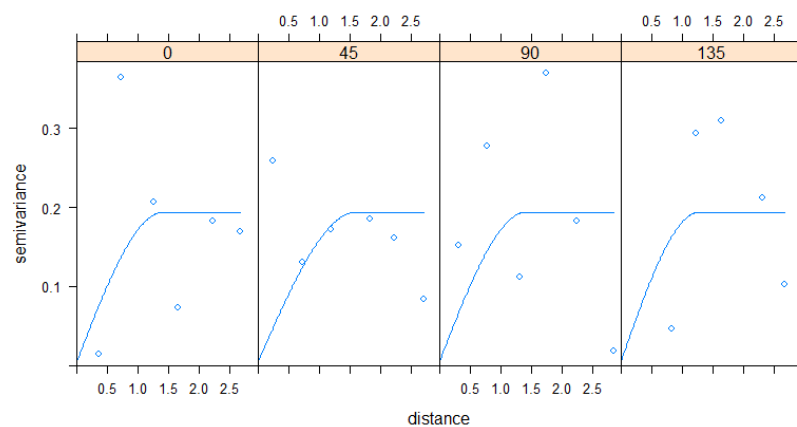
Para la construcción del semivariograma, primero se consideró realizar un mapeo por defecto el cual permitió visualizar tanto la estructura como la dependencia espacial de las observaciones dentro del marco georreferencial. El Gráfico 3.2, elaborado a partir de los datos de la semana 46 del año 2021, se evidencia que la nube de puntos del semivariograma presentó una tendencia en una dirección de  $45^\circ$ , la cual puede ser validada mediante la proyección de un gráfico de semivarianzas direccionales.



**Gráfico 3. 2** Nube de puntos del semivariograma y dirección.

**Fuente:** Elaboración propia.

En el Gráfico 3.3, se puede visualizar un ejemplo del ajuste realizado en cada uno de los semivariogramas semanales. Al usar un modelo esférico se evidenció que la curva se ajusta mejor al conjunto de observaciones. Adicional a esto, el modelo tuvo un mejor ajuste en un ángulo de  $45^\circ$ , en el que se presentó un nugget cercano a 0.001 y se dividió un psill cercano a 0.2. Finalmente, se consideraron varios ángulos de tolerancia entre  $70^\circ$  y  $80^\circ$ .



**Gráfico 3. 3** Ajuste de los semivariogramas direccionales

**Fuente:** Elaboración propia.

En la tabla 3.1, se pudo incorporar cada uno de los parámetros resultantes del modelo esférico implementado a cada uno de los semivariogramas semanales del periodo de estudio.

**Tabla 3. 1 Parámetros de cada modelo ajustado a los semivariogramas semanales.**

**Fuente: Elaboración propia.**

Semana	Nugget/Pepita	Psill/Meseta	Rangos	Tolerancias
1	0,000	0,2002	0,7931	0,8
2	0,000	0,1927	0,5183	0,8
3	0,000	0,2038	0,9172	0,7
4	0,000	0,1923	0,6148	0,8
5	0,000	0,1993	0,9297	0,7
6	0,040	0,1511	0,6345	0,7
7	0,069	0,1271	1,1545	0,7
8	0,000	0,1963	1,1198	0,7
9	0,055	0,1479	0,8949	0,8
10	0,041	0,1735	1,1028	0,8
11	0,037	0,1789	1,1913	0,7
12	0,000	0,2185	1,1786	0,7
13	0,000	0,2289	1,0747	0,7
14	0,000	0,2308	0,9564	0,7
15	0,000	0,2339	0,8834	0,8
16	0,000	0,2347	0,8749	0,8
17	0,000	0,2371	0,8730	0,8
18	0,000	0,2359	0,8415	0,7
19	0,000	0,2468	1,1220	0,7
20	0,000	0,2474	1,1218	0,7
21	0,000	0,2478	1,0297	0,8
22	0,000	0,2480	1,0292	0,8
23	0,000	0,2482	1,1211	0,7
24	0,000	0,2482	1,1211	0,7

### 3.2 Interpolación mediante Kriging.

En el Gráfico 3.4, se puede visualizar un ejemplo de interpolación kriging implementado en cada uno de los modelos esféricos ajustados a los semivariogramas semanales. Mediante el kriging General se identificó que la sierra norte de Ecuador presenta un mayor número de estimaciones relacionados a la cantidad de contagios, con mayor concentración en la provincia de Pichincha; de igual manera, se pudo observar que la zona costera sur presenta un elevado número de estimaciones, teniendo como mayor concentración en la provincia del Guayas.

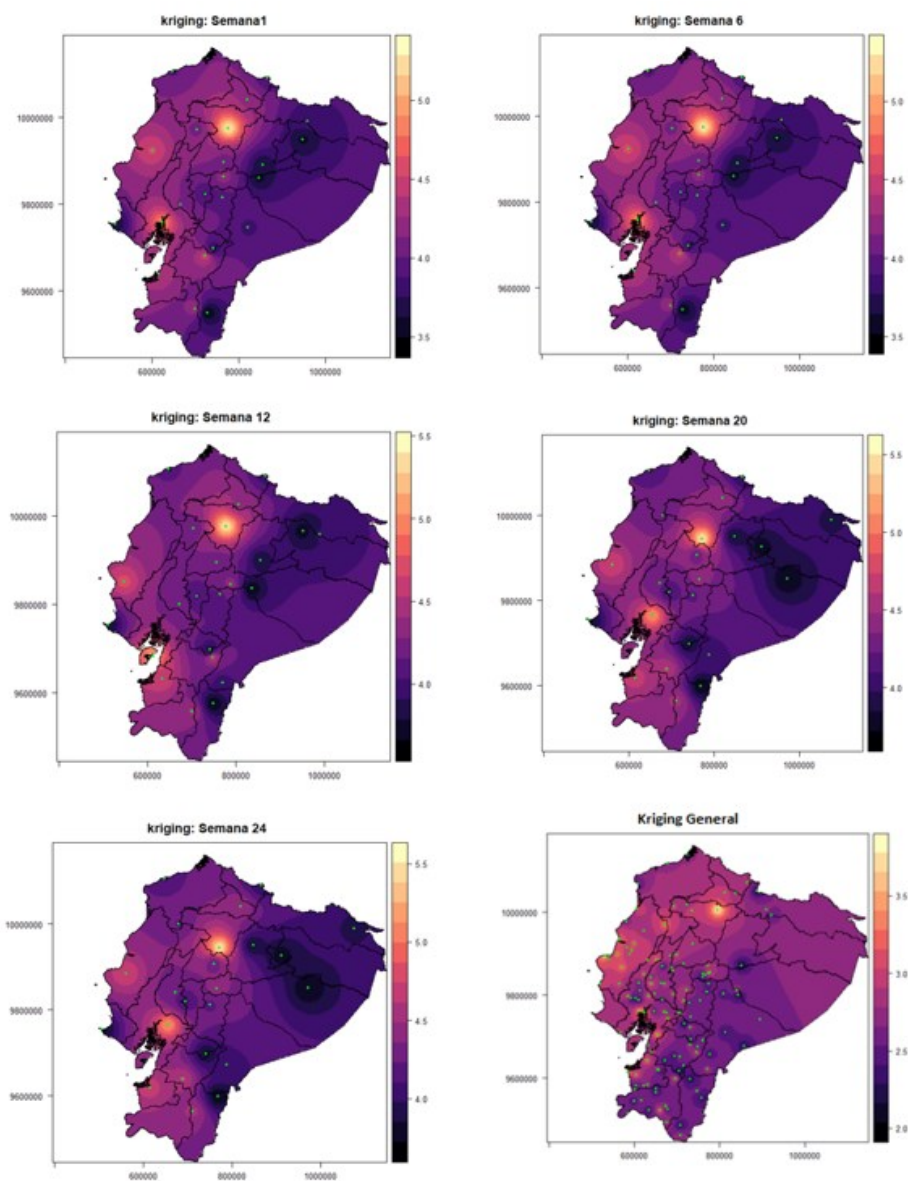


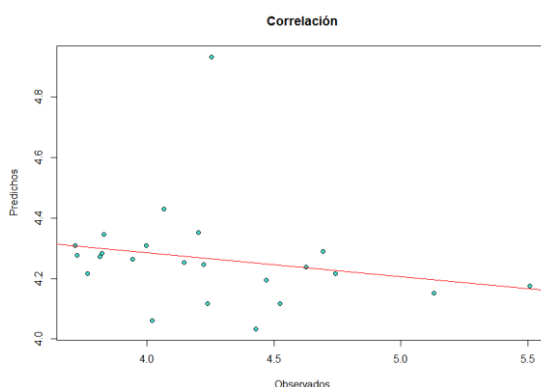
Gráfico 3. 4 Interpolación de Kriging para los contagios semanales

Fuente: Elaboración propia.

Por otro lado, se pudo visualizar que las provincias costeras junto a las provincias de la sierra sur presentan una cantidad de contagios ligeramente creciente, a diferencia de las estimaciones de contagio establecidas para toda la región amazónica, donde crecimiento de contagios es menor.

### 3.3 validación cruzada para definir mejor kriging.

En el Gráfico 3.5, se puede visualizar el resultado del ajuste realizado entre los valores observados y los valores predichos, luego de realizar la validación cruzada al kriging. Como resultado se obtuvo una correlación de -0.2111, la cual se interpretó como no significativa e indicó que no existe una relación fuerte entre las estimaciones realizados por el kriging y las observaciones iniciales. Por último, se obtuvo un ratio de desviación cuadrática media de 1.233.

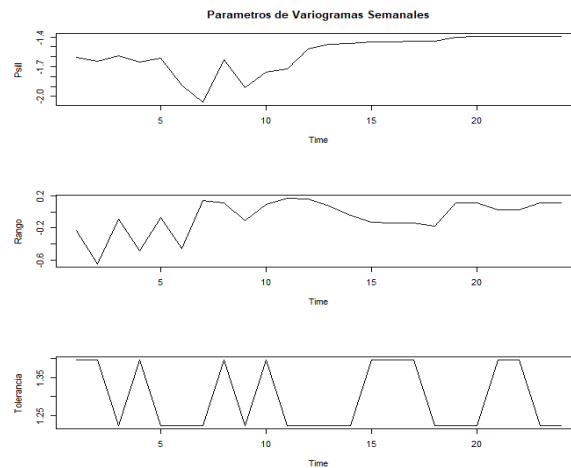


**Gráfico 3. 5 Correlación entre valores observados y las estimaciones.**

**Fuente: Elaboración propia.**

### 3.4 Modelización VAR sobre los parámetros del semivariogramas.

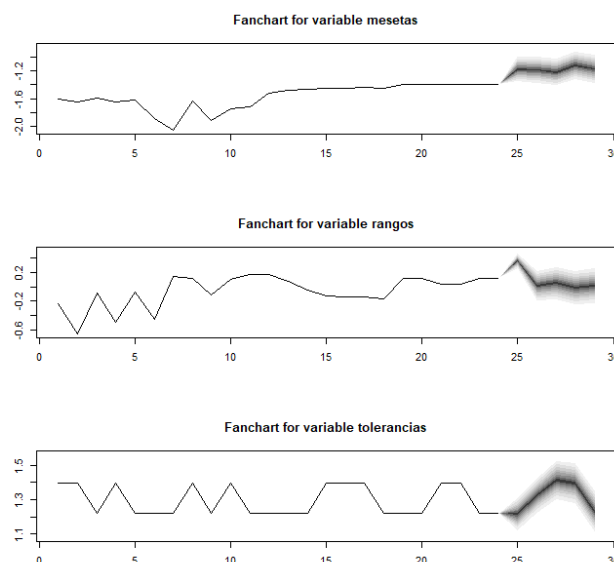
En el Gráfico 3.6, se puede visualizar la evolución de los parámetros resultantes de cada uno de los semivariogramas ajustados, luego de haber validado que cumplan con el supuesto de estacionariedad y haber seleccionado las series temporales de los parámetros indispensables para la construcción del modelo de series de tiempo múltiple vectores autorregresivos VAR.



**Gráfico 3. 6 Series temporales de parámetros de los semivariogramas.**

**Fuente: Elaboración propia.**

El Gráfico 3.7 indica los pronósticos realizados durante 5 semanas posteriores al último dato seleccionado previo a la realización proyecto. Se puede visualizar que en el parámetro meseta los pronósticos registraron una tendencia leve a incrementar para luego estabilizarse y descender. Por otro lado, en el parámetro rango se puede observar que, así como la serie tiene una tendencia rápida al crecimiento, de la misma forma comienza a decrecer rápidamente luego de un punto limite. Por último, se puede visualizar que en el parámetro tolerancia el crecimiento es cíclico de acuerdo con los registros históricos de dicha serie.



**Gráfico 3. 7 Pronósticos del modelo VAR de cada uno de los parámetros.**

**Fuente: Elaboración propia.**



### 3.5 Implementación de parámetros pronosticados en semivariogramas

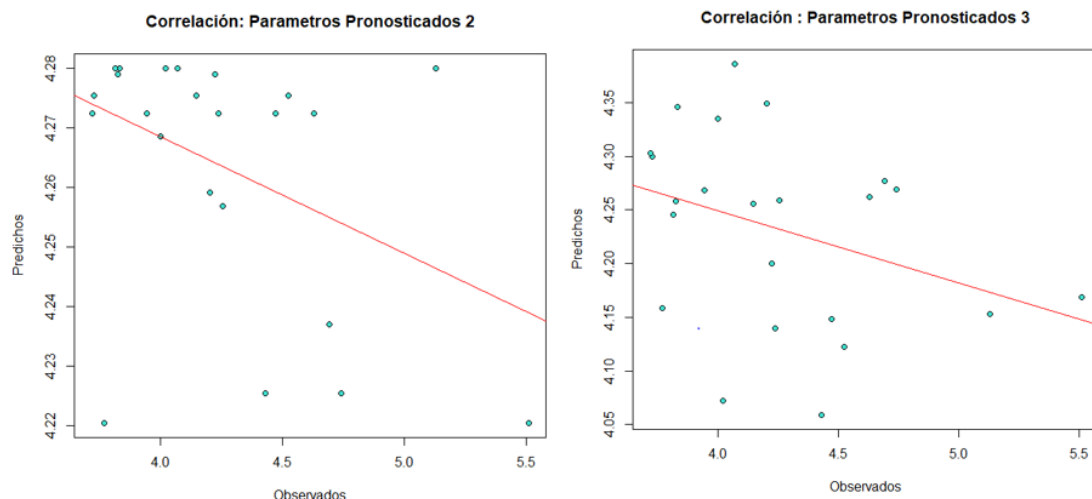
En la Tabla 3.2, se puede observar los valores relacionados a cada uno de los parámetros pronosticados para las 5 semanas siguientes, a partir de estos parámetros podemos elaborar nuevos variogramas iniciales y tratar de estimar un nuevo conjunto de kriging y estimar el número de contagios.

**Tabla 3. 2 Parámetros Pronosticados para 5 semanas.**

**Fuente: Elaboración propia.**

Parámetros	1ra Semana	2da Semana	3ra Semana	4ta Semana	5ta Semana
Meseta	1,17	1,18	1,21	1,12	1,17
Rango	0,36	0,02	0,05	0,01	0,02
Tolerancias	57	76	81	80	71

El Gráfico 3.8 muestra la correlación resultante de dos conjuntos de parámetros pronosticados, se puede observar que al implementar estos parámetros en la construcción de semivariogramas y posterior a esto realizar un ajuste para al final aplicar un kriging ordinario. Se denota que los valores de correlación aumentaron a  $-0.4214$  y  $-0.3461$  respectivamente. Por lo que se puede mencionar que debido a este nuevo conjunto de parámetros las estimaciones mejoraron.



**Gráfico 3. 8 Correlación resultante de nuevos ajustes Kriging.**

**Fuente: Elaboración propia.**

La Tabla 3.3 contiene los contagios pronosticados mediante los parámetros de la semana 2, se puede establecer mediante las coordenadas de latitud y longitud, las provincias y cantones a las cuales pertenecían estos pronósticos.

**Tabla 3. 3 Contagios pronosticados dado el nuevo ajuste kriging**

**Fuente: Elaboración propia.**

Provincias	Contagios	Longitud	Latitud
AZUAY	18726	-79,30	-3,27
BOLIVAR	19054	-79,26	-1,63
CAÑAR	19006	-78,83	-2,73
PASTAZA	18556	-77,72	0,81
CHIMBORAZO	18160	-78,76	-1,70
COTOPAXI	18062	-78,67	-0,84
EL ORO	17254	-79,96	-3,45
LOS RIOS	18854	-79,67	0,95
GUAYAS	19054	-79,60	-2,13
IMBABURA	18726	-78,13	3,63
ZAMORA CHINCHIPE	18854	-79,11	-3,94
LOS RIOS	18726	-79,46	-1,44
MANABI	16803	-80,45	-1,05
TUNGURAHUA	19054	-78,42	-2,97
NAPO	19054	-77,88	-0,45
ORELLANA	16611	-77,32	-0,67
SUCUMBIOS	18854	-76,78	-1,35
PICHINCHA	16611	-78,57	-0,50
SANTA ELENA	19054	-80,95	-2,21
SANTO DOMINGO	19006	-79,38	0,00
SUCUMBIOS	18726	-75,85	-0,10
TUNGURAHUA	16803	-78,62	-1,35
MORONA SANTIAGO	18726	-78,59	-3,62

# CAPÍTULO 4

## 4. CONCLUSIONES Y RECOMENDACIONES

Posterior a la elaboración del presente proyecto integrador el cual permitió definir el espacio geográfico o las provincias dentro de Ecuador continental y adicional es esto, analizar el comportamiento de los contagios confirmados en frecuencias semanales. Por lo que es necesario introducir a este capítulo de conclusiones y recomendaciones relacionadas estrechamente con los resultados que se obtuvo en la implementación de las metodologías descritas en el Capítulo 2.

### 4.1 Conclusiones

Al implementar la metodología de semivariogramas semanales relacionados a la cantidad de contagios provinciales, se determinó que un modelo esférico era el más adecuado al conjunto de datos.

La interpolación de Kriging inicial nos ayudó a determinar un conjunto de estimaciones para el Ecuador continental, donde se observó que en las provincias de Pichincha y Guayas existe la mayor cantidad de concentración de casos confirmados; adicional a esto, se observó que una correlación no significativa entre las estimaciones y las observaciones iniciales

El modelo de series de tiempo de vectores autorregresivos genero pronósticos de los parámetros del semivariograma más precisos lo cual influyo en que el kriging realice mejores estimaciones y por ende se observó un aumento en la correlación entre los valores observados y los valores estimados.

En los pronósticos de contagios a causa del COVID no se visualizó mucha variabilidad entre las provincias de cada una de las provincias en las que se obtuvo las estimaciones de los contagios.

## **4.2 Recomendaciones**

Para futuras implementaciones de modelos geoestadísticos en casos de contagio por COVID, se considera que en un inicio puedan agruparse un subconjunto de provincias de la costa y aplicar los semivariogramas y kriging, de la misma forma se puede realizar este agrupamiento a otro subconjunto de provincias de la sierra y el Oriente.

Para ello es necesario contar con un mayor número de datos espacialmente referenciados, sugiriendo el uso de la información a nivel cantonal y no solo a nivel provincial.

Consecuente a la recomendación previa, al tener una mejor estructuración de las zonas provinciales, permitirá realizar un ajuste del semivariograma más focalizado y la interpolación del kriging será más precisa, lo que permitirá obtener pronósticos más precisos.

# BIBLIOGRAFÍA

- Aguirre-Duarte, N. (2020). Can people with asymptomatic or pre-symptomatic COVID-19 infect others: a systematic review of primary data. <https://doi.org/10.1101/2020.04.08.20054023>.
- Bedoya, S. Z., Acero, D. M. W., & Mercado, M. (2022). Modelos geoespaciales para control de brotes de SARS-CoV-2 en Cartagena y Barranquilla, Colombia, 2020. *Revista Panamericana de Salud Pública*, 46. doi: <https://doi.org/10.26633/RPSP.2022.26>.
- Díaz Pinzón, J. E. (2020). Perspectiva del COVID-19 en Colombia para el año 2021. *Revista Repertorio De Medicina Y Cirugía*, 128–133. <https://doi.org/10.31260/RepertMedCir.01217372.1136>.
- Inca Ruiz, G., & Inca León, A. (2020). Evolución de la enfermedad por coronavirus (COVID-19) en Ecuador. *La Ciencia Al Servicio De La Salud*, 11(1), 5-15. doi:10.47244/cssn.Vol11.Iss1.441.
- Izar, Juan. (2019). Método de Holt. 10.13140/RG.2.2.18763.23844. [https://www.researchgate.net/publication/331844153\\_Metodo\\_de\\_Holt](https://www.researchgate.net/publication/331844153_Metodo_de_Holt).
- Maleki, M., Mahmoudi, M. R., Heydari, M. H., & Pho, K. H. (2020). Modeling and forecasting the spread and death rate of coronavirus (COVID-19) in the world using time series models. *Chaos, Solitons & Fractals*, 140, 110151. doi: 10.1016/j.chaos.2020.110151.
- Masero Bravo, L. (2020). Una aplicación de interpolación Kriging para el estudio de la evolución de la pandemia Covid-19 en España y en la Comunidad Valenciana. <http://hdl.handle.net/10234/193962>.
- Ministerio de salud Pública (2020). “Ministra de Salud confirma muerte de paciente por COVID-19” Descargado de <https://www.salud.gob.ec/ministra-de-salud-confirma-muerte-de-paciente-por-covid>.

- Ordoñez Mercado, A. (2021). Modelos híbridos SARIMA-ANN para pronósticos de la COVID-19 en el Perú. *Revista IECOS*, 22(1), 7-22.  
<https://doi.org/10.21754/iecos.v22i1.1332>.
- Ortiz Cardona, I. A. (2021). Propuesta de modelo ARIMA para la serie temporal de los casos de Covid-19 en Colombia aplicando la metodología Box and Jenkins.  
<http://hdl.handle.net/11371/3594>
- Prades Escobar, Eliecer, & Marín Sánchez, Dayana. (2020). Modelos estadísticos para las predicciones de la COVID-19 en Cuba. *Revista Cubana de Higiene y Epidemiología*, 57, e303. Epub 01 de septiembre de 2020. Recuperado en 20 de julio de 2022, de  
[http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1561-30032020000100005&lng=es&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1561-30032020000100005&lng=es&tlng=es).
- Rodríguez Arana Zumaya, A. (2021). La expansión de Covid-19 en México en 2020: un enfoque desde la econometría de series de tiempo.  
<http://ri.iberomx.mx/handle/iberomx/4810>.
- Sotomayor, Daniel Angel Córdova, & Carlos, Flor Benigna Santa Maria. (2021). Aplicación del método autorregresivo integrado de medias móviles para el análisis de series de casos de covid-19 en Perú. *Revista de la Facultad de Medicina Humana*, 21(1), 65-74. <https://dx.doi.org/10.25176/rfmh.v21i1.3307>.
- World Health Organization(2020).TIMELINE COVID -19 Descargado de  
<https://www.who.int/es/news/item/27-04-2020-who-timeline---covid-19>.