

**UNIVERSIDAD DE CIENCIAS Y
ARTES DE CHIAPAS
FACULTAD DE INGENIERÍA
PROGRAMA EDUCATIVO DE INGENIERÍA AMBIENTAL**

INFORME TÉCNICO

**MÉTODO DE REGRESIÓN ESPACIAL PARA
DETERMINAR LA GENERACIÓN DE RSU EN
BERRIOZÁBAL, CHIAPAS.**

**PARA OBTENER EL TÍTULO DE
INGENIERO AMBIENTAL**

**PRESENTA:
MÁRQUEZ PÉREZ DAVID FRANCISCO**

**DIRECTOR:
DR. JUAN ANTONIO ARAIZA AGUILAR**

TUXTLA GUTIÉRREZ, CHIAPAS

JUNIO 2023





UNIVERSIDAD DE CIENCIAS Y ARTES DE CHIAPAS
 SECRETARÍA GENERAL
 DIRECCIÓN DE SERVICIOS ESCOLARES
 DEPARTAMENTO DE CERTIFICACIÓN ESCOLAR
 AUTORIZACIÓN DE IMPRESIÓN

Lugar: Tuxtla Gutiérrez, Chiapas
 Fecha: 05/06/2023

C. David Francisco Márquez Pérez

Pasante del Programa Educativo de: Ingeniería ambiental

Realizado el análisis y revisión correspondiente a su trabajo recepcional denominado:
 Método de regresión espacial para determinar la generación de RSU en Berriozábal, Chiapas.

En la modalidad de: Informe técnico

Nos permitimos hacer de su conocimiento que esta Comisión Revisora considera que dicho documento reúne los requisitos y méritos necesarios para que proceda a la impresión correspondiente, y de esta manera se encuentre en condiciones de proceder con el trámite que le permita sustentar su Examen Profesional.

ATENTAMENTE

Revisores

Mtro. Pedro Vera Toledo

Dr. Hugo Alejandro Nájera Aguilar

Dr. Juan Antonio Araiza Aguilar

Firmas:

Ccp. Expediente

ÍNDICE GENERAL	
I. INTRODUCCIÓN.....	4
I.1. PLANTEAMIENTO DEL PROBLEMA.....	4
I. 2. JUSTIFICACIÓN.....	6
I. 3. OBJETIVOS	7
I. 3. 1. OBJETIVO GENERAL	7
I. 3. 2. OBJETIVOS ESPECÍFICOS	7
II. MARCO TEÓRICO	8
II.1 RESIDUOS SÓLIDOS URBANOS	8
II.1.1 FACTORES QUE CONTRIBUYEN CON LA GENERACIÓN DE RSU	9
II.1.2. MÉTODOS DE CUANTIFICACIÓN DE RSU	11
II.2 MODELOS DE REGRESIÓN.....	12
II.2.1 TIPOS DE MÉTODOS DE REGRESIÓN.....	13
II. 2. 1. 1 MODELO DE REGRESIÓN ESPACIAL.....	13
II. 2. 1. 2 TIPOS DE DATOS ESPACIALES	14
II. 2. 1. 3 CORRELACIÓN ESPACIAL.....	15
II. 2. 1. 4 TIPOS DE MODELOS DE REGRESIÓN ESPACIAL	15
II. 3 VALIDEZ DEL MODELO DE REGRESIÓN.....	17
III. METODOLOGÍA.....	18
III.1 ZONA DE ESTUDIO	18
III.2 GENERACIÓN DE LA BASE DE DATOS	18
III.3 ANÁLISIS EXPLORATORIO DE VARIABLES.....	19
III.4 MULTICOLINEALIDAD.....	20
III.5 DESARROLLO DEL MODELO.....	22
III.6 VALIDACIÓN DEL MODELO DE REGRESIÓN	22
III. 7 DIAGRAMA GENERAL DE LA METODOLOGÍA	23
IV. RESULTADOS Y DISCUSIÓN	24
IV.1 ANÁLISIS EXPLORATORIO DE VARIABLES.....	24
IV. 1. 1 CORRELACIÓN DE LAS VARIABLES.....	25
IV. 1. 2 VARIABLES NO SIGNIFICATIVAS.....	27
IV.2 MULTICOLINEALIDAD.....	28
IV. 2. 1. VARIABLES SIGNIFICATIVAS	29
IV.3 VALIDACIÓN DEL MODELO	30

IV.4 ESTIMACIÓN DE RSU DEL MODELO DE REGRESIÓN.....	31
IV. 5 MODELO DE REGRESIÓN.....	35
V. CONCLUSIÓN	37
V. 1 CONCLUSIÓN POR OBJETIVOS.....	37
V.2 CONCLUSIÓN GENERAL.....	38
VI. BIBLIOGRAFÍA	39

ÍNDICE DE TABLAS.

Tabla 1. Fuente y composición de los RSU.....	8
Tabla 2. Factores generales de la generación de RSU e información del cómo influye.....	10
Tabla 3. Métodos de regresión con su respectiva definición y fórmulas	13
Tabla 4. Variables socioeconómicas y demográficas.....	19
Tabla 5. Variables que aprobaron normalidad	26
Tabla 6. Prueba de Spearman	26
Tabla 7. Primera corrida de modelo	28
Tabla 8 Cuarta corrida del modelo	29
Tabla 9 Diagnostico de OSL	30

ÍNDICE DE FIGURAS

Figura 1 Zona de estudio de Berriozábal, Chiapas.....	18
Figura 2 Diagrama de proceso de la metodología	23
Figura 3 Comportamiento de la variable Y1 aplicando el índice de moran.....	24
Figura 4 Grafica de normalidad de tasa de empleo.....	25
Figura 5 Grafica de normalidad de Grad. Prom. Esc.	25
Figura 6 Índice de moran de residuos del modelo	31
Figura 7 Mapa de estimación de RSU en Berriozábal, Chiapas.....	32
Figura 8 Histograma de viviendas totales de Berriozábal.....	33
Figura 9 Histograma de comercios totales de Berriozábal, Chiapas	33
Figura 10 Histograma de densidad poblacional de Berriozábal, Chiapas	34
Figura 11 Mapa de residuos estandarizados del modelo	36

I. INTRODUCCIÓN

I.1. PLANTEAMIENTO DEL PROBLEMA

A principios del siglo XX comenzó a incrementar la urbanización en diferentes ciudades, múltiples comunidades se trasladaban a las zonas urbanas con el objetivo de obtener oportunidades de trabajo, esto trajo consigo un aumento de población en esos lugares provocando un crecimiento en la generación de Residuos Sólidos Urbanos (RSU). El desarrollo urbano causó un incremento de empresas que empezaron a generar productos para la población y como resultado empezó un consumo excesivo y una irresponsabilidad por parte de los consumidores [1].

La Ley General Para La Prevención y Gestión Integral de Residuos (LGPGIR) define a los RSU como

Los generados en las casas habitación, que resultan de la eliminación de los materiales que utilizan en sus actividades domésticas, de los productos que consumen y de sus envases, embalajes o empaques; los residuos que provienen de cualquier otra actividad dentro de establecimientos o en la vía pública que genere residuos con características domiciliarias, y los resultantes de la limpieza de las vías y lugares públicos [2].

Los RSU son una problemática que tomó relevancia en las últimas décadas. Aunque no se logre visualizar a simple vista es un hecho que nuestro planeta se está quedando sin espacio en donde depositar estos desperdicios. La composición de los RSU es variada, lo cual quiere decir que están compuestos de diversos materiales como son, materia orgánica, inorgánica, aparatos electrónicos, etc, por lo que se debe tener una adecuada clasificación a la hora de gestionarlos. Las consecuencias de los RSU no solamente se reflejan en lo ambiental, sino también se refleja en la sociedad, por ejemplo, los crecientes aumentos de residuos generan una dificultad para su recolección, provocando un rápido agotamiento de la vida útil de los rellenos sanitarios. Otros problemas son la falta de aprovechamiento de materiales y el aumento de sitios de disposición final de residuos que no cuentan con un tratamiento previo [3].

El problema de los RSU es considerado uno de los más grandes a nivel mundial, principalmente por los altos volúmenes de generación que afectan la salud de los habitantes como al medio ambiente, por lo que empiezan a surgir estrategias para mejorar la gestión de ellos. Se han realizado estudios a nivel mundial sobre la generación de RSU, en donde se tomaron como sitios de estudio a países de Latinoamérica y el Caribe. En esos estudios se ha estimado que en las últimas décadas la tasa de generación de residuos ha incrementado del 0.5 a 1 kg/habitante-día, por lo que se observa que en cada año aumenta la generación per cápita [4].

El crecimiento demográfico es una de las razones de las cuales se hablan más dentro del tema de generación de residuos, ya que por el aumento de la población empiezan a surgir nuevas causas de generación de RSU, las cuales son la falta de conciencia de la población, servicios de limpieza ineficientes y la falta de cultura sobre el reciclaje [1]. México es uno de los países que ocupa el onceavo lugar en población a nivel mundial. Gracias a datos del INEGI se puede notar una explosión demográfica entre los años 2015 y 2020, en el año 2015 la población de México fue de 119, 530, 753 habitantes y en el año 2020 se estimaron 126, 014, 024 habitantes, es necesario tener un método para determinar la generación de RSU con el fin de crear medidas para contrarrestarlos [5]. La LGPGIR propone que las entidades federativas y municipios deben de elaborar programas locales para la gestión y prevención de RSU, aunque primero se tiene que realizar una cuantificación de los productos [2]. INEGI cuenta con bases de datos por Área Geoestadística Básica (AGEBS) de distintos censos ya sean demográficos, económicos o sociales. Numerosos investigadores han utilizado este tipo de variables para cuantificar la generación de RSU, aunque la mayoría no cuenta con bases de datos por lo que tienen que recurrir a encuestas u otros métodos para obtener este tipo de información, cabe mencionar que las encuestas realizadas para obtener los datos no siempre serán fiables por lo que habrá presencia de errores a la hora de utilizar los datos [6].

I. 2. JUSTIFICACIÓN

El estado de Chiapas se coloca en el 10º lugar a nivel nacional de los que generan altos volúmenes de RSU ya que producen por día 5,188 toneladas, de los cuales se recolectan 2,886 t/día, cabe mencionar que esta recolección no es la verdadera, es decir Chiapas no cuenta con un servicio de recolección eficiente, esto se debe por que los camiones son inadecuados puesto que no cuentan con un mantenimiento correcto [3]. En cada año que transcurre el número de población va creciendo gradualmente por lo que se empieza a consumir más productos, esto trae consigo un aumento de RSU [3]. Otro de los posibles factores del aumento de basura en un municipio es por el cambio de consumo de la población o estilos de vida, por lo que se tiene que tomar en cuenta esos factores. Actualmente el municipio de Berriozábal, Chiapas, cuenta con servicio de recolección, la última vez que se realizó un estudio de cuantificación fue en el 2016 sin embargo se requieren estudios por lo menos una vez al año para actualizar los planes de manejo de RSU. Adicionalmente, hoy en día se está enfrentando el problema de la pandemia por COVID 19, lo que dificulta realizar la cuantificación de RSU de manera presencial, así que una manera de llevar a cabo este tipo de estudios es mediante la creación de modelos predictivos [1].

Un plan de gestión de residuos se considera fundamental, ya que proporciona un sistema viable para manejar los RSU de una manera ambiental y económica, sin embargo, se necesita información de la cantidad de residuos que se generan al día para planificar la operación y optimización de los sistemas de gestión [7]. Existen métodos para cuantificar la generación de RSU, uno de ellos es la NMX-AA-061-1985. El problema de aplicar la metodología de esta norma es que se necesita el apoyo de los H. ayuntamientos Municipales para que brinden las herramientas necesarias como los instrumentos, almacenamiento, y transporte por lo que puede influir económicamente. Otro de los problemas que más se observa en este tipo de métodos, es que no siempre la comunidad en donde se realizará el estudio estará dispuesta a cooperar, esto se debe a la falta de interés o de confianza, esto perjudica de gran manera el estudio por la falta de datos por lo que se propone un buen número de casas para evitar este problema. Otro de los métodos que se trabajan a nivel mundial, son los modelos de regresión lineal, en donde ocupan variables explicativas. Estos modelos se consideran una herramienta eficiente para determinar la generación de RSU, la cuestión es que solamente se pueden aplicar en ciudades que cuenten con los suficientes datos para aplicar los modelos. De esta manera se propone utilizar un método de regresión espacial utilizando tanto datos tabulados como espaciales, con el fin de cuantificar la generación de RSU del municipio de Berriozábal y así ayudar al municipio que actualice sus planes de gestión de residuos. El

tipo de investigación de este trabajo es cuantitativo, ya que se pretende utilizar herramientas de análisis matemático con el objetivo de describir o explicar la generación de RSU.

I. 3. OBJETIVOS

I. 3. 1. OBJETIVO GENERAL

Determinar las tasas de generación de RSU del municipio de Berriozábal, a través de un método de regresión espacial.

I. 3. 2. OBJETIVOS ESPECÍFICOS

- Construir una base de datos que incluya variables explicativas de tipo social, económico y demográfico.
- Analizar estadísticamente el comportamiento de las variables incluidas el análisis.
- Aplicar el modelo construido a nivel de áreas geoestadísticas espaciales básicas, en el municipio de Berriozábal, Chiapas.
- Validar el modelo de regresión para determinar tasas de generación, a partir de diversas medidas estadísticas.

II. MARCO TEÓRICO

II.1 RESIDUOS SÓLIDOS URBANOS

El aumento de RSU provoca problemas en diferentes ámbitos como el ambiental, social y económico por lo que se considera un problema multidimensional, estos ámbitos son afectados ya que usualmente los municipios que no están totalmente urbanizados carecen de herramientas para tratar con los RSU y que además la población no está muy consciente de esta problemática. Con eso en mente se necesita rápidamente adoptar medidas de gestión de RSU para mitigar estos problemas [8].

Para abordar una correcta gestión de RSU es necesario cuantificar la generación per cápita, así como también hay que conocer su composición. Existen antecedentes en donde aplican la NMX-AA-015-1985 para determinar la generación de residuos con el fin de actualizar los planes de gestión integral de los RSU que se está manejando en ese municipio, también mencionan que es necesario que se realicen estudios por lo menos una vez al año [9]. El inicio de la gestión de RSU fue alrededor del año 1964 bajo un enfoque de regulación sanitaria, poco después se empezó un manejo básico de los residuos bajo la legislación ambiental, aunque este no era muy apoyado. En el año 2003 se empezó a emplear enfoques sustentables bajo la dirección de la LGPGIR [10].

La composición de RSU suele ser variada debido a que se generan por distintas fuentes como por ejemplo, domicilios, oficinas, empresas o industrias. Esta información se representa en la tabla 1.

Tabla 1. Fuente y composición de los RSU

Fuente	Lugar o forma de generación	Tipo
Doméstica	Vivienda, bloques de vivienda	Residuos de comida, cartón, plásticos, textiles, cuero, residuos de jardín, aluminio, hojalata y otros metales, electrodomésticos pequeños o de hogar, pilas, baterías, aceites y residuos domiciliarios peligrosos.
Comercial	Restaurantes, bares, tiendas, negocios en general, talleres, etc.	Papel, cartón, vidrio, comida, metales, residuos peligrosos
Institucional	Escuelas, municipios, dependencias locales, hospitales, centros de salud, etc.	Símil comercial

Servicios municipales	Calles, jardinería, limpieza urbana	residuos de poda y arbolado, cartón y papel, plástico.
-----------------------	-------------------------------------	--

II.1.1 FACTORES QUE CONTRIBUYEN CON LA GENERACIÓN DE RSU

Existen diferentes factores que influyen en la generación de RSU, una de las principales es la densidad demográfica, pero gracias a múltiples estudios que se realizaron para predecir la generación de residuos, se lograron encontrar más factores. Para modelar la generación de RSU se empleaban variables socioeconómicas y demográficas, las cuales dependen de la ciudad en donde se realizará el estudio, ya que dicho sitio de estudio puede o no contar con los datos necesarios [11].

Ya se ha mencionado algunos factores socioeconómicos, pero no los más importantes mayormente en los estudios de determinación se toman en cuenta los ingresos de la familia, el PIB, gastos, impuestos, empleo, desempleo, consumo de energía, etc., estos factores se toman más en cuenta para la predicción a nivel hogar. En México es muy común que comunidades rurales se trasladen a zonas urbanas con el objetivo de mejorar su calidad de vida esto se debe por el constante aumento de la urbanización, México no es el único lugar en donde ocurre esto, en Tailandia por ejemplo, se presencié un aumento rápido de la urbanización, lo cual produjo un incremento de municipios de 1131 a 2000, esto concluyó que la urbanización es un factor importante que contribuye a la generación de RSU ya que modifican el estilo de vida y de consumo [12].

Otros factores que se han considerado en otros estudios han sido los hábitos alimenticios, el nivel de vida, el grado de actividades comerciales y las estaciones del año. Hay que añadir que estos factores son tomados en cuenta en los estudios en donde se quiere averiguar la composición de los RSU [11]. En algunos estudios realizados los factores que más se toman en cuenta son, el tamaño del hogar, que se refiere al número de habitantes por vivienda y los ingresos de los hogares, también existen algunos factores que muchos investigadores han mencionado que puede influir en la generación de RSU, los cuales son la variación estacional, el consumo de gas, la cantidad de agua, electricidad, temperatura y precipitaciones [13]. Los factores generales de generación más comunes que se pueden encontrar en distintos estudios se representan en la tabla 2.

Tabla 2. Factores generales de la generación de RSU e información del cómo influye

Factores	Explicativa	Listado de variables
Económicos	<p>Se identificó que variables como los ingresos de la familia puede contribuir a la generación de RSU, ya que cuentan con un fondo monetario para comprar ciertos productos que pueden o no ser de primera necesidad, lo cual influiría en la composición de los residuos. Otro detalle que destacar es que la alta demanda de consumo también influye a que las empresas empiezan a generar más productos de lo que se debe y esto incita a cambiar los hábitos de consumo de la comunidad.</p>	<p>Gasto de consumo anual per cápita de los hogares urbanos [11], [13] y [14] Tasa de desempleo [12], [15]. Tasa de empleo [12], [15]. Ingresos [11], [12], [13], [16] PIB [12], [15]</p>
Demográficos	<p>La población aumenta gradualmente en cada año, esto ocasiona una demanda por parte de los consumidores ya que necesitan cambiar sus hábitos de consumo, en otras palabras, una familia de dos personas no genera la misma cantidad de residuos que una familia de tres personas, por lo que gradualmente empiezan a generar más residuos. Por lo que es fundamental tener en cuenta los habitantes por vivienda, ya que</p>	<p>Densidad de población [12] [15] Tamaño del hogar [11], [12] [14] [15]. Población [7] [15] Edad de las personas [15] [10] [9] Tasa de mortalidad [15] Educación [11] [12] [13]</p>

	cada vivienda genera una cantidad diferente de residuos.	
Urbanización	La urbanización es otro factor fundamental para tener en cuenta, ya que la generación de más empresas influye en muchos factores los cuales son los económicos y sociales. En lo económico se encarga de que las personas inviertan más en productos de segunda mano que al poco tiempo dejaran de utilizar y lo reemplazaran por otros más nuevos, en lo social generara más empleos y contribuirá a un aumento monetario para que las personas puedan consumir más.	Viviendas [12], [14], [15] Grado de urbanización [12]

II.1.2. MÉTODOS DE CUANTIFICACIÓN DE RSU

La cuantificación de RSU es una metodología de varias etapas en donde se determina la limitación del sitio del estudio y herramientas a utilizar. Se tiene antecedentes de un estudio realizado en Berriozábal en el año 2016 empleando la NMX-AA-061-1985, la cual establece un método para determinar la generación de Residuos Sólidos Municipales (RSM) a partir de un muestreo estadístico aleatorio [17]. Otras normas mexicanas que se aplican son NMX-AA-015-1985, que define el método de cuarteo para RSM para la obtención de especímenes para los análisis de laboratorio. Por último, se encuentra la NMX-AA-019-1985, que establece un método para determinar el peso volumétrico de los RSM donde se efectuó la operación del cuarteo. De acuerdo con un artículo de Colombia se propuso un plan de muestreo que se pueda adaptar a ciudades de tamaño mediano, este se aplicó en la ciudad de Cali, Colombia, consiste en un muestreo bietápico en el cual se selecciona un conglomerado de viviendas de un determinado estrato socioeconómico y que además sean exclusivamente habitacional, se aplican distintas fórmulas para estimar la generación de residuos por persona por día, por estrato económico y lados de manzanas (18). Existen

otros tipos de residuos que se pueden cuantificar los cuales son micro plásticos (MP), para ello se tiene la siguiente metodología: selección de zona de muestreo en donde se determina el tiempo a muestrear y los criterios que se tienen que llevar a cabo, luego se realiza el muestreo de campo para RSU creando transectos de 100 m, dependiendo de la zona de muestreo se determina el número de transectos y se recolectan, para el muestreo de Micro Plásticos se lleva a cabo la metodología de Alvarez-Zeferino [2] citado por Salas [19, p.7] en donde se emplea un muestreador cilíndrico de policloruro de vinilo, posteriormente se clasifican los RSU y MP. En Ecuador se llevó a cabo un estudio de cuantificación de RSU en la universidad Politécnica Salesiana que llevaba como línea de base la planificación, cuantificación y caracterización de los RSU conforme establecía la norma ASTM, D5231-92, 2008 “Método de ensayo estándar para la determinación de la composición de residuos sólidos municipales sin procesar”, para ello primero se determina el número de individuos a muestrear, el tamaño de la muestra y el número de muestras a caracterizar empleando distintas fórmulas estadísticas, posteriormente se categorizan los residuos y se determina la generación per cápita para más información consultar la norma ASTM, D531-92, 2008 [20]

II.2 MODELOS DE REGRESIÓN

Los métodos de regresión se aplican para la estimación, explicación o predicción de la totalidad de RSU a partir de un conjunto de datos históricos. De acuerdo con Beigl et al, [7], los modelos de generación de residuos toman en cuenta los siguientes criterios: escala regional, tipo de flujos de residuos modelados, tipos de variables independientes y el método de modelado.

- Escala regional: Se refiere al tamaño del estudio, es decir se realizará a nivel hogar, distrito o país, dependiendo del tamaño del estudio influirá en el número de variables independientes que se tomaran en cuenta.
- Tipo de flujos de residuos: Da entender el origen de los residuos (doméstico o urbanos)
- Variables independientes: Se refieren a las posibles causas que puedan predecir la generación de RSU.
- Método de modelado: Que son el número total de variables independientes y los métodos para la validación del modelo.

Para aplicar un modelo de regresión espacial es fundamental conocer si la variable independiente cuenta con correlación espacial, lo cual se puede identificar a través del índice de moran. En los modelos de

regresión se debe tener cuidado a la hora de ocupar las variables demográficas y socioeconómicas, ya que muchos de esos datos no van a servir para el estudio, por lo que es de vital importancia realizar un análisis de multicolinealidad, ya que esto permitirá observar si tres o más variables tienen una relación lineal con otra, y si ese es el caso, reduciría el poder predictivo del modelo de regresión [15].

II.2.1 TIPOS DE MÉTODOS DE REGRESIÓN

Existen múltiples métodos de regresión por aplicar en este tipo de investigaciones, los cuales se representan en la tabla 3.

Tabla 3. Métodos de regresión con su respectiva definición y fórmulas

Método	Definición	Fórmula	Antecedentes de estudios
Regresión simple	Su objetivo es estimar la relación existente entre dos variables (X y Y), para emplearse con fines predictivos.	$Y = \beta_0 + \beta_1 X$ β_0 y β_1 : son parámetros de regresión, es decir son desconocidos Y y X: son la variable dependiente e independiente	[13]
Regresión múltiple	Estudia las relaciones que se establecen entre la variable que se va a explicar y las que se usan para explicarlas.	$Y = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \epsilon_i$ Donde: $i = 1, 2, \dots, n$ y los coeficientes β son desconocidos.	[11]
Regresión espacial	Se utilizan principalmente para representar información geográfica	Se puede utilizar software para este tipo de método. $Y = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \epsilon_i$	[15]

II. 2. 1. 1 MODELO DE REGRESIÓN ESPACIAL

El modelado de regresión espacial es una representación geográfica de la zona de estudio en donde se puede visualizar el fenómeno estudiado en base a acciones que se correlacionen espacialmente. Existen

dos tipos de modelados espaciales, uno de ellos es el modelado basado en entidades que es una vista geográfica que reúne dentro de un objeto espacial puntos del espacio que comparten propiedades similares. En este modelo los tipos de fenómenos espaciales que se analizan se identifican por dimensionalidad [21].

- Punto: Son objetos cero-dimensionales que ubica entidades cuya superficie es pequeña en relación con la del espacio (ejemplos. Edificios, personas, epicentros de terremotos).
- Línea: Son objetos unidimensionales que se usan para representar entidades del espacio con forma de redes (ejemplos. Ríos, caminos, comunicaciones).
- Volumen: Son objetos que tienen longitud anchura y profundidad, y por lo tanto son tridimensionales.

El segundo tipo de modelo se denomina basado en campo, en donde cada punto del espacio está asociado con distintos valores de atributos, definidos como una función sobre dos coordenadas X e Y. Las mediciones sobre los fenómenos se reúnen como valores de atributos variando con la ubicación en el plano [21].

II. 2. 1. 2 TIPOS DE DATOS ESPACIALES

A la hora de describir los datos espaciales es importante distinguir el carácter discreto o continuo del espacio en el que se miden las variables y el carácter discreto o continuo de las variables. El modelo basado en campo los valores de las variables se deben de interpretar de forma continua ya que la continuidad del campo no se puede preservar en variables de valores discretos. Si el modelo es de objetos los valores de las variables se pueden valorar de forma continua o discreta (valores ordinales o nominales). Se puede distinguir cuatro tipos de datos espaciales [21]:

- Datos de patrones puntuales: Un conjunto de datos que consiste en una serie de ubicaciones de puntos en alguna región del estudio, en la que se han producido eventos de interés, como casos de enfermedad o incidencia de un tipo de crimen.
- Datos de campo: También se denomina datos geoestadísticos, se relacionan con variables que son conceptualmente continuas y cuyas observaciones se han muestreado en un conjunto predefinido y fijo de ubicaciones en punto.

- Datos de área: Los valores de datos son observaciones asociadas con un número fijo de unidades de área que pueden formar una red regular, como con imágenes de detección remota, o un conjunto de áreas o zonas irregulares.
- Datos de interacción espacial: también denominados flujo de origen-destino o datos de enlace, consisten en mediciones tal que cada una de las cuales está asociada con un par de ubicaciones de puntos o un par de áreas.

II. 2. 1. 3 CORRELACIÓN ESPACIAL

La geografía fue evolucionando alrededor de los años 70 en donde se comenzó a definir el término de autocorrelación espacial con el objetivo de analizar espacialmente diferentes fenómenos, su término fue acuñado por Andrew Cliff y Keith Ord en 1967 que se determinaba como la medición del grado de asociación que una variable desarrolla a través de un espacio definido como marco geográfico [22]. La esencia de la autocorrelación espacial es analizar la variabilidad de un fenómeno a través del espacio geográfico para determinar patrones espaciales y descubrir su comportamiento, por lo que puede ser un medio para comprender como se distribuye el fenómeno en el espacio analizado y en qué grado los elementos locales pueden verse afectados por sus vecinos [22]. Existen distintas definiciones de la autocorrelación espacial por diferentes autores Upton y Fingleton lo definen como una propiedad que los datos poseen cuando muestran un patrón de comportamiento, estos mismos autores exponen que la autocorrelación espacial existe siempre que haya una variación espacial sistemática en los valores a lo largo de un mapa, o patrones en los valores registrados en las localizaciones [19]. Goodchild menciona que en su sentido más general la autocorrelación espacial se refiere al grado en que los objetos o actividades en algún lugar de la superficie son similares a otros objetos o actividades ubicados cerca y refleja la primera ley de geografía de Tobler “todo está relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las cosas distantes” [23].

II. 2. 1. 4 TIPOS DE MODELOS DE REGRESIÓN ESPACIAL

El modelo de regresión espacial tiene como objetivo establecer relaciones entre las observaciones de una variable y las observaciones de otras variables, registradas para cada unidad del área. Existen distintos tipos de regresiones espaciales los cuales son los siguientes:

- Modelo de retardo espacial: Estos modelos presentan la correlación espacial en la variable dependiente (y). Son extensiones de modelos de regresión tipo (a) donde y_i es una observación de la variable dependiente o de interés, x_{iq} es una observación en una variable explicativa, β es el

coeficiente de regresión que mide la influencia por si sola de la q-variable explicativa en la variable dependiente. Permiten a las observaciones de la variable dependiente Y en el área i depender de las observaciones en áreas vecinas.

$$y_i = \sum_{q=1}^q x_{iq} \beta_q + \epsilon_i \quad i = 1, \dots, n \quad (1)$$

- Modelos de error espacial: Estos modelos explican la dependencia espacial en el término de error. La dependencia del error espacial puede surgir de variables latentes no observables que están correlacionadas espacialmente. También pueden surgir de los límites del área que no reflejan con precisión la vecindad que dan lugar a las variables recopiladas para el análisis. La especificación más común es un proceso autorregresivo espacial de primer orden, dado por:

$$\epsilon_i = \lambda \sum_{j=1}^n W_{ij} \epsilon_j + \mu_i \quad i = 1, \dots, n \quad (2)$$

Donde λ es el parámetro autorregresivo, y μ_i es un término aleatorio de error, que asumimos que es independiente e idénticamente distribuido. Si $|\lambda| \leq 1$, entonces nos queda:

$$\epsilon = (I - \lambda W)^{-1} u \quad (3)$$

Insertando esta ecuación en el modelo de regresión estándar obtenemos:

$$Y = X\beta + (I - \lambda W)^{-1} Uu \quad (4)$$

Con $E[uu^t] = \sigma^2 I$. En consecuencia, la matriz de la varianza-covarianza del vector de errores aleatorio es:

$$E[\epsilon\epsilon^t] = \sigma^2 (I - \lambda W)^{-1} (I - \lambda W^t)^{-1} \quad (5)$$

El modelo (4), denominado modelo de error espacial (SEM) puede verse como una combinación de un modelo de regresión estándar con un modelo autorregresivo espacial en el término de error ϵ .

II. 3 VALIDEZ DEL MODELO DE REGRESIÓN

Para conocer si las variables explicativas tienen una relación con la variable dependiente se tiene que aplicar la prueba de Spearman, que tiene como objetivo medir la fuerza o grado de asociación entre dos variables aleatorias cuantitativas que poseen una distribución normal bivariada conjunta, en pocas palabras, conocer la relación lineal entre las dos variables (dependiente e independiente), cabe recalcar que el coeficiente fluctúa entre $-1 \leq \rho \leq 1$, indicando que los valores mayores a 1 no cuentan con una relación lineal [24]. Para aplicar el coeficiente de Pearson la distribución conjunta de las variables X y Y deben ser normal, si la variable no presenta normalidad significa que la variable explicativa no tiene correlación con la variable dependiente. En caso de que las variables explicativas no presenten una distribución normal se puede recurrir a las pruebas no paramétricas que en este caso sería a través del coeficiente de Spearman, que tiene como misión medir la tendencia de X y Y a relacionarse de modo monótona creciente o decreciente, al medir el grado de asociación de forma monótona entre las variables X, Y no se encuentra restringido a descubrir sólo una asociación lineal entre las variables [24]. Para validar el modelo de regresión espacial se emplean distintos tipos de pruebas como la presencia de multicolinealidad, que interpreta la correlación entre dos o más variables explicativas con el objetivo de eliminar aquellas que tengan valores altos de 10, este hecho también incluye que algunas variables importantes no se han incluido en el modelo de regresión [25], la solución es eliminar aquellas variables explicativas que dependen unas de otras.

En los modelos de regresión se tiene que cumplir el supuesto de homocedasticidad que establece que la variabilidad de los residuos es independiente a las variables explicativas, en general la variabilidad de los residuos está fuertemente correlacionada con la variable dependiente. [26], la prueba que más se aplica es la de Koenker que indica que el valor P tiende a ser mayor a 0.05 para aceptar que presenta homocedasticidad, en caso de tener valores menores a 0.05 el modelo es heterocedástico y que la varianza de error es diferente en cada valor de x [27]. La prueba de Jarque Bera es otro tipo de prueba que se realiza para la validez del modelo de regresión que tiene como objetivo validar el supuesto de normalidad de los residuos del modelo, la hipótesis está establecida de la siguiente manera “ H_0 : La muestra sigue una distribución normal; H_1 : La muestra no sigue una distribución normal”, en otras palabras si el valor de $P \geq 0.05$ cumple con el supuesto H_0 , si los valores fueran menores se rechaza la hipótesis lo cual quiere

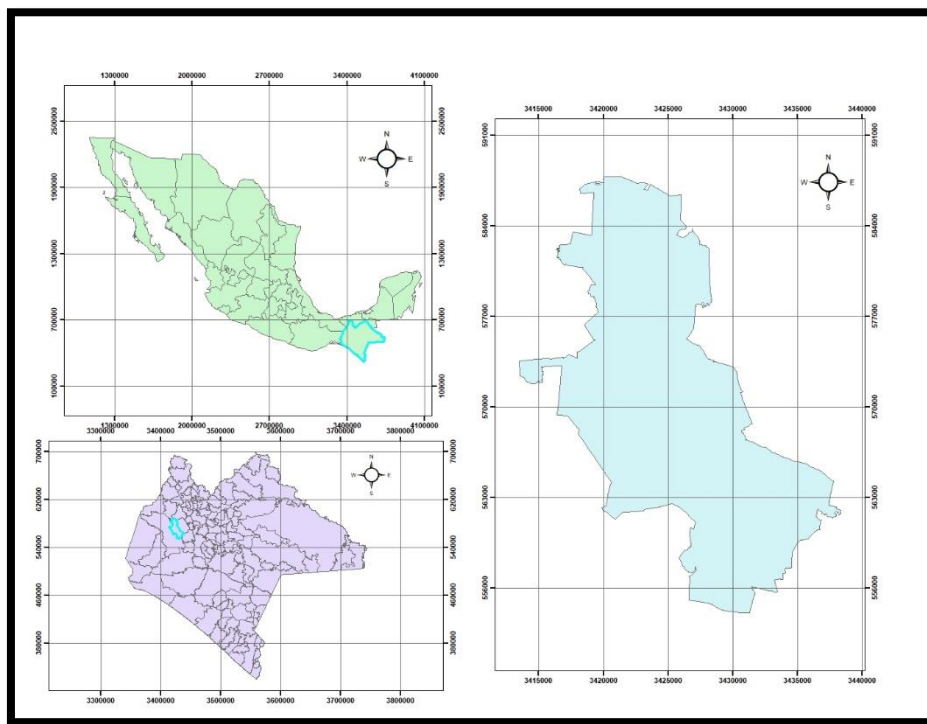
decir que falta una variable fundamental y que el modelo se encuentra sesgado [28]. La R^2 ajustada es un grado de confiabilidad o bondad de ajuste de los modelos ajustados de un conjunto de datos, mayormente los modelos de regresión tienen un valor de R^2 .90 por lo tanto explica el 90% del problema, la R^2 nos ayuda a comparar los modelos que se crean para conocer cuál de ellos es el más fiable.

III. METODOLOGÍA

III.1 ZONA DE ESTUDIO

El municipio de Berriozábal, Chiapas se ubica en el suroeste de la república mexicana, las coordenadas de la cabeza de su municipio son $16^{\circ}48'00''$ de latitud norte y $93^{\circ}16'22''$ de longitud oeste, se ubica a 900 m sobre el nivel del mar. El municipio de Berriozábal forma parte de la región metropolitana que incluye a Tuxtla Gutiérrez, su cabecera municipal se divide en un total de 22 barrios (figura 1) [29].

Figura 1 Zona de estudio de Berriozábal, Chiapas



Fuente: Propia

III.2 GENERACIÓN DE LA BASE DE DATOS

Para cumplir con el primer objetivo se consultó bibliografías relacionadas al tipo de trabajo que se realizó, algunos autores han identificado que los factores sociodemográficos y socioeconómicos son comunes en los modelos de generación de desechos, principalmente porque los datos son accesibles. Estudios

realizados al tema de investigación han concluido que este tipo de modelos es adecuado en zonas desarrolladas ya que cuentan con bases de datos estadísticos existentes que ayudan al objetivo del trabajo [16].

Este trabajo es a nivel municipal por lo que se obtuvieron los datos demográficos y socioeconómicos en dependencias como INEGI y DENUE del año 2020. Se tomaron en total 10 variables tabulares y 4 variables espaciales las cuales se representa en la tabla 4.

Tabla 4. Variables socioeconómicas y demográficas

Tabular	Espacial
Y: Generación de residuos	X7: Comercios
X1: Grado promedio de escolaridad	X10: Temperatura
X2: Viviendas totales	X11: Precipitación
X3: Promedio de ocupantes por vivienda	X14: Densidad de la población.
X4: Viviendas particulares habitadas con refrigerador	
X5: Población nacida en la entidad	
X6: Población afiliada a servicios médicos	
X8: Ingresos trimestrales por vivienda	
X9: Gastos trimestrales por vivienda	
X12: Grado de urbanización	
X13: Tasa de empleo	

Fuente: INEGI 2020

III.3 ANÁLISIS EXPLORATORIO DE VARIABLES

Para cumplir con el segundo objetivo del trabajo se tuvo que conocer la autocorrelación de la variable dependiente, en donde se aplicó el índice de Moran para conocer su comportamiento, si los datos tienden

a agruparse se habla de una autocorrelación positiva, si los datos se dispersan se convierte en una autocorrelación negativa, pero si estos valores se encuentran dispersos o distribuidos de forma aleatoria no cuenta con una autocorrelación espacial [30]. En el caso de tener una autocorrelación positiva se procederá a utilizar la regresión lineal múltiple espacial.

El análisis exploratorio se realizó en Minitab que es un software que permite análisis estadísticos y visualizaciones de gráficas. Para las 14 variables se necesitó conocer su comportamiento con el objetivo de averiguar qué tan frecuentes son esas variables en la aparición del fenómeno que se quiere conocer en este caso la generación de residuos. El software aplicó la prueba de Kolmogorov-Smirnov, con un valor de $\alpha = 0.05$ de acuerdo con las bibliografías consultadas [10], descubriendo la normalidad de los datos nos ayuda a conocer que tan frecuentes son esas variables para que se generen los RSU. Cuando la prueba no cumple la hipótesis de la normalidad, no se puede comprobar la presencia o ausencia de homocedasticidad por lo que procederá a transformarse con el objetivo de normalizar la distribución de no paramétrica de los datos originales [31].

III.4 MULTICOLINEALIDAD

Las variables que mostraron una distribución normal pasaron por el método de multicolinealidad en donde se fueron eliminando variables que demostraron una relación lineal con otras o con la combinación de otras. Si hay presencia de multicolinealidad esto aumenta los errores estándar y la incertidumbre en las estimaciones de los coeficientes de regresión, por lo que es de vital importancia eliminar cualquier multicolinealidad. Se empleó el factor de inflación de la varianza (FIV) y el coeficiente de correlación Pearson (r), cada uno de estos métodos tienen un objetivo diferente, FIV se encargará de detectar la correlación multivariada, mientras tanto r será empleado para detectar la asociación bivariado (*ver ecuaciones 6 y 7*) [7].

$$(6) \quad FIV = \frac{1}{(1 - R_k^2)} \quad r = \sqrt{1 - \frac{SSE}{SS_{YY}}} \quad (7)$$

Donde (1):

FIV: Factor de la inflación de la varianza.

R_k^2 : Es la ecuación de regresión de las variables explicativas.

Donde (2):

r: coeficiente de correlación Pearson.

SSE: Mínimos cuadrados.

SS_{yy}: Suma de los cuadrados de la diferencia de los datos cuadrados.

El valor de r indica la relación entre dos o más variables, en este método no se tiene como tal un valor umbral definido, por lo que se tuvo que recurrir a las bibliografías y se comprobó que un valor umbral adecuado para r oscilaba entre 0.50 y 0.60, de acuerdo con esto se trabajó con un valor umbral de 0.60 con el fin de identificar la correlación entre las variables explicativa [16].

Las variables transformadas no presentaron normalidad por lo que se empleó el coeficiente de spearman que es una prueba no paramétrica utilizando un valor de corte de 0.5 (*ver ecuación 8*) [24].

$$rR = 1 - \frac{6 \sum_i d_i^2}{n(n^2 - 1)} \quad (8)$$

rR: Coeficiente de consecuencia de spearman

d_i: diferencia entre los dos rangos de cada observación

n: número de observaciones

6: que es una constante

El valor de FIV fue calculado por R² de la ecuación de regresión, las variables explicativas que fueron denotadas por k se analizaron como variables dependientes, mientras que las otras se examinaron como variables independientes, por último se calculó el FIV para cada variable explicativa k.

Existen distintos valores de corte de FIV que es utilizado, los valores más comunes de umbral suelen ser 10 y 4, por lo que el valor de umbral que se empleó en este trabajo fue de 7.5 [13]. Si una variable hubiera presentado valores umbrales mayores de 7.5 indicaría presencia de multicolinealidad. [12].

III.5 DESARROLLO DEL MODELO

El tercer objetivo de trabajo se cumplió cuando se obtuvo las variables que pasaron el método de Spearman y FIV. El desarrollo del modelo busca establecer relaciones entre las variables dependientes que están registradas en el área de observación. El modelo se basa en la regresión lineal múltiple (Ecuación 9):

$$Y = \beta_0 + \beta_1 X_{1I} + \beta_2 X_{2I} + \dots + \beta_P X_{PI} + \epsilon_I \quad (9)$$

Donde Y_i es la observación de la variable de interés, X_{iq} son las variables explicativas, β_q es el coeficiente de regresión que mide la influencia de la variable explicativa, ϵ_i es el error aleatorio, que puede ser causado por variables no controladas o variabilidad muestral. Este modelo es válido para entender cuáles son las causas de la variable independiente con las variables explicativas.

Para determinar los coeficientes de regresión se aplicó el método de mínimos cuadrados con el objetivo de minimizar la suma de los errores cuadráticos (Ecuación 10).

$$SSE = \sum_{i=1}^n (Y_i - \bar{Y}_i)^2 \quad (10)$$

Donde Y_i es el valor de cada variable explicativa y \bar{Y}_i es el valor predicho.

III.6 VALIDACIÓN DEL MODELO DE REGRESIÓN

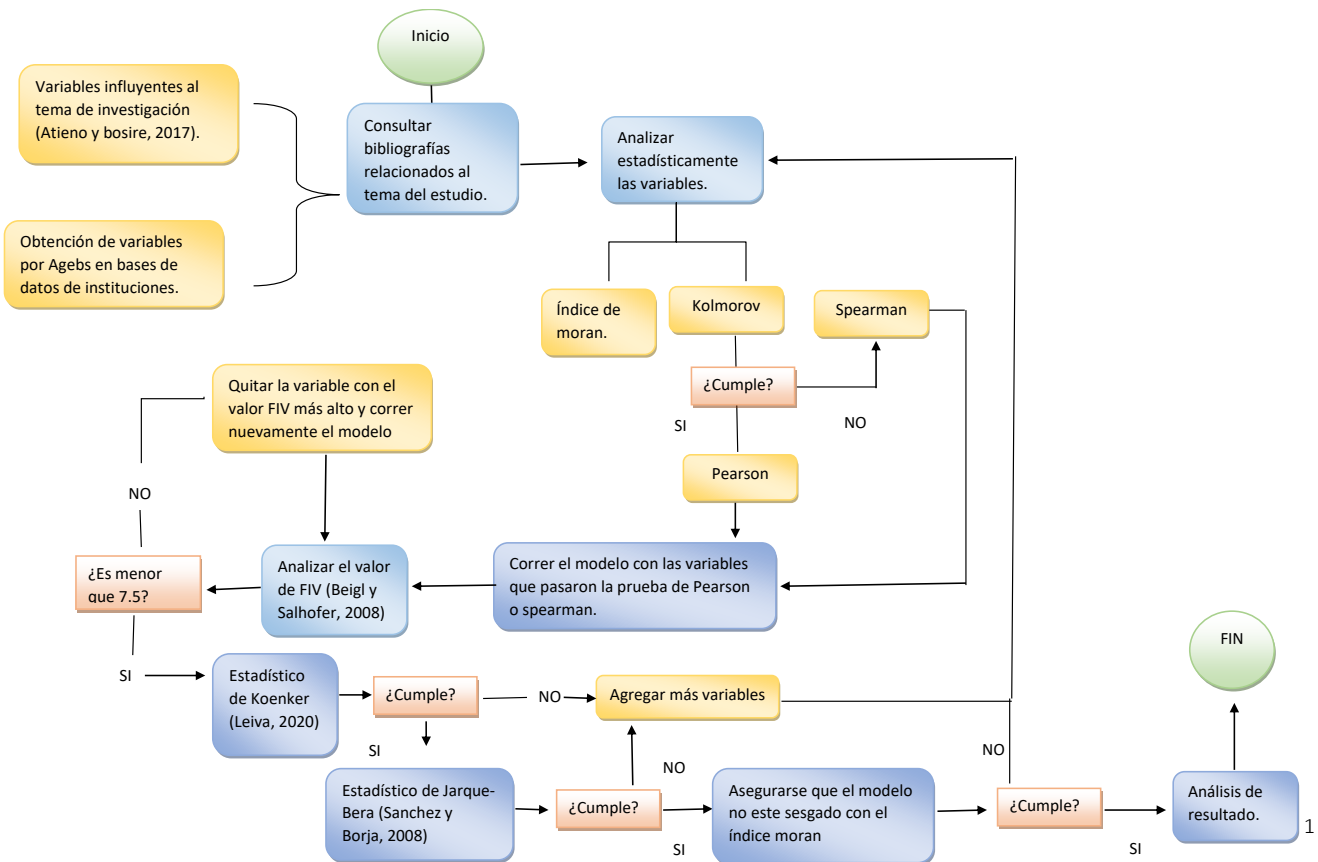
Esta sección se enfocó en cumplir con el último objetivo del trabajo, en donde se validó el modelo de regresión, para la significancia de las variables que se presentan en el modelo de regresión se aplicó el estadístico de Koenker con el objetivo de averiguar si existe una constante varianza de error lo que se denomina como homocedasticidad, si no se encuentra una constante varianza de error quiere decir que el modelo es heterocedástico [26]. El estadístico de Jarque-Bera se empleó con el fin de averiguar si los residuos del modelo tienen una distribución normal, este tipo de estadístico es muy utilizado en los modelos de regresión con el objetivo de ver si el modelo está sesgado o no [28].

Para la efectividad de las variables independientes en explicar la variable dependiente se utilizó la R^2 Ajustada, Caridad et. al [32] afirma que muchos estudios de modelos se aceptan aquellos con R^2 sumamente bajos que están en el rango de 0.4 a 0.5 (40 a 45%) o menores, en este trabajo se propuso alcanzar una R^2 Ajustada de 0.9 (90%). Por último, se analizó los residuos del modelo utilizando el índice de Moran con el objetivo de que los residuos estén libres de autocorrelación espacial.

III. 7 DIAGRAMA GENERAL DE LA METODOLOGÍA

Se realizó un diagrama de flujo en donde se puede observar un resumen general de la metodología que se implementó en el trabajo actual (figura 2), en donde primero se consultó bibliografías relacionadas al tema de trabajo y a la vez se obtuvo las variables formando una base de datos por AGEBS del municipio de Berriozábal, se aplicó distintas pruebas como lo son el índice de moran, Pearson, Spearman, FIV con el fin de descartar variables que no influyan con la variable Y1, se aplicó también los estadísticos de Koenker, Jarque Bera y nuevamente el índice de moran pero en los residuos del modelo con el fin de observar si están correlacionados espacialmente.

Figura 2 Diagrama de proceso de la metodología

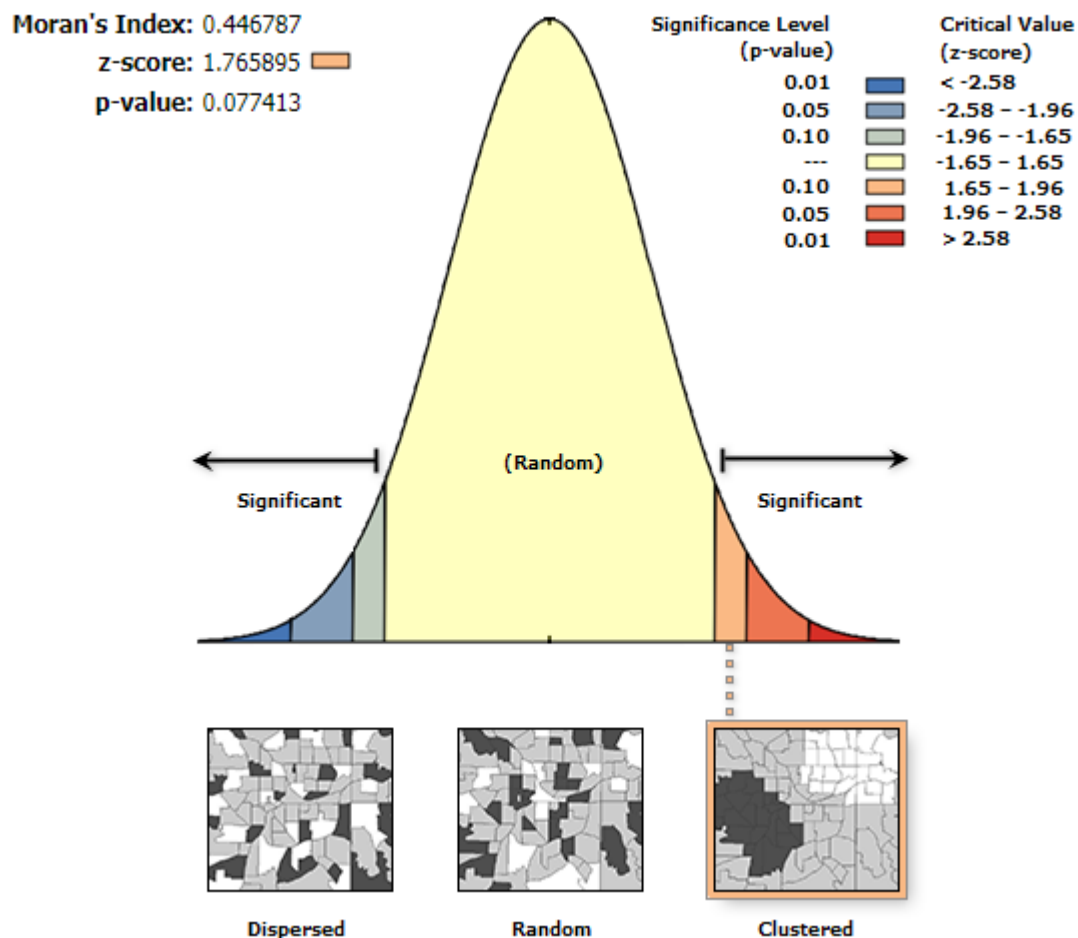


IV. RESULTADOS Y DISCUSIÓN

IV.1 ANÁLISIS EXPLORATORIO DE VARIABLES

Se aplicó el índice de Moran para conocer el comportamiento de la variable Y1, en donde se puede observar que la variable Y1 cuenta con un comportamiento de agrupación por lo que se cumple la hipótesis de que cuenta con una correlación espacial, por lo cual se puede emplear un método de regresión espacial (figura 3). Si la variable Y1 contara con un comportamiento disperso tiende a decir que no existe correlación espacial, por lo tanto lo indicado sería emplear una regresión múltiple convencional.

Figura 3 Comportamiento de la variable Y1 aplicando el índice de Moran



IV. 1. 1 CORRELACIÓN DE LAS VARIABLES

Se aplicó la prueba de normalidad de Kolmogorov-Smirnov para explorar las variables en donde se presentó normalidad en dos variables la cuales se representan en las figuras 4 y 5, se puede apreciar puntos de color azul que se ajusta a la variable dependiente que está siendo representada como una línea en diagonal, esto indica que las variables explicativas están correlacionadas con la variable dependiente, otra forma de verlo es que esas variables tuvieron un valor de $\text{Alpha} \geq 0.05$

Figura 4 Gráfica de normalidad de tasa de empleo

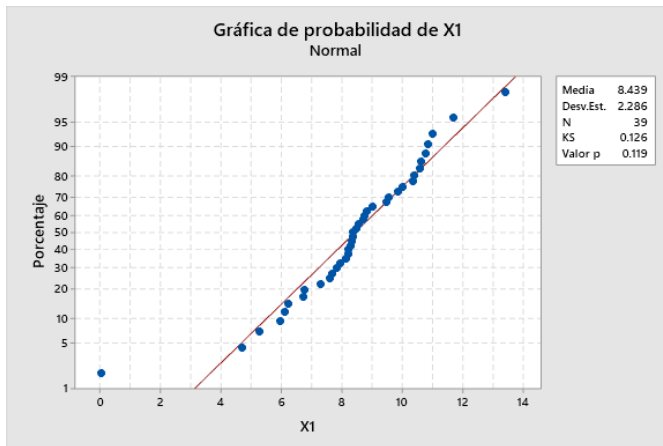
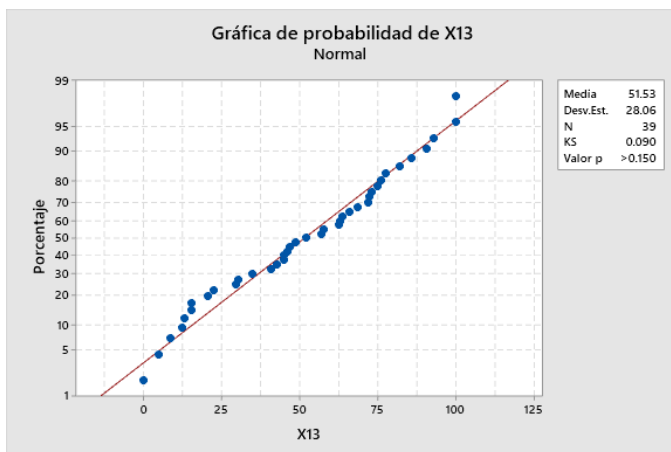


Figura 5 Gráfica de normalidad de Grad. Prom. Esc.



Las variables que no demostraron normalidad fueron las que obtuvieron un valor de P inferior al Alpha que es de 0.05, por lo que se procedió la transformación de variables con el fin de normalizar la distribución, se aplicó el método de logaritmo natural (Ln) con el objetivo de obtener una distribución simétrica de los datos, en donde se consiguió 5 variables que aprobaron el proceso de normalidad (tabla 5).

Tabla 5. Variables que aprobaron normalidad

Variable	Variables normales		Variables transformadas
	Kolmogorov-Smirnov		Kolmogorov-Smirnov
	Valor-P		Valor-P
X1: Grad. Prom. Esc.	0.119		
X4: VPH. Refrig	0.010		0.081
X5: Pob. Nac. Ent	0.010		0.093
X6: Pob. Afi. Serv. Med.	0.010		0.150
X8: Ingr. Trim. Viv.	0.010		0.102
X9: Gas. Trim. Viv.	0.010		0.102
X13; Tasa de empleo	0.150		

En total serian 7 variables con las que se trabajaría, pero se descarta la variable dependiente por lo que se tuvo que recurrir a la prueba no paramétrica de Spearman, con el objetivo de no destacar la variable Y1 (tabla 6).

Tabla 6. Prueba de Spearman

	Y1	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
	0.239										
X2	0.964	0.397									
X3	0.084	-0.594	-0.086								
X4	0.994	0.302	0.977	0.028							
X5	0.448	0.003	0.402	0.196	0.435						
X6	0.361	-0.037	0.338	0.156	0.352	0.894					
X7	0.767	-0.132	0.628	0.395	0.737	0.380	0.252				
X8	0.997	0.267	0.974	0.047	0.997	0.441	0.357	0.747			
X9	0.997	0.267	0.974	0.047	0.997	0.441	0.357	0.747	1.000		
X10	-0.038	0.007	0.013	-	-	-	0.007	-	-	-	
				0.076	0.037	0.099		0.184	0.048	0.048	
X11	0.124	0.242	0.190	-	0.154	-	-	0.040	0.138	0.138	-
				0.174		0.093	0.130				0.059
X12	1.000	0.239	0.964	0.084	0.994	0.448	0.361	0.767	0.997	0.997	-
											0.038
X13	0.326	0.111	0.388	-	0.321	0.133	0.183	0.234	0.325	0.325	0.154
				0.114							
X14	0.615	0.312	0.626	-	0.618	0.403	0.360	0.451	0.614	0.614	0.026
				0.037							
	X11	X12	X13								
X1											

X2			
X3			
X4			
X5			
X6			
X7			
X8			
X9			
X10			
X11			
X12	0.124		
X13	0.069	0.326	
X14	0.121	0.615	0.260

De acuerdo con la tabla 6 se puede observar que las variables que pasaron la prueba fue X2, X4, X7, X9, X12 y X14, esto se debe a que obtuvieron valores mayores a 0.5, lo cual indica correlación con la variable dependiente.

IV. 1. 2 VARIABLES NO SIGNIFICATIVAS

El análisis de spearman arrojó que las variables X1, X3, X5, X6, X10, X11 y X13 (tabla 4) no están estadísticamente correlacionadas con la variable dependiente, por valores ≤ 0.5 (tabla 6). Se puede observar que la X10 es la menos correlacionada con la generación de residuos por tener valores -0.0038 esto se debe a que la variable de temperatura es más utilizada para cuantificar los subproductos de los residuos, pero de igual manera se tomó en cuenta esta variable para conocer si se puede relacionar con la generación al día de RSU esto también aplica con la segunda variable menos correlacionada X11 (precipitación).

IV.2 MULTICOLINEALIDAD

Las variables que pasaron la prueba de spearman se introdujeron en el software de Arcgis para crear el modelo de regresión con el método de mínimos cuadrados, se generó un reporte para analizar el resultado del modelo de regresión (tabla 7).

Tabla 7. Primera corrida de modelo

Variable	Coficiente [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	-0.023459	0.010454	-2.243980	0.031883*	0.011270	-2.081508	0.045471*
X2	0.000074	0.000144	0.510462	0.613230	0.000110	0.670465	0.507371	65.872722
X4	0.000261	0.000693	0.376950	0.708699	0.000327	0.798303	0.430580	608.768282
X7	0.000130	0.000213	0.610637	0.545752	0.000098	1.327828	0.193629	3.304677
X9	-0.000000	0.000000	-0.377993	0.707931	0.000000	-0.613008	0.544203	≥1000.0
X12	283.662404	0.058532	4646.312140	0.000000*	0.042799	6627.814780	0.000000*	557.592030
X14	0.000000	0.000000	0.398532	0.692884	0.000000	0.618885	0.540373	1.321570

Se hizo énfasis en la columna de FIV (Factor de inflación de la varianza), como se puede notar los valores exceden del valor de corte que es de 7.5 lo cual quiere decir que hay correlación multivariada en dichas variables, por lo que se eliminó la variable con el valor FIV más alto del modelo y nuevamente se corrió el modelo con las variables sobrantes.

Tabla 8 Cuarta corrida del modelo

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_PR [b]	VIF [c]
Intercept	-28.335829	58.508136	-0.484306	0.631185	30.547592	- 0.927596	0.359970
X2	1.639457	0.122410	13.393138	0.000000*	0.179063	9.155739	0.000000*	1.351380
X7	2.626375	0.779977	3.367247	0.001857*	1.256986	2.089423	0.044006*	1.260096
X14	0.001239	0.000607	2.042185	0.048725*	0.000415	2.983269	0.005170*	1.100960

De acuerdo con la tabla 8 la cuarta corrida el valor de FIV ya no excede del valor de corte por lo que se puede concluir que no hay correlación multivariada.

IV. 2. 1. VARIABLES SIGNIFICATIVAS

Las variables que se emplearon en el modelo de regresión fueron X2, X7 y X14 (ver tabla 4) ya que obtuvieron valores bajos de FIV. Durante el proceso de Spearman se encontró la variable más significativa del modelo que fue viviendas totales con un valor de Spearman de 0.964 lo cual quiere decir que es la variable que más se relaciona a la generación de residuos, esta es una de las variables que se ha empleado en los trabajos de [13] [33].

La variable de comercios es una que no se ha encontrado en las literaturas, tuvo un valor de spearman de 0.764 lo cual nos indica que es la segunda variable que se relaciona a la generación de residuos, por lo que se tiene que tomar en cuenta para los posteriores estudios que se realicen en otros lugares.

La variable de densidad de la población es una de las más utilizadas en los métodos predictivos de residuos, ya que entre más poblada este la zona más residuos se genera en ella. La densidad de población ha sido empleada en el trabajo de Keser y Dusgun [15] dando valores de FIV de 1.55, los valores de FIV que arrojó en el trabajo actual fueron de 1.30 por lo que es una variable que se tiene que en pocas ocasiones se relaciona con otras variables explicativas.

IV.3 VALIDACIÓN DEL MODELO

Se enfatizó en la columna de Probability [b] y Robust_Pr [b] en donde los valores de las variables X2, X7 y X14 fueron menores al valor de corte de 0.05 por lo que nos indica que las variables de “viviendas totales”, “comercios” y “densidad de la población” son significativas para el modelo (tabla 8).

Tabla 9 Diagnostico de OSL

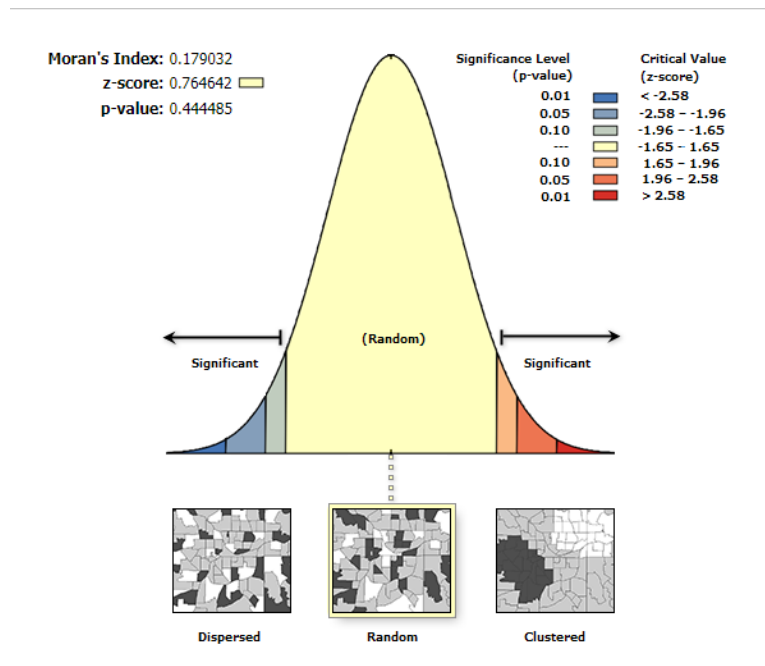
Input Features:	Base_de_datos_2	Dependent Variable	Y1
Number of observations	39	Akaike's Information Criterion (AICc) [d]:	549.864161
Multiple R-Squared [d]:	0.906016	Adjusted R-Squared [d]:	0.897960
Join F-Statistic [e]:	112.467392	Prob($F \geq F$), (3.35) degrees of freedom:	0.000000*
Join Wald Statistic [3]:	962.192413	Prob($F \geq \text{Chi-squared}$), (3) degrees of freedom	0.000000*
Koenker (BP)Statistic [f]:	26.119564	Prob($F \geq \text{Chi-squared}$), (3) degrees of freedom	0.000009*
Jarque-Bera Statistic [g]:	0.245824	Prob($F \geq \text{Chi-squared}$), (2) degrees of freedom	0.884341

De acuerdo con la tabla 9 se observó que el estadístico de Jarque-Bera es mayor que 0.05 con valores de 0.245824 por lo que se cumple la hipótesis nula, lo cual nos indica que los residuos se distribuyen de manera normal, en el dado caso en donde no se cumpla la hipótesis sostiene que falta una o más variables claves al modelo de regresión, el software nos indica que si no cuenta con un asterisco al lado del resultado

quiere decir que no está sesgado el modelo, así que se cumple la hipótesis. La R2 ajustada nos dio un valor de 0.90 lo cual quiere decir que nuestro modelo explica el 90% del problema de la generación de RSU.

Por último, se aplicó el índice de Moran para asegurarse que los residuos del modelo no tengan una correlación espacial (figura 6).

Figura 6 Índice de moran de residuos del modelo

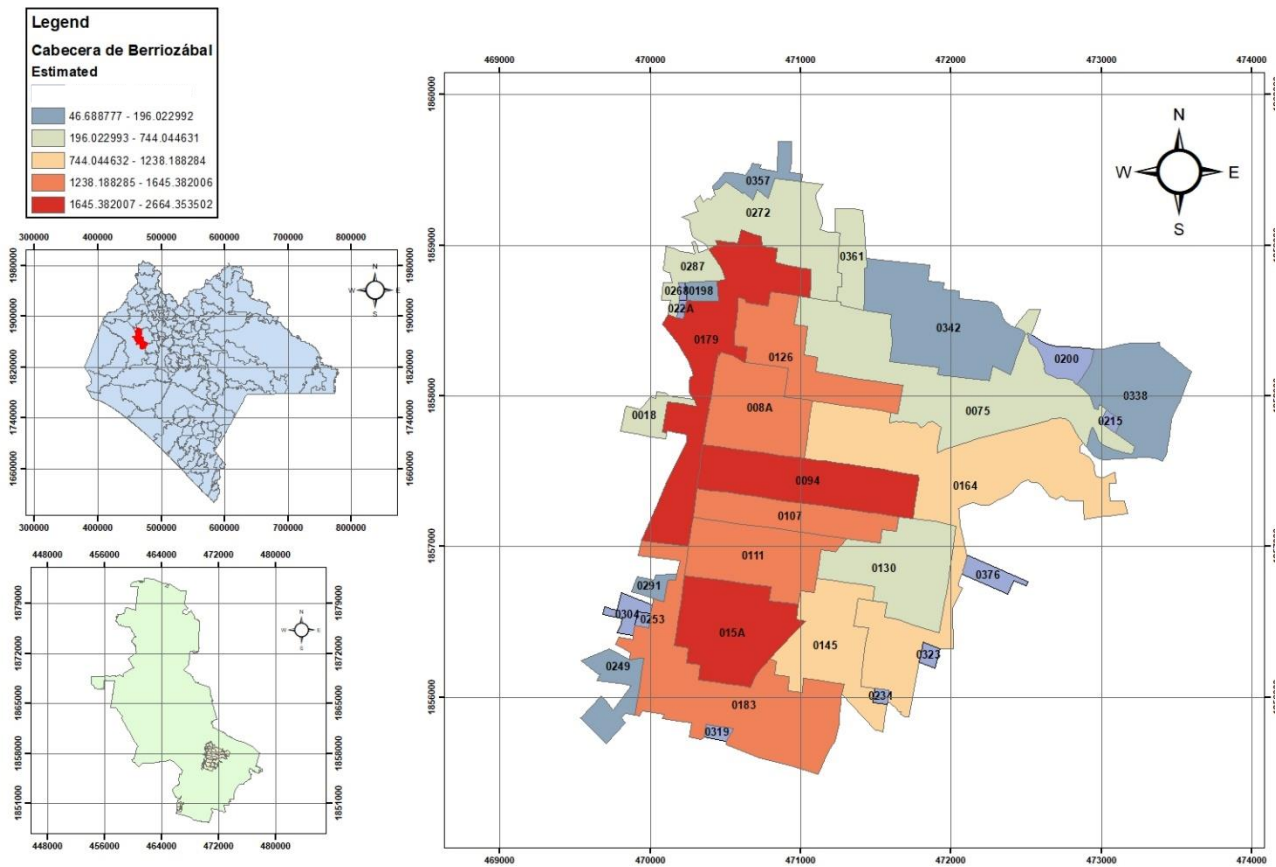


De acuerdo con la figura 6 el valor p de los residuos es un total de 0.444485 lo cual quiere decir que no tiene una correlación espacial. Si el valor de los residuos fueran ≥ 1 tendrían correlación espacial positiva y esto indica que el modelo está sesgado y requiere una o más variables independientes.

IV.4 ESTIMACIÓN DE RSU DEL MODELO DE REGRESIÓN

El análisis final que nos arrojó muestra un mapa de la cantidad de residuos en kg/día que se generan en distintas áreas de la cabecera municipal de Berriozábal (figura 7), se puede observar que el rango oscila entre 46.68 kg/día – 2664.35 kg/día, esto se debe por la concentración urbanística de las distintas zonas. En el centro del municipio rumbo al oeste, las zonas con un tono rojo son aquellas en donde se genera una gran cantidad de RSU, mayormente son zonas céntricas en donde se encuentran mercados, establecimientos comerciales y viviendas por lo cual su generación de RSU es mayor que las otras.

Figura 7 Mapa de estimación de RSU en Berriozábal, Chiapas



Las variables que demostraron significancia en el modelo de regresión de RSU fueron “Comercios”, “Viviendas totales” y “densidad poblacional” por lo que se realizaron histogramas con el objetivo de visualizar si las áreas con mayor generación de RSU concuerda con la alta tasa de comercios, viviendas y densidad poblacional (figuras 8, 9 y 10).

Figura 8 Histograma de viviendas totales de Berriozábal

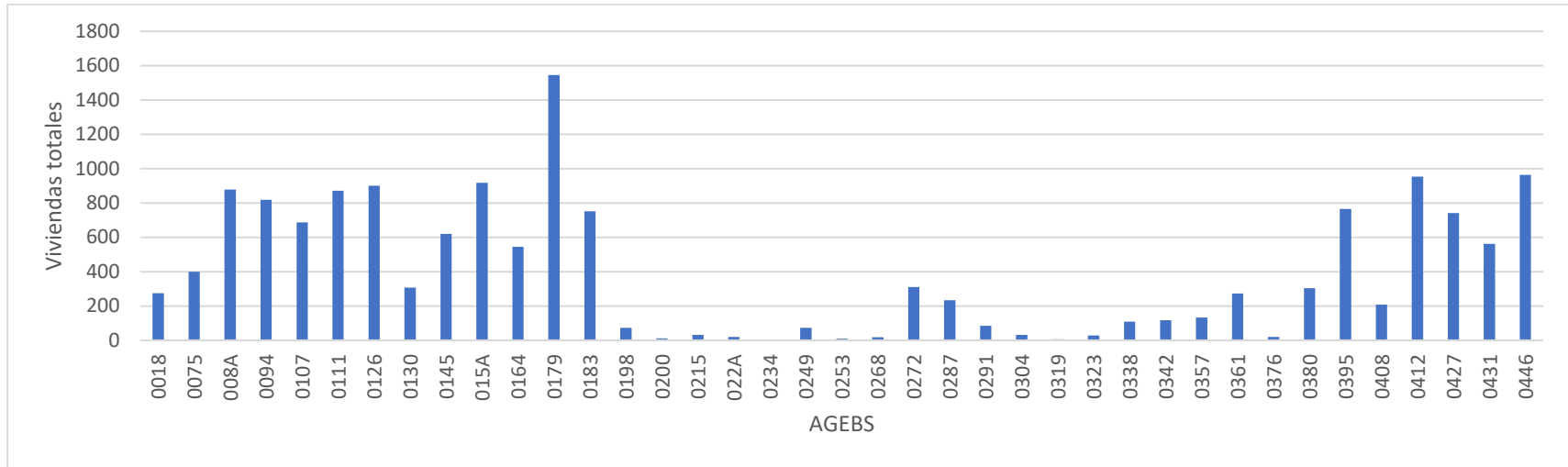
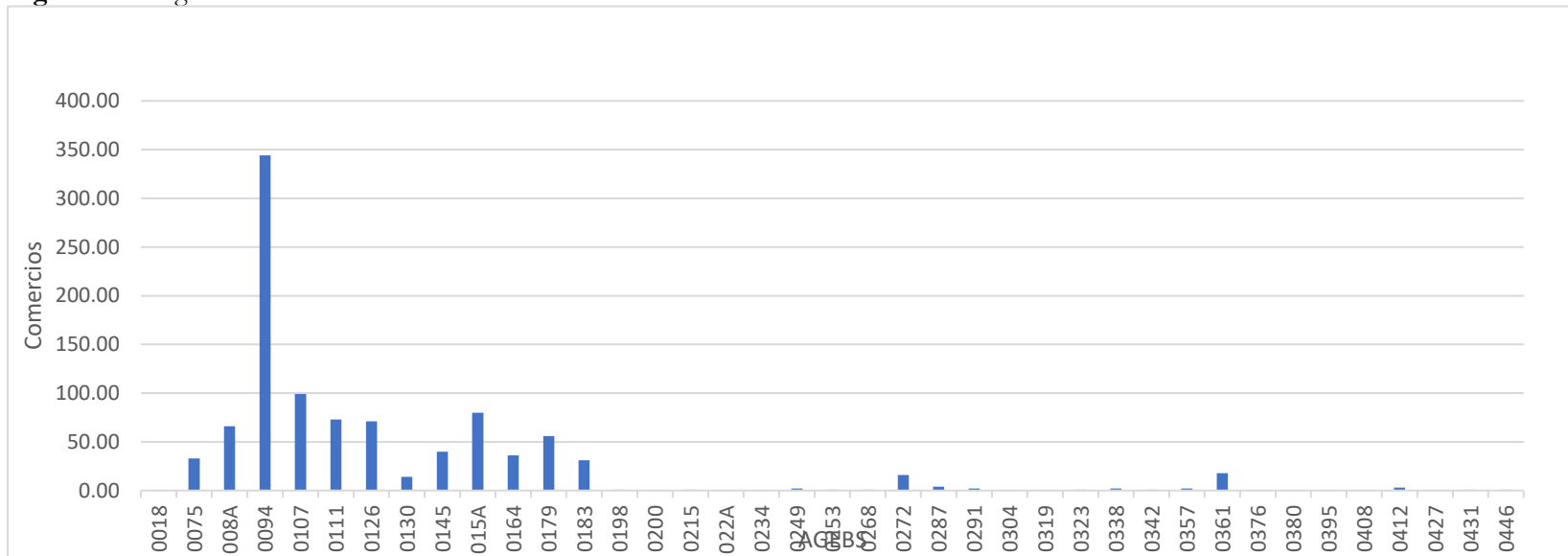
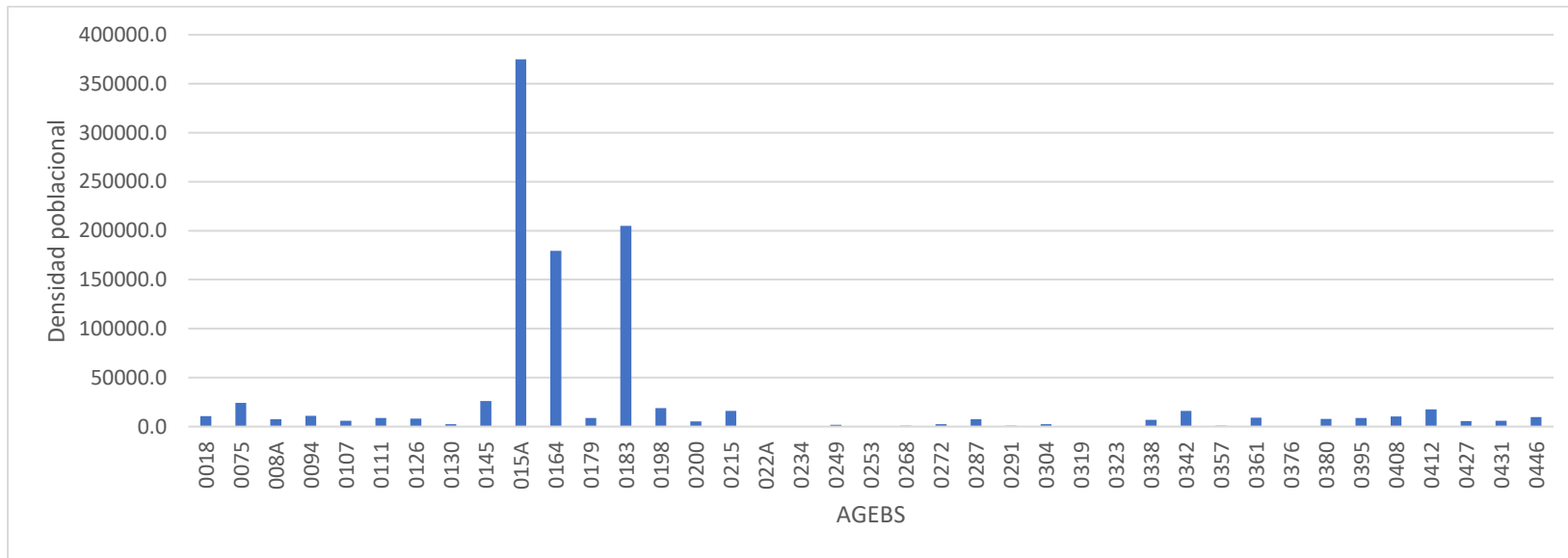


Figura 9 Histograma de comercios totales de Berriozábal



l, Chiapas

Figura 10 Histograma de densidad poblacional de Berriozábal, Chiapas



De acuerdo con los histogramas se puede observar que las Agebs 0179, 015A y 0094 son las que más presentan densidad poblacional, establecimientos de comercios y viviendas totales, por lo que respecto al mapa de estimación de RSU concuerda que esas zonas son las que más se generan RSU, por lo que se comprueba que las características sociales y demográficas influyen significativamente para explicar la generación de RSU.

IV. 5 MODELO DE REGRESIÓN

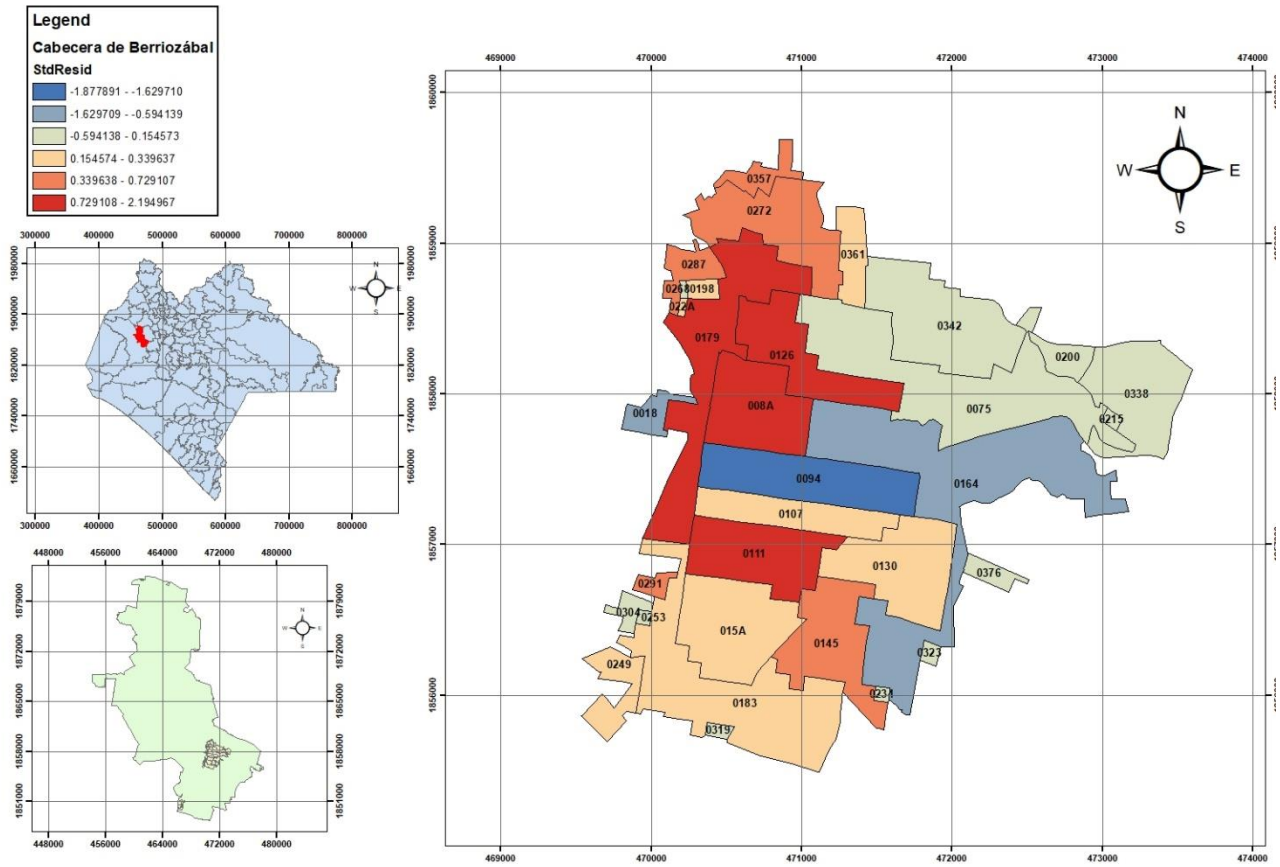
Los coeficientes de regresión fueron determinados por Arcgis, en donde se emplearon las variables que cumplieron el criterio de multicolinealidad (ecuación 11).

$$Y1 = -28.3 + 1.639 X2 + 2.626 X7 + 0.001239 X14 \quad (11)$$

El modelo consta de tres variables las cuales fueron “comercios” “viviendas totales” y “densidad poblacional”. El estadístico de Jarque-Bera nos indicó que todas las variables del modelo son estadísticamente significativas, por lo que el modelo puede ser empleado para la predicción de residuos.

Para el análisis de los residuos se aplicó el índice de Moran con el objetivo de averiguar si cuentan con correlación espacial. De acuerdo con Bucheli [30] la autocorrelación espacial está indicada por medio de valores que oscilan entre +1 y -1, en donde +1 indica autocorrelación positiva perfecta, -1 expresa autocorrelación negativa perfecta y un valor de 0 muestra presencia de patrones completamente aleatorios, como resultado nos arrojó un valor de P-Value: 0.44448 por lo que los residuos cuentan con una distribución aleatoria y no están correlacionados espacialmente (Figura 6). Se realizó un mapa de residuos estandarizados con el objetivo de visualizar en que zonas el modelo de regresión es factible (figura 11). Logramos notar que en las zonas con color rojo el modelo no es adecuado esto se debe por los valores que oscilan entre 0.729 – 2.19 de acuerdo con Rojo, 2007 [26] los residuos no tienen que tener valores ≥ 2 ni ≤ -2 , por lo que en este caso hay valores mayores a 2.

Figura 11 Mapa de residuos estandarizados del modelo



V. CONCLUSIÓN

V. 1 CONCLUSIÓN POR OBJETIVOS

- Este modelo fue construido con la finalidad de visualizar la generación de RSU por Agebs utilizando factores demográficos, económicos y sociales que han sido empleados previamente en distintos estudios de modelos predictivos. Los datos fueron proporcionados por INEGI por cada AGEB del municipio de Berriozábal, se han realizado estudios similares con datos recopilados mediante encuestas pero que carecen de confiabilidad, por lo que se recomienda utilizar datos que estén censados por instituciones gubernamentales con el objetivo de evitar errores en el modelo de regresión.
- Analizando las variables se encontró que el factor social fue uno de los que más se correlaciona con la generación de RSU dando valores de Spearman de 0.964 (viviendas totales). La variable de densidad de la población es una de las más utilizadas a nivel mundial, ya que entre más concentración de la población en ciertas áreas más se suele generar los residuos sólidos urbanos. Los comercios fue una variable nueva que se empleó para determinar la generación de RSU y fue la segunda más correlacionada en la generación de RSU por lo que se propone tener en cuenta esta variable para los posteriores estudios que se realicen. Dependiendo del trabajo de estudio se puede arreglar distintas variables, por ejemplo, la variable de temperatura se ha utilizado para cuantificar los subproductos de los residuos, en este caso fue la variable menos correlacionada a la variable dependiente dando valores de -0.038 por lo que no es indicada utilizarla en modelos que solo se enfocan en predecir la generación de RSU.
- Visualizando el mapa de la generación de residuos junto con los histogramas se aprecia que existe una gran correlación con las variables explicativas del modelo, por lo cual se puede concluir que entre más concentración de densidad de población, comercios y viviendas más se tiende a generar los residuos sólidos urbanos. Conocer la generación de RSU es importante con el fin de crear una gestión adecuada de RSU, por lo que es necesario crear herramientas que nos permitan visualizar la generación de RSU.
- El modelo de regresión obtuvo un valor de R² ajustada de 0.90 por lo cual el modelo explica un 90% la generación de residuos, existen varios modelos que explican el 98%, 78.2%, 67.6% [9, 33,13], los resultados difieren por las variables explicativas utilizadas y el sitio del estudio.

V.2 CONCLUSIÓN GENERAL

Este estudio concluyó en que las variables densidad poblacional, comercios y viviendas totales son buenos predictores para la generación de residuos sólidos urbanos. Tanto las variables demográficas y socioeconómicas permiten una exactitud de las cantidades generadas de RSU, por lo que la inclusión de otras variables que se enfoquen en esos ámbitos puede mejorar la precisión del modelo generado. De acuerdo con el modelo la tasa per cápita de RSU fue de 0.670 kg/hab/día, en cada AGEBS cambiaba la cantidad de RSU esto dependía de las variables, en otras palabras aumentaba o disminuía los residuos sólidos por la concentración de comercios, viviendas o la densidad poblacional en el lugar. Los distintos modelos que se han desarrollado en distintos países son herramientas útiles para el diseño de programas de gestión de residuos.

VI. BIBLIOGRAFÍA

- [1]. Hernández, M. y Rivera, E. “Diagnóstico De La Generación De Residuos Sólidos Urbanos En El Residencial Río Viejo”. *JEEOS* 1(1), 17-34. 2017.
- [2]. Ley General para la Prevención y Gestión Integral de los Residuos- LGPGIR. (2003). México: DOF 22-05-2015.
- [3]. Secretaria de medio ambiente e historia natural “*Guía para el manejo adecuado de los RSU y RME*” 1° Ed. Tuxtla Gutiérrez, 2019.
- [4]. Sáez, A. y Urdaneta, J. “Manejo de residuos sólidos en América Latina y el Caribe”. *Omnia* 20(3) pp. 121-135. Diciembre, 2014.
- [5]. INEGI, “*Censo de población y vivienda 2020*”. Recuperado de 21 de agosto del 2021 de la base de datos de INEGI. 2020. <https://www.inegi.org.mx/app/scitel/Default?ev=10>.
- [6]. Mendenhall, W., Beaver, R. y Beaver, B. “*Introducción a la probabilidad y estadística*” 13° Ed California, 2010.
- [7]. Beigl, P., Lebersorger, S. and Salhofer, S. “Modelling municipal solid waste generation: A review”. *Waste management* 28(1), 200-214. Diciembre, 2008.
- [8]. Castillo, L., Maya, J. y García, N. “Diagnóstico de la gestión de residuos sólidos urbanos (RSU) en la ciudad bacalar, Quintana Roo mediante el enfoque del nuevo institucionalismo”. *Quivera*, 18(1), 75-87. Junio, 2016.
- [9]. Rendón, A. “Caracterización de residuos sólidos”. *Activa* 1(4), 67-72. Febrero. 2012.
- [10]. Crescencio, C. y Caldelas, R. “Diagnóstico de la gestión de residuos sólidos urbanos en el municipio de Mexicali, México: Retos para el logro de una planeación sustentable”. *Información tecnológica* 25(3), 59-72. Febrero, 2014.
- [11]. Atieno, J. and Bosire, B. “Modeling Household Solid Waste Generation in Urban Estates Using Socio-Economic and Demographic Data, Kisumu City, Kenya”. *Scholars journal of arts, humanities and social Sciences* 5(5), 468-477 Junio, 2017.

- [12]. Intharathirat, R., Salam, A., Kumar, S. and Untong, A. "Forecasting of municipal solid waste quantity in a developing country using multivariate grey models". *Waste management* 39(1), 3-14. Julio, 2015.
- [13]. Kolekar, K., Hazra, T. and Chakrabarty, S. (2016). A Review on Prediction of Municipal Solid Waste Generation Models. *Procedia environmental sciences* 35(1), 238-244.
- [14]. L. Chen and W. Xin-wu. "Factors influencing municipal solid waste generation in China: A multiple statistical analysis study" *Waste manag res onlinefirst* 0(0), 1-8, 2010.
- [15]. Keser, S., Duzgun, S. and Aksoy, A. "Application of spatial and non-spatial data analysis in determination of the factors that impact municipal solid waste generation rates in Turkey". *Waste management* 32(1), 359-371. Mayo, 2012.
- [16]. Grazhdani, D. "Assessing the variables affecting on the rate of solid waste generation and recycling: An empirical analysis in Prespa Park". *Waste Management* 48(1), 3-13 September, 2016.
- [17]. Araiza, J., Moreno, C. y Pérez, J. "Cuantificación de residuos sólidos urbanos generados en la cabecera municipal de Berriozábal, Chiapas, México". *Revista Internacional de contaminación ambiental* 33(4), 691-699. DOI 10.20937/RICA.2017.33.04.12 2017 abril, 2017.
- [18]. Kingler, R., Olaya, J., Marmolejo, L. y Madera, C. "Plan de muestreo para la cuantificación de residuos sólidos residenciales generados en las zonas urbanas de ciudades de tamaño intermedio" *Rev. Fac. Ing. Univ. Antioquia* 48(1), 76-86, 2009.
- [19]. Salas, A., Zeferino, J., Salvador, C., Rosado, M., Ortiz, M. Morillas, A. y Benitez S. "Cuantificación y caracterización de micro plásticos y residuos sólidos urbanos en playa Zipolite, Oaxaca" *Ciencia y mar* 24 (71), 3-21, 2020.
- [20]. Coyago, E., Gonzales, K., Heredia, E. y Sánchez, R. "Recomendaciones para la caracterización y cuantificación de residuos sólidos universitarios. Caso de estudio: Universidad politécnica Salesiana, Campus Sur, Quito" *Ciencias de la vida* 23 (1), 60-71, 2016.
- [21]. Borrego, J. "Modelos de regresión para datos espaciales" (Trabajo fin de grado), Facultad de matemáticas, Universidad de Sevilla, España. 2018..
- [22]. Siabato, W. y Manrique, J. "La autocorrelación espacial y el desarrollo de la geografía cuantitativa" *Cuadernos de Geografía: revista colombiana de geografía* 28(01), 1-22, 2019.

- [23]. Tobler, W. "A computer model simulation of urban growth in the Detroit region". *Economic Geaography* 46 (2), 234-240. 1970.
- [24]. Barrera, M. "Uso de la correlación de Spearman en un estudio de intervención en fisioterapia" *Movimiento científico* 8 (1), 98-104, 2014.
- [25]. Moreno, J. y Bustillo, W. "La multicolinealidad en modelos de regresión lineal múltiple" *Ciencias técnicas agropecuarias* 21 (4), 80-83, 2012.
- [26]. J. Rojo. "Regresión lineal múltiple". 1º Ed. Madrid: Instituto de economía y Geografía, 2007.
- [27]. Leiva, M. "Determinación de la competitividad y los factores que afectan las exportaciones de jengibre peruano (2012-2019)" (Trabajo de investigación), Facultad de ciencias empresariales, Universidad San Ignacio de Loyola (USDL), Lima, 2020.
- [28]. Sánchez, F. y Borja, J. "Distribución de la estadística de Jarque y Bera para la prueba de normalidad en una serie temporal estacionaria con datos faltantes" *Entre ciencia e ingeniería* 2(4), 99-114, Noviembre, 2008.
- [29]. Camotlan, S. "Plan municipal de desarrollo Berriozábal, Chiapas". 2º Ed. Berriozábal, 2019.
- [30]. Bucheli, E. "Uso del Índice de Moran y LISA para explicar el ausentismo electoral rural en ecuador" *Revista geográfica* 1(1), 91-108 Julio, 2019.
- [31]. Díaz, V. "Errores estadísticos frecuentes al comparar dos poblaciones independientes" *Chil Nutr* 36(4) 1136-1138 diciembre, 2019.
- [32]. Caridad, W., Cabrera, A. y Fernández, L. "Criterios para la selección de modelos estadísticos en la investigación científica" *Rev. Cubana de Ciencia Agrícola* 37(1), 3-10, 2003.
- [33]. Nayseang, S. and Supachart, C. "Development of an appropriate model for forecasting municipal solid waste generation in Bangkok" *Energy procedia* 138 (2017) 907-912, 2017.