



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería Química

**DISEÑO DE INVESTIGACIÓN PARA LA CONSTRUCCIÓN DE UN MODELO DE
REGRESIÓN PARA LA PREDICCIÓN DEL COSTO DEL TRATAMIENTO DE AGUAS
RESIDUALES Y DESECHOS SÓLIDOS INDUSTRIALES DE UNA PLANTA EN GUATEMALA**

Ana Beatriz Márquez Alvarado

Asesorado por Msc. Ing. Gabriel Fong Mazariegos

Guatemala, noviembre de 2022

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**DISEÑO DE INVESTIGACIÓN PARA LA CONSTRUCCIÓN DE UN MODELO DE
REGRESIÓN PARA LA PREDICCIÓN DEL COSTO DEL TRATAMIENTO DE AGUAS
RESIDUALES Y DESECHOS SÓLIDOS INDUSTRIALES DE UNA PLANTA EN GUATEMALA**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

ANA BEATRIZ MÁRQUEZ ALVARADO
ASESORADO POR MSC. ING. GABRIEL FONG MAZARIEGOS

AL CONFERÍRSELE EL TÍTULO DE

INGENIERA QUÍMICA

GUATEMALA, NOVIEMBRE DE 2022

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANA	Inga. Aurelia Anabela Cordova Estrada
VOCAL I	Ing. José Francisco Gómez Rivera
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Kevin Vladimir Cruz Lorente
VOCAL V	Br. Fernando José Paz González
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANA	Inga. Aurelia Anabela Cordova Estrada
EXAMINADOR	Ing. César Ariel Villela Rodas
EXAMINADOR	Inga. Hilda Piedad Palma Ramos
EXAMINADOR	Inga. Mercedes Esther Roquel Chávez
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

**DISEÑO DE INVESTIGACIÓN PARA LA CONSTRUCCIÓN DE UN MODELO DE
REGRESIÓN PARA LA PREDICCIÓN DEL COSTO DEL TRATAMIENTO DE AGUAS
RESIDUALES Y DESECHOS SÓLIDOS INDUSTRIALES DE UNA PLANTA EN GUATEMALA**

Tema que me fuera asignado por la Dirección de la Escuela de Ingeniería Química con fecha 10 de noviembre de 2022.

Ana Beatriz Márquez Alvarado



EEPFI-PP-1786-2022

Guatemala, 10 de noviembre de 2022

Director
Williams G. Álvarez Mejía
Escuela De Ingenieria Quimica
Presente.

Estimado Ing. Álvarez

Reciba un cordial saludo de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería.

El propósito de la presente es para informarle que se ha revisado y aprobado el Diseño de Investigación titulado: **CONSTRUCCION DE UN MODELO DE REGRESION PARA LA PREDICCIÓN DEL COSTO DEL TRATAMIENTO DE AGUAS RESIDUALES Y DESECHOS SÓLIDOS INDUSTRIALES DE UNA PLANTA EN GUATEMALA**, el cual se enmarca en la línea de investigación: **Todas las áreas - Diseño y análisis de experimentos**, presentado por la estudiante **Ana Beatriz Márquez Alvarado** camé número **200819383**, quien optó por la modalidad del "PROCESO DE GRADUACIÓN DE LOS ESTUDIANTES DE LA FACULTAD DE INGENIERÍA OPCIÓN ESTUDIOS DE POSTGRADO". Previo a culminar sus estudios en la Maestría en ARTES en Estadística Aplicada.

Y habiendo cumplido y aprobado con los requisitos establecidos en el normativo de este Proceso de Graduación en el Punto 6.2, aprobado por la Junta Directiva de la Facultad de Ingeniería en el Punto Décimo, Inciso 10.2 del Acta 28-2011 de fecha 19 de septiembre de 2011, firmo y sello la presente para el trámite correspondiente de graduación de Pregrado.

Atentamente,

"Id y Enseñad a Todos"

Mtro. Gabriel Fong Mazariegos
Asesor(a)

Gabriel Fong Mazariegos
Ingeniero Mecánico Electricista
MSc. Ingeniería Sanitaria con
Especialización en Aguas Residuales
Colegado No. 20,640

Mtro. Edwin Adalberto Bracamonte Orozco
Coordinador(a) de Maestría



Mtro. Edgar Darío Álvarez Coti
Director
Escuela de Estudios de Postgrado
Facultad de Ingeniería





EEP.EIQ.1440.2022

El Director de la Escuela De Ingenieria Quimica de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del Asesor, el visto bueno del Coordinador y Director de la Escuela de Estudios de Postgrado, del Diseño de Investigación en la modalidad Estudios de Pregrado y Postgrado titulado: **CONSTRUCCION DE UN MODELO DE REGRESION PARA LA PREDICCIÓN DEL COSTO DEL TRATAMIENTO DE AGUAS RESIDUALES Y DESECHOS SÓLIDOS INDUSTRIALES DE UNA PLANTA EN GUATEMALA**, presentado por el estudiante universitario **Ana Beatriz Márquez Alvarado**, procedo con el Aval del mismo, ya que cumple con los requisitos normados por la Facultad de Ingeniería en esta modalidad.

ID Y ENSEÑAD A TODOS

Ing. Williams G. Álvarez Mejía; Mg.I.Q., M.U.I.E.
Director
Escuela De Ingenieria Quimica

Guatemala, noviembre de 2022

Decanato
Facultad de Ingeniería
24189101- 24189102
secretariadecanato@ingenieria.usac.edu.gt

LNG.DECANATO.OI.808.2022

La Decana de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería Química, al Trabajo de Graduación titulado: **DISEÑO DE INVESTIGACIÓN PARA LA CONSTRUCCIÓN DE UN MODELO DE REGRESIÓN PARA LA PREDICCIÓN DEL COSTO DEL TRATAMIENTO DE AGUAS RESIDUALES Y DESECHOS SÓLIDOS INDUSTRIALES DE UNA PLANTA EN GUATEMALA**, presentado por: **Ana Beatriz Márquez Alvarado**, después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:



Inga. Aurelia Anabela Cordova Estrada
Decana



Guatemala, noviembre de 2022

AACE/gaoc

ACTO QUE DEDICO A:

- Dios** Por ser el pilar de mi vida y darme la fuerza y sabiduría para persistir en mis metas.
- Mis padres** Edwin Rolando Márquez Arresis y Roxana Isabel Alvarado Marcos por apoyarme y creer en mí. Por sus ánimos en los momentos de debilidad y ser mi mayor ejemplo de perseverancia y tenacidad.
- Mi hermana** Yesenia Márquez, por siempre darme palabras de ánimo y su amor cuando sentía no poder más.
- Mi hermano** Edwin Márquez, por sus consejos y apoyo durante mi carrera.
- Mi tío** Edgar Márquez (q. e. p. d.) por su amor, ejemplo y celebrar mis triunfos con gran alegría.
- Mi familia** Porque gracias a su ejemplo de perseverancia y trabajo duro me enseñaron a ser una mujer fuerte para enfrentar la vida. Además de su amor a lo largo de mi vida. Especialmente a mi familia materna.

AGRADECIMIENTOS A:

Universidad de San Carlos de Guatemala	Por ser mi casa de estudios, la cuál me lleno de conocimientos, aprendizajes de vida y muchas vivencias.
Mis amigos de la Facultad	Por su apoyo, cariño y compañía hicieron este viaje más ligero, gracias a Damián Ochoa, Luis Pedro Ríos, Gerson Ortíz, José Guas, Diego Alfonzo.
Ing. William Fagiani	Por su apoyo y conocimientos en mi proceso académico.
Ing. Eduardo Herrera	Por ser el mejor amigo, novio y apoyo incondicional para concluir esta etapa de mi vida.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
LISTA DE SÍMBOLOS	VII
GLOSARIO	IX
1. INTRODUCCIÓN	1
2. ANTECEDENTES	3
3. PLANTEAMIENTO DEL PROBLEMA	9
3.1. Contexto general	9
3.2. Descripción del problema	9
3.3. Formulación del problema	10
3.4. Delimitación del problema	10
4. JUSTIFICACIÓN	11
5. OBJETIVOS	13
5.1 General.....	13
5.2 Específicos	13
6. NECESIDADES A CUBRIR Y ESQUEMA DE LA SOLUCIÓN	15
7. MARCO TEÓRICO.....	17
7.1. Estadística.....	17
7.1.1. Muestreo.....	17

7.1.1.1.	Muestreo no aleatorio.....	18
7.1.1.2.	Muestreo aleatorio.....	18
7.1.1.2.1.	Muestreo aleatorio simple.....	19
7.1.1.2.2.	Muestreo aleatorio estratificado.....	19
7.1.2.	Análisis de Regresión Múltiple	19
7.1.2.1.	Estimación de los coeficientes	20
7.1.3.	Métodos Estadísticos Gráficos	22
7.2.	El agua	27
7.2.1.	Usos del agua	27
7.2.2.	Características físicas	28
7.2.3.	Características químicas	28
7.2.4.	Características microbiológicas.....	28
7.2.5.	Tipo de agua	29
7.2.5.1.	Agua Cruda	29
7.2.5.2.	Agua Residual	30
7.2.5.3.	Agua Potable.....	31
7.2.6.	Métodos de Filtrado.....	31
7.2.7.	Calidad del Agua	32
7.2.8.	Norma COGUANOR	33
8.	PROPUESTA DE ÍNDICE DE CONTENIDOS	37
9.	METODOLOGÍA	39
9.1.	Características del estudio	39
9.2.	Unidad de análisis	39
9.3.	Operativización de las variables.....	40
9.4.	Fases del estudio	43

10.	TÉCNICAS DE ANÁLISIS DE INFORMACIÓN.....	47
11.	CRONOGRAMA.....	49
12.	FACTIBILIDAD DEL ESTUDIO	51
12.1.	Recurso humano	51
12.2.	Recursos tecnológicos.....	51
12.3.	Recursos financieros	51
12.4.	Acceso a información y permisos	52
12.5.	Equipo e infraestructura.....	52
13.	REFERENCIAS.....	53
14.	APÉNDICE	57

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Ejemplo de un modelo de regresión lineal múltiple en R.....	22
2.	Histograma con datos no normales.....	23
3.	Histograma con datos normales.....	24
4.	Representación de Box-Plot.....	25
5.	Representación gráfica de un diagrama de dispersión	26
6.	Ciclo hidrológico	29
7.	Cronograma	49

TABLAS

I.	Uso del agua en el mundo.....	28
II.	Calidad promedio del agua de mar.....	30
III.	Relación de las sustancias inorgánicas cuya presencia en el agua es significativa para la salud.	35
IV.	Operativización de Variables.....	40
V.	Presupuesto	52

LISTA DE SÍMBOLOS

Símbolo	Significado
%	Porcentaje
Q	Quetzales

GLOSARIO

Agua potable	Agua apta para el consumo humano que sirve para beber agua, preparar alimentos, higiene y fines domésticos.
Agua residual	Aguas procedentes de procesos productivos o del consumo humano. Previo tratamiento puede convertirse en aguas regeneradas susceptibles de reutilización si alcanzan los criterios de calidad adecuados para cada tipo de uso.
Calidad	Propiedad o conjunto de propiedades inherentes a algo, que permiten juzgar su valor.
Concentración	Magnitud que expresa la cantidad de una sustancia por unidad de volumen. Su unidad en el Sistema Internacional es el mol por metro cúbico (mol/m^3).
Costo	El coste es un concepto de la contabilidad que se refiere al valor del consumo de los recursos que han sido necesarios para poder producir productos o prestar servicios.
DBO	La demanda bioquímica de oxígeno consumida por los microorganismos para descomponer la materia

orgánica presente en una muestra durante un período de tiempo y temperatura específicos.

Desechos sólidos	Residuos generados por los humanos día a día, los cuales se caracterizan por presentarse en estado sólido no biodegradable, atributo que los diferencia de los desechos líquidos y gaseosos. Son considerados basura porque terminan siendo tirados en vertederos al creerse que han cumplido con su ciclo de uso.
DQO	La demanda química de oxígeno necesaria para oxidar la materia orgánica en una muestra por medio de un agente químico.
Filtrar	Hacer pasar un fluido por un filtro.
Muestra	Parte o porción extraída de un conjunto por métodos que permiten considerarla como representativa de él.
Nocivo	Dañoso, pernicioso, perjudicial.
Norma	Regla que se debe seguir o a que se deben ajustar las conductas, tareas, actividades, etc.
Parámetro	Variable que, en una familia de elementos, sirve para identificar cada uno de ellos mediante su valor numérico.
Partícula	Parte pequeña de materia.

Poroso	Que tiene poros.
Proceso	Conjunto de las fases sucesivas de un fenómeno natural o de una operación artificial.
PTAR	Planta de tratamiento de aguas residuales.
PTDS	Planta de tratamiento de desechos sólidos.
Rango	Amplitud de la variación de un fenómeno entre un límite menor y uno mayor claramente especificados.

1. INTRODUCCIÓN

El agua es un líquido vital para varias actividades diarias del ser humano, tanto en el hogar como en la industria. La presente propuesta de investigación es una sistematización ya que es una secuencia ordenada de pasos para resolver el problema de investigación, partiendo de las aguas residuales, su tratamiento, el tiempo y materiales que se utilizaron para la obtención de agua potable con el fin de estimar el costo óptimo.

El problema en la calidad del agua ha sido motivo de discusión en los últimos años debido al aumento de contaminación causada por el mal manejo de los desechos en la industria en los hogares, ya que muchas veces desembocan en ríos, mares, lagos, etc. Como consecuencia de esto, el agua residual debe ser tratada para obtener nuevamente agua potable.

El estudio de este fenómeno se realizará con el interés de conocer cuál es el costo óptimo del tratamiento de agua y desechos sólidos en una planta de tratamiento de aguas residuales (PTAR), ya que no se cuenta con un sistema estandarizado de cobro, que tome en cuenta el tiempo del proceso, así como los materiales a emplearse, porque esto depende mucho de la muestra a tratarse, es decir, no será lo mismo si se procesa agua de fosas sépticas al agua de una industria alimenticia.

Se pretende llevar a cabo un análisis no experimental retrospectivo con alcance correlacional que contempla los datos históricos de la planta de tratamiento de aguas residuales y desechos sólidos para determinar el costo óptimo del proceso.

Con los datos de las variables DQO, DBO, peso, tiempo, relacionadas al tratamiento de aguas residuales se procederá a realizar un análisis de regresión para la construcción de un modelo matemático con variable dependiente el costo. Se espera conocer la relación entre las variables implicadas en la PTAR para estandarizar el costo en quetzales por metro cúbico.

En la primera fase se procederá a revisar los antecedentes, es decir, estudios previos y la información técnica que da soporte a la investigación. En la segunda fase se ordenará la información histórica de la PTAR. En la tercera fase se realizará el análisis de regresión en programas informáticos. En la cuarta fase se analizarán los resultados obtenidos y así obtener el modelo matemático que se ajuste mejor a los datos. El trabajo de investigación es factible ya que se cuentan con todos los recursos necesarios para llevarse a cabo todas las fases del estudio.

En el primer capítulo se describirá el marco referencial, las investigaciones que tienen relación con el tema a analizar, sirviendo de base para el estudio.

En el segundo capítulo se hace referencia al fundamento teórico del estudio, que está dividido en dos partes, la primera es la estadística que se conforma por los fundamentos teóricos que respaldan la investigación y la segunda parte se relaciona con una PTAR, para conocer las variables que están involucradas en el proceso.

En el tercer capítulo se presentarán los modelos matemáticos más eficientes, para luego seleccionar el que presenta un mejor ajuste.

En el cuarto capítulo se presentará la discusión de resultados, que consiste en detallar los hallazgos en la construcción del modelo de regresión.

2. ANTECEDENTES

Las investigaciones relacionadas al consumo de agua para el ser humano son de suma importancia dado que es un líquido primordial para la vida. Y por eso es importante que el agua cumpla con los parámetros establecidos por las normas vigentes en cada país.

En la investigación de Saravia (2016) evalúa la calidad del agua, a través de la determinación del índice ICA-NSF a partir de muestreos fisicoquímicos y bacteriológicos de los años 2007 al 2014. Y así deduce si el agua requiere o no un tratamiento previo para catalogarla como apta para consumo humano a partir de una escala del 0 al 100, que se divide en 5 rangos, donde para un ICA de 91-100 la categoría es excelente y el agua no requiere tratamiento o bien si se tiene un valor de ICA de 0-25 la categoría es muy mala y el agua no es apta para realizar tratamiento de agua para consumo humano, es decir, que no es aceptada para las actividades que impliquen un contacto directo con ella.

Debido a la contaminación actual de las aguas superficiales en muchos lugares del mundo, Hernández (2018) procede a analizar la percepción de la contaminación de arroyos urbanos en México, por el deterioro ambiental que se ha notado en los últimos años ya que el objetivo ha sido el poder encontrar medios para abastecer a la población de agua sin tomar en cuenta el lado ecológico. Para conocer el punto de vista de las personas, en su procedimiento estadístico realizó entrevistas y encuestas.

Chan y Peña (2014) investigan si es posible consumir el agua de la cuenca alta del Sis Iacán en Guatemala. Según la información presentada por el Instituto

de Agricultura, Recursos Naturales y Ambiente e Instituto de Incidencia Ambiental en 2006, afirma que lamentablemente los principales ríos de Guatemala están contaminados con microbios, elementos tóxicos y sedimentos. Para ello, deciden tomar dos puntos de muestreo en los ríos, uno en la cabecera y otro en la desembocadura, planteando la hipótesis que los resultados de los análisis de coliformes son mayores en la cabecera, encontrando que no existe suficiente evidencia para afirmar que las concentraciones son mayores en la cabecera, por ende, presenta mayor calidad microbiológicamente que en la desembocadura.

El agua es parte fundamental de la vida, pero cuando se trata de agua para consumo humano, es necesario monitorear parámetros que protejan la salud pública para evitar cualquier daño. A raíz de ello González (2018) realiza el estudio para estimar estadísticamente que riesgos potenciales son los que se deben monitorear para el control de calidad del agua y así cumplir con lo establecido en la Normatividad Nacional para el Agua Potable en Colombia. Dado que los parámetros fisicoquímicos y microbiológicos explican menos del 60 % el modelo estadístico, no lograron inferir a que se debe el riesgo asociado con la calidad del agua.

El estudio de Valdes, Samboni y Zanjón (2016) consiste en desarrollar un indicador de la Calidad del Agua a través de la Estadística Aplicada. Surge de la necesidad de conocer la calidad del agua de la subcuenca porque es utilizada para regar los campos de caña de azúcar, pero dado que, en el recorrido se unen aguas residuales domésticas e industriales, esto puede causar la contaminación en la industria del azúcar. Este índice está conformado por distintos parámetros fisicoquímicos que establecen el límite cualitativo y cuantitativo de la calidad del agua y así seleccionar la curva óptima que explique de mejor manera los datos. Para luego compararlo con métodos tradicionales para medir la calidad del agua

y saber que tanto difieren entre ellos. Como resultado, observaron que existe una gran similitud con los Índices de Calidad del Agua empleados usualmente.

El tener acceso a agua para beber garantiza que el ser humano sobreviva en el planeta, por ello, es importante el tener certeza de su calidad porque de no ser así puede causar malestares y enfermedades. Con el fin de conocer el estado actual de las dos quebradas que drenan la microcuenca El Limón en Honduras, Mejía (2005) realiza muestreos del agua en estudio para analizar la relación que existe entre la calidad y el manejo del agua que se está desarrollando en la zona y así poder inferir los contaminantes que afectan su calidad. Encontrando que los factores que afectan principalmente la calidad son la turbidez y sedimentación además de los coliformes fecales. El método estadístico empleado fue mediante encuestas para conocer la opinión de los pobladores de las posibles causas de contaminación y la disponibilidad para implementar métodos de desinfección.

El estudio de Zúñiga (2018) investiga la calidad del agua potable en la Región de Antofagasta en la costa del océano pacífico. Con el principal objetivo de conocer la calidad del agua potable, ya que, considera importante el que las personas tengan acceso a un suministro seguro y continuo tanto para la higiene personal como la preparación de alimentos. Con los resultados obtenidos, se infiere que el suministro del agua potable está dentro del rango aceptado por la normativa chilena. Realizando encuestas para conocer el punto de vista de los consumidores de la región.

Díaz (2015), observó potencial en estudiar la cuenca del Río Motagua por el mal uso que se da al agua, causando escasez y contaminación superficialmente y de forma subterránea. Para el análisis fisicoquímico se utilizó la base de datos del INSIVUMEH con el fin de conocer su comportamiento y

calcular el parámetro Índice de Calidad del Agua (ICA), y con ello proponer soluciones con el fin de mejorar y conservar los recursos naturales actuales.

Debido a la importancia que tiene el conocer la calidad del agua superficial y una forma de tomar decisiones en beneficio del futuro, Sotomayor (2016) se propone estudiar la evaluación de la calidad de las aguas superficiales mediante técnicas de estadística multivariante. Para ello utilizará los datos obtenidos de las muestras de la cuenta del Río Paute al sur de Ecuador por cinco años. Dicha información la analizará a través de regresiones múltiples y un análisis de componentes principales, con el fin de obtener un modelo matemático que sea más fácilmente interpretable, proporcionando una mejor idea sobre la calidad del agua y los factores que influyen en ello.

Aguilar (2017) realiza el análisis de la calidad y cantidad del agua para consumo humano en La Parroquia San Isidro de Patulú ubicado en el país de Ecuador. A raíz de la importancia que tiene el agua para el ser humano y aun así no se le da el lugar que merece, ya que no se tiene un adecuado manejo ni control. Por lo que, con la investigación pretende asegurar el bienestar de las personas a través de contar con la disponibilidad del agua. La forma de manejar los datos es a través de encuestas a los pobladores que ayudarán a identificar las actividades para las cuáles el agua es de vital importancia. Adicional análisis fisicoquímicos para conocer si el agua es apta o no para consumo humano.

Luego de esta revisión puede verse que la tendencia de las investigaciones está en conocer la calidad del agua y así determinar si es apta o no para consumo humano. Partiendo de la vital importancia de dicho líquido para el ser humano, tanto para consumo como para las actividades del diario vivir.

El tener estudios previos sobre el tema sirve de guía y apoyo para conocer como han abordado el problema y tomar en cuenta las recomendaciones para hacer un trabajo más eficiente.

3. PLANTEAMIENTO DEL PROBLEMA

3.1. Contexto general

El agua es un líquido vital para el ser humano, por ello la importancia de garantizar la calidad de esta, porque de lo contrario puede ocasionar graves problemas de salud en la población como dificultades respiratorias, gastrointestinales y en la piel.

El problema en la calidad del agua ha sido motivo de discusión en los últimos años debido al aumento de contaminación por el mal manejo de los desechos en la industria en los hogares, ya que muchas veces desembocan en ríos, mares, lagos, etc. Como consecuencia de esto, el agua residual debe ser tratada para obtener nuevamente agua potable.

Sumado a ello, el fin principal de toda empresa es generar rentabilidad y para ello es necesario conocer cuál sería el costo óptimo de la operación.

3.2. Descripción del problema

El estudio se realizará con base a la información de una planta de tratamiento de aguas residuales y desechos sólidos.

No se conoce la ecuación de regresión lineal que modele el costo por realizar el tratamiento de aguas residuales a las industrias en quetzales por metro cúbico. Por lo que, es necesario identificar cuáles son las variables que afectan de forma independiente al costo del tratamiento de agua no potable.

3.3. Formulación del problema

- Pregunta central

¿Cuál es el modelo de regresión que predice adecuadamente el costo del tratamiento de aguas residuales y desechos sólidos industriales de la planta?

- Preguntas auxiliares
 - ¿Qué variables tienen un efecto significativo sobre el costo de tratamiento de aguas residuales y desechos sólidos?
 - ¿Cuál es el grado de influencia e interacción entre las variables?
 - ¿Cuál es el modelo más adecuado o que se ajusta mejor al comportamiento del costo?

3.4. Delimitación del problema

Espacial: Ciudad de Guatemala.

Poblacional: Aguas residuales.

Categoría: Industrial.

4. JUSTIFICACIÓN

La línea de investigación en que se enfocará este estudio es la Estadística Multivariada porque enmarca el estudio de explicar la variable dependiente a través de la relación de otras variables.

El problema presentado es que no se tiene una ecuación que modele el costo de tratar aguas residuales para obtener agua potable ya que actualmente se realiza por volumen, pero no se toma en cuenta la industria de la que proviene, que es de suma importancia porque los materiales y tiempo llegan a diferir.

El tener una ecuación matemática para el costo del proceso es garantizar que la empresa no tenga pérdidas y cumpla su fin primordial que es el generar utilidad para mantenerse a flote.

El presente estudio es muy importante porque se logrará generar valor a la empresa de Tratamiento de Aguas Residuales ya que se le propondrá un modelo matemático que modele el costo óptimo de sus operaciones.

La motivación de realizar el presente estudio es aplicar herramientas estadísticas para el análisis de datos para perfeccionar los procesos relacionados.

El beneficio que se obtendrá es tener un modelo matemático para el costo explicado por variables intrínsecas del proceso.

El beneficiario será la planta de tratamientos de aguas residuales y desechos sólidos.

5. OBJETIVOS

5.1 General

Construir un modelo de regresión lineal para el costo del tratamiento de aguas residuales y desechos sólidos industriales a través de un análisis multinomial.

5.2 Específicos

- Identificar las variables más influyentes sobre el costo de procesos en la planta de tratamiento de agua y residuos sólidos a través de un análisis de correlación.
- Cuantificar el grado de influencia que tienen las variables identificadas, sobre el costo del tratamiento a través de un análisis de regresión para evaluar sus interacciones y maximizar la parsimonia del modelo construido.
- Proponer el modelo estadístico más adecuado para predecir el precio del tratamiento de agua y desechos sólidos a partir de las variables del proceso, a través del diagnóstico de las posibles ecuaciones de regresión analizando residuos y estadísticos asociados al ajuste.

6. NECESIDADES A CUBRIR Y ESQUEMA DE LA SOLUCIÓN

El propósito de este trabajo es el analizar qué variables explican mejor el costo de tratar aguas residuales y desechos sólidos para suplir una necesidad administrativa en una planta de tratamiento de agua. Para ello, como se menciona anteriormente, el estudio se realizará con base en un análisis de regresión con datos históricos de la empresa.

Lo más importante es estandarizar la forma en la que se calcula el costo del proceso. Para seleccionar las variables que explican el costo se realizará un análisis de correlación lineal previamente con el fin de construir una menor cantidad de posibles modelos para su evaluación.

Por tanto, se procederá a construir distintos modelos matemáticos para lograr observar cuál es el que presenta un mejor ajuste de los datos utilizando el análisis de residuos y los distintos estadísticos asociados con el nivel de ajuste.

Esto ayuda a transformar sistemas complejos de distintos parámetros a una forma más simple y eficiente.

7. MARCO TEÓRICO

7.1. Estadística

La estadística consiste en la recogida, cuantificación, síntesis, análisis e interpretación de la información relevante contenida en unos datos. Puede dividirse en dos grandes campos: estadística matemática y estadística aplicada.

La estadística matemática es el terreno de los profesionales de las Matemáticas y puede resultar inaccesible a especialistas en otras áreas. Exige una gran fuerza creativa, ya que implementa nuevos procedimientos que se utilizan para resolver problemas en los distintos campos del saber. Además, requiere un conocimiento minucioso de los principios matemáticos y exige el nivel de abstracción y generalización propio de las ciencias matemáticas.

La estadística aplicada versa sobre cómo y cuándo utilizar cada procedimiento y cómo interpretar sus resultados. Estudia, por tanto, la transferencia de los métodos de la estadística matemática a otras disciplinas, como la economía, la publicidad, la sociología o la medicina. Bioestadística Amigable.

7.1.1. Muestreo

Una muestra es un subconjunto de una población. Para que las inferencias que hacemos sobre la población a partir de la muestra sean válidas, debemos obtener muestras que sean representativas de ella. Con mucha frecuencia nos sentimos tentados a elegir una muestra seleccionando a los miembros más

convenientes de la población. Tal procedimiento podría conducir a inferencias erróneas respecto a la población. Se dice que cualquier procedimiento de muestreo que produzca inferencias que sobreestimen o subestimen de forma consistente alguna característica de la población está sesgado. Probabilidad y estadística para ingenieros.

La selección de las unidades debe ser aleatoria, lo que equivale a decir probabilística, con el fin de obtener una muestra representativa de la población, respecto de todas o algunas de sus características, para las cuales el valor del parámetro es desconocido. En caso contrario se habla de una muestra no aleatoria. Estadística y muestreo.

7.1.1.1. Muestreo no aleatorio

Una muestra no es aleatoria cuando los elementos son elegidos por medio de métodos no aleatorios. Este procedimiento está sujeto a errores, ya que se confía al juicio subjetivo de seres humanos, como en el caso de selección dirigida (no aleatoria), donde las unidades son tomadas en forma caprichosa y en la mayoría de las veces por conveniencia; tal sucede en las encuestas de opinión que realizan los locutores deportivos, cuando se juega un partido de fútbol o en las entrevistas para televisión, radio o prensa, hechas en las principales vías de la ciudad, con el ánimo de conocer la opinión sobre algún asunto de interés para la comunidad.

7.1.1.2. Muestreo aleatorio

Una muestra es aleatoria cuando los elementos que constituyen la población o universo investigado tienen la misma posibilidad de ser seleccionados.

Existen numerosos procedimientos en la aplicación del muestreo aleatorio. Los más importantes son:

7.1.1.2.1. Muestreo aleatorio simple

El método de muestreo aleatorio simple es recomendable, en especial, cuando la población no es numerosa y las unidades se concentran en un área pequeña; por otra parte, la característica no debe tener gran variabilidad, porque implicaría un tamaño muestral muy amplio lo que, a su vez, incrementaría costos y tiempo; por último, la población debe facilitar su enumeración para que permita la aplicación de este método.

7.1.1.2.2. Muestreo aleatorio estratificado

Se le denomina también como muestreo aleatorio restringido. Este procedimiento implica una división de la población en grupos, denominados estratos, en tal forma que el elemento presenta una característica tan definida que sólo le permite pertenecer a un único estrato. Por lo tanto, para que la división por grupos sea efectiva, los factores de estratificación deberán guardar estrecha relación con las características que se investigan y con el objetivo o finalidad del estudio. De esta manera se logra una mayor precisión en los resultados.

7.1.2. Análisis de Regresión Múltiple

Cuando se intenta caracterizar un modelo matemático, una variable independiente a menudo no es suficiente. Por ejemplo, el grado de desgaste de un cojinete lubricado en una máquina puede depender tanto de la carga del cojinete como de las propiedades físicas del lubricante. Las ecuaciones de

desgaste basadas únicamente en la carga o las propiedades del lubricante son predictores fallidos. En tal situación se tienen varias variables independientes x_1, x_2, \dots, x_p que están asociados con la variable dependiente y . Las técnicas de regresión múltiple se utilizan cuando la relación entre las variables dependientes e independientes es lineal.

Para explicar el modelo de regresión múltiple, se tiene una muestra de n artículos, y para cada artículo mide la variable dependiente y y p variables independientes x_1, \dots, x_p . Por lo tanto, el i -ésimo elemento muestral es el conjunto ordenado $(y_i, x_{1i}, \dots, x_{pi})$. Por lo tanto, el modelo se representa como:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i$$

7.1.2.1. Estimación de los coeficientes

Para el cálculo de los estimadores $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$, se calculan por el método de mínimos cuadrados:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_p x_p$$

Para los residuos tenemos:

$$e_i = y_i - \hat{y}_i$$

Se pretende minimizar la suma de los cuadrados residuales por lo que e_i se expresa como:

$$e_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \dots - \hat{\beta}_p x_{pi}$$

Minimizando lo siguiente:

$$\sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \dots - \hat{\beta}_p x_{pi})^2$$

Por lo tanto, el modelo de regresión múltiple, cumpliendo los siguientes supuestos:

1. Los errores $\varepsilon_1, \dots, \varepsilon_n$ son aleatorios e independientes. En particular, la magnitud de cualquier error ε_i no influye en el valor del siguiente error ε_{i+1} .
2. Los errores $\varepsilon_1, \dots, \varepsilon_n$ tienen media 0.
3. Los errores $\varepsilon_1, \dots, \varepsilon_n$ tienen la misma varianza, que se denota por medio de σ^2 .
4. Los errores $\varepsilon_1, \dots, \varepsilon_n$ están distribuidos normalmente.

Implica que las observaciones y_1, \dots, y_n sean variables aleatorias independientes que tienen un comportamiento normal. Donde la media y varianza se definen como:

$$\mu_{y_i} = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi}$$

$$\sigma_{y_i}^2 = \sigma^2$$

A la vez, cada coeficiente β_i representa el cambio en la media de y relacionada al aumento de una unidad en el valor de x_i , al permanecer el resto de variables x constantes.

Figura 1. **Ejemplo de un modelo de regresión lineal múltiple en R**

```
modelo <- lm(esp_vida ~ habitantes + ingresos + analfabetismo + asesinatos +  
            universitarios + heladas + area + densidad_pobl, data = datos )  
summary(modelo)
```

```
##  
## Call:  
## lm(formula = esp_vida ~ habitantes + ingresos + analfabetismo +  
##     asesinatos + universitarios + heladas + area + densidad_pobl,  
##     data = datos)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -1.47514 -0.45887 -0.06352  0.59362  1.21823   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept)  6.995e+01  1.843e+00  37.956 < 2e-16 ***  
## habitantes   6.480e-05  3.001e-05   2.159  0.0367 *   
## ingresos     2.701e-04  3.087e-04   0.875  0.3867   
## analfabetismo 3.029e-01  4.024e-01   0.753  0.4559   
## asesinatos  -3.286e-01  4.941e-02  -6.652 5.12e-08 ***  
## universitarios 4.291e-02  2.332e-02   1.840  0.0730 .   
## heladas      -4.580e-03  3.189e-03  -1.436  0.1585   
## area         -1.558e-06  1.914e-06  -0.814  0.4205   
## densidad_pobl -1.105e-03  7.312e-04  -1.511  0.1385   
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 0.7337 on 41 degrees of freedom  
## Multiple R-squared:  0.7501, Adjusted R-squared:  0.7013   
## F-statistic: 15.38 on 8 and 41 DF,  p-value: 3.787e-10
```

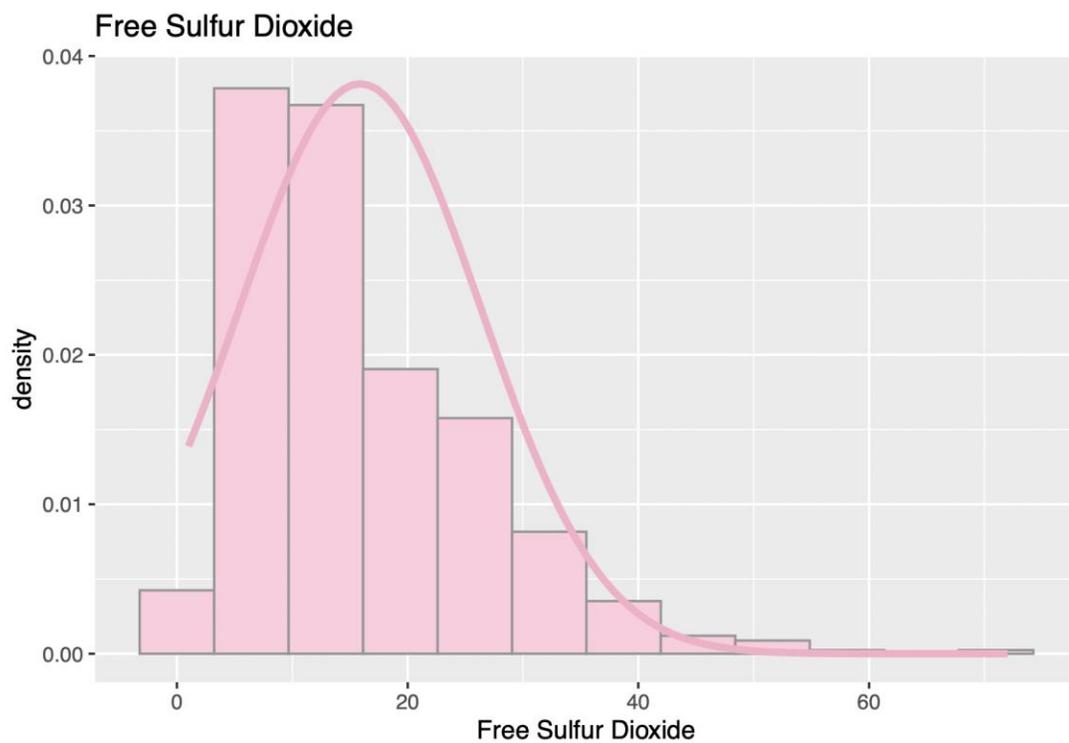
Fuente: Rodrigo, J. (2016). *Introducción a la Regresión Lineal Múltiple* [Imagen]. RPubS by RStudio. Recuperado de: https://rpubs.com/Joaquin_AR/226291

7.1.3. Métodos Estadísticos Gráficos

- Histograma de Densidad

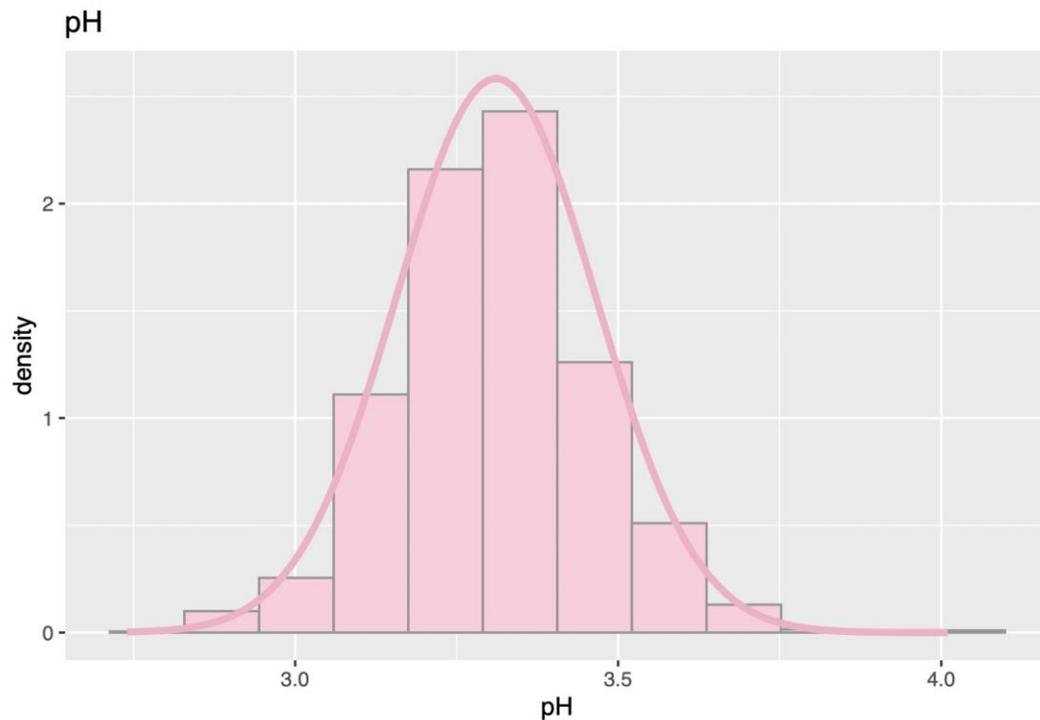
Se entiende por histograma a la representación gráfica en forma de barras, donde cada barra es proporcional a la frecuencia de los datos representados. Nos da una idea sobre la distribución de los datos, rápidamente podemos inferir si estamos ante una serie de datos normales.

Figura 2. **Histograma con datos no normales**



Fuente: elaboración propia, realizado con RStudio, 2022.

Figura 3. **Histograma con datos normales**

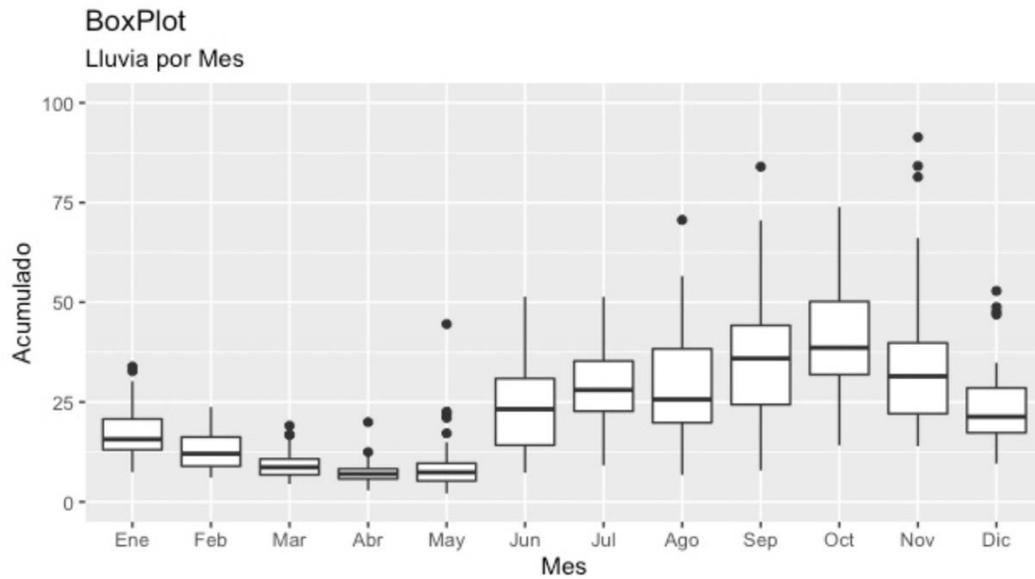


Fuente: elaboración propia, realizado con RStudio, 2022.

- **Box-Plot** o Diagrama de Cajas

El diagrama de cajas es una representación empleada en la estadística con el fin de observar la tendencia de su distribución, es decir, que tanto los datos están centrados respecto a su media y si se tienen puntos atípicos. La importancia radica en que con una inspección visual se pueden realizar inferencias sobre los datos a estudiar.

Figura 4. Representación de Box-Plot

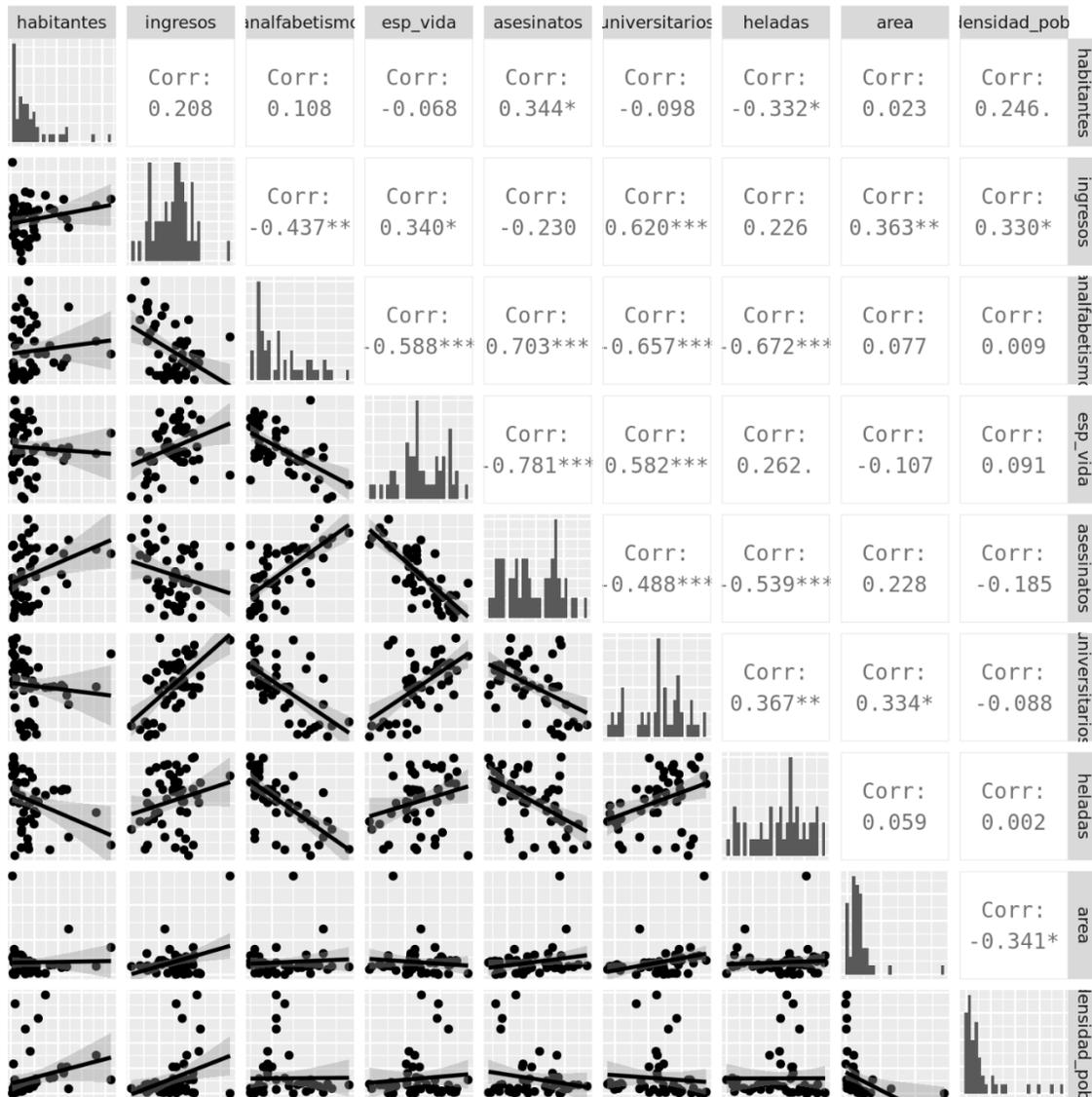


Fuente: elaboración propia, realizado con RStudio, 2022.

- Diagramas de dispersión

Estos diagramas permiten observar la relación que existe entre cada par de variables, así como su distribución y dispersión.

Figura 5. Representación gráfica de un diagrama de dispersión



Fuente: Rodrigo, J. (2016). *Introducción a la Regresión Lineal Múltiple* [Imagen]. RPubs by RStudio. Recuperado de: https://rpubs.com/Joaquin_AR/226291

7.2. El agua

Según Forde, Izurieta, Ôrmeci, Arellano y Mitchell (2019):

La relación entre el agua y la salud humana es inmutable. El agua es imprescindible no sólo para la vida, sino para las buenas condiciones de salud y bienestar humanos. Dado que más de 60 % del cuerpo humano es agua, poder contar con ella en cantidades y calidad suficientes es fundamental tanto para nuestra salud como para nuestro bienestar. (p. 29)

7.2.1. Usos del agua

El agua puede ser empleada en distintos usos, es decir, para uso doméstico o industrial. Y dependiendo de ello serán sus propiedades, porque no se requiere de la misma calidad. Entre estos podemos mencionar:

- Bañarse
- Riego industrial
- Cocinar
- Beber
- Lavar ropa
- Producir energía hidráulica
- Transferencia de calor

En la siguiente tabla observamos un resumen general del uso del agua en el mundo.

Tabla I. **Uso del agua en el mundo**

USO DEL AGUA EN EL MUNDO							
Año	Consumo total	Agricultura		Industria		Agua potable	
	km³/año	km³/año	%	km³/año	%	km³/año	%
1900	400	350	87.5	30	7.5	20	5
1950	1100	820	74.5	220	20	60	5.5
1975	3000	2200	73	650	22	150	5
2000	5000	3400	68	1250	25	350	7

Fuente: Auge, 2007, p.6.

7.2.2. Características físicas

Son las características que se detectan por medio de los sentidos, es decir, vista, olfato, sabor o bien por mediciones analíticas.

7.2.3. Características químicas

Son las características que dependen de elementos orgánicos e inorgánicos.

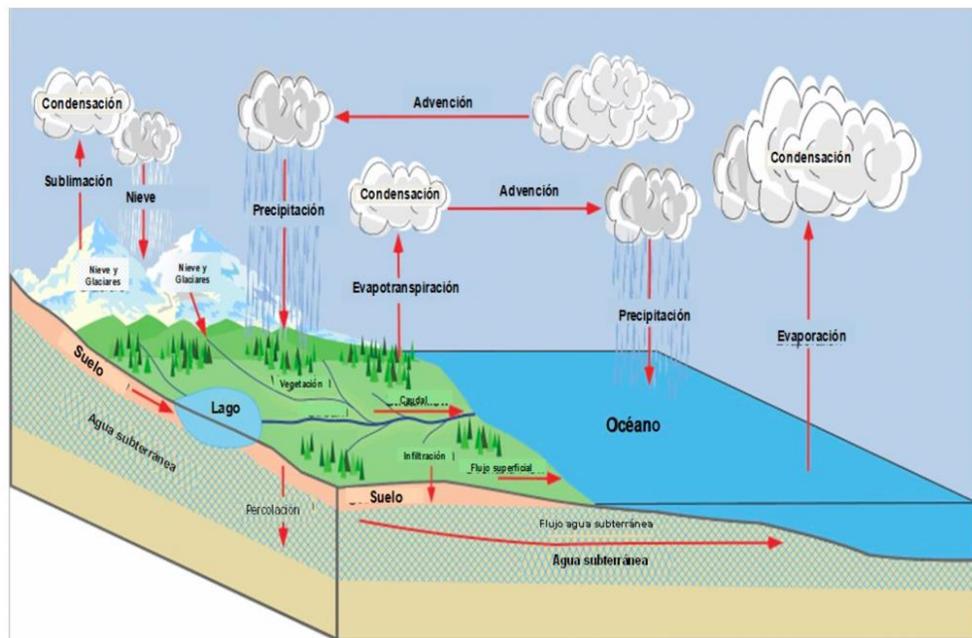
7.2.4. Características microbiológicas

Son las características que se derivan por trazas de microorganismos.

7.2.5. Tipo de agua

Acorde a sus características químicas, físicas y biológicas el agua se puede clasificar entre agua cruda, residuales y tratada que es el agua potable la cuál es de interés para esta investigación.

Figura 6. **Ciclo hidrológico**



Fuente: Ordoñez, 2011, p. 9.

7.2.5.1. Agua Cruda

Se comprende al agua cruda como al agua que se encuentra en el medio ambiente, por mencionar ejemplos, lagos, ríos, océanos, agua subterránea, lluvia. Dado que no ha recibido ningún tratamiento previo es que se le conoce también como agua natural.

El agua natural es la fuente de abastecimiento de muchas comunidades, donde la calidad puede depender de su origen, es decir, si proviene del mar, si es subterránea o superficial. O bien, de la actividad humana, ya que muchas veces el hecho de no tener drenajes hace que industrias a orillas de ríos o lagos desemboquen sus desechos en esa área. También la explotación minera y agroquímicos afectan considerablemente la calidad del agua.

A continuación, se mostrará la cantidad promedio de minerales y sales con mayor concentración en el agua de mar que es muy similar a la que se encuentra en los océanos.

Tabla II. **Calidad promedio del agua de mar**

Parámetro	ppm CaCO₃
Calcio	1,000
Magnesio	5,234
Sodio	23,000
Potasio	500
Bicarbonato	120
Cloruros	27,000
Sulfatos	2,800
Bromuros	40
Dureza	6,200

Fuente: Sierra, 2011, p.50.

7.2.5.2. Agua Residual

Es el agua que ya ha sido previamente utilizada para cualquier proceso. Entre ellas está el agua residual industrial, comercial y doméstica.

7.2.5.3. Agua Potable

Es el agua que ha sido previamente tratada para modificar sus características físicas, químicas y biológicas con el fin de ser utilizada posteriormente. Cabe resaltar que su calidad dependerá mucho del fin con el que se está tratando. Ya que si es agua apta para consumo humano debe cumplir con distintas características que si es para riego de un beneficio de café.

7.2.6. Métodos de Filtrado

Existen diversos métodos de filtrado para obtener agua apta para consumo humano, con el fin de que no sea nociva para el ser humano al ingerirla.

Entre los filtros más conocidos podemos encontrar:

- Carbón activado: su principal fin es el poder eliminar malos olores del agua y químicos que puedan dañar la salud de las personas.
- Intercambio iónico: es de utilidad para disminuir la concentración de sales, nitratos del agua y descalcificar.
- Filtros cerámicos: eliminan microorganismos y partículas en suspensión por medio del material poroso del filtro.
- Luz ultravioleta: logran eliminar las bacterias que están presentes en el agua.

- Osmosis inversa: estos filtros emplean membranas que tienen microporos con el fin de captar sólidos disueltos en el agua y eliminar dichas impurezas por el desagüe.

7.2.7. Calidad del Agua

Es importante garantizar agua libre de contaminantes para el ser humano, porque de lo contrario puede ser muy perjudicial para la salud. Tomando en cuenta que el agua no es solo de utilidad para consumo sino actividades cotidianas como el cocinar. Entre los efectos nocivos se pueden mencionar, enfermedades del aparato digestivo y cáncer si contiene metales pesados.

Según, Sierra (2011):

En vista de la complejidad de los factores que determinan la calidad del agua y la gran cantidad de variables utilizadas para describir el estado de los cuerpos de agua en términos cuantitativos, es difícil dar una definición simple de calidad del agua. Además, los conocimientos sobre calidad del agua han evolucionado a través del tiempo a medida que ha aumentado su demanda en diferentes usos y han mejorado los métodos para analizar e interpretar sus características. (p. 47)

Por ende, podemos inferir que la calidad del agua es cuando la composición y el estado de la biota acuática son las debidas. Y se ve alterada cuando factores externos o internos modifican su estado.

Esta puede medirse de tres formas:

- Variables físicas:
 - Sólidos totales
 - En suspensión.
 - Disueltos.
 - Turbidez
 - Color
 - Olor
 - Sabor

- Variables químicas:
 - pH
 - Acidez
 - Conductividad
 - Metales
 - Cloro
 - Detergentes
 - Pesticidas

- Variables biológicas
 - Hongos
 - Bacterias coliformes

7.2.8. Norma COGUANOR

Según la Comisión Guatemalteca de Normas Ministerio de Economía,
(2013):

La Comisión Guatemalteca de Normas (COGUANOR) es el Organismo Nacional de Normalización, creada por el Decreto No. 1523 del Congreso de la República del 05 de mayo de 1962. Sus funciones están definidas en el marco de la Ley del Sistema Nacional de la Calidad, Decreto 78-2005 del Congreso de la República. (p. 3)

COGUANOR tiene como fin principal el apoyar a entidades públicas y privadas en estandarizar sus procesos. Por lo tanto, brindan asesorías técnicas.

La norma que proporciona los estándares para afirmar que el agua es apta para consumo humano, es decir, agua potable, es la norma COGUANOR NTG 29001.

Su campo de aplicación es para toda agua para consumo humano, preparación de alimentos y uso doméstico. Dicha norma no aplica para el agua purificada envasada y carbonatada porque para ellas existen otras normativas.

Según COGUANOR la definición de Agua Potable “es aquella que por sus características organolépticas, físicas, químicas y bacteriológicas, no representa un riesgo para la salud del consumidor y cumple con lo establecido en la presente norma” (Comisión Guatemalteca de Normas Ministerio de Economía, 2013, p. 4).

Las características que deben tener el agua para consumo humano se muestran en las siguientes tablas:

Tabla III. **Relación de las sustancias inorgánicas cuya presencia en el agua es significativa para la salud**

Substancia	LMP (mg/L)
Arsénico (As)	0.010
Bario (Ba)	0.700
Boro (B)	0.300
Cadmio (Cd)	0.003
Cianuro (CN ⁻)	0.070
Cromo total (Cr)	0.050
Fluoruro (F ⁻)	1.500
Mercurio total (Hg)	0.001
Plomo (Pb)	0.010
Selenio (Se)	0.010
Nitrato (NO ₃ ⁻)	50.000
Nitrito (NO ₂ ⁻)	3.000

Fuente: COGUANOR, 2013, p.7.

8. PROPUESTA DE ÍNDICE DE CONTENIDOS

ÍNDICE DE ILUSTRACIONES

ÍNDICE DE TABLAS

LISTA DE SÍMBOLOS

GLOSARIO

RESUMEN

PLANTEAMIENTO DEL PROBLEMA

OBJETIVOS

MARCO METODOLÓGICO

INTRODUCCIÓN

1. MARCO DE REFERENCIA

2. MARCO TEÓRICO

2.1. Estadística

2.1.1. Muestreo

2.1.1.1. Muestreo no aleatorio

2.1.1.2. Muestreo aleatorio

2.1.1.2.1. Muestreo aleatorio simple

2.1.1.2.2. Muestreo aleatorio
estratificado

2.1.2. Análisis de Regresión Múltiple

2.2. El Agua

2.2.1. Usos del agua

2.2.2. Características físicas

2.2.3. Características químicas

- 2.2.4. Características microbiológicas
- 2.2.5. Tipo de agua
 - 2.2.5.1. Agua Cruda
 - 2.2.5.2. Agua Residual
 - 2.2.5.3. Agua Potable
- 2.2.6. Métodos de filtrado
- 2.2.7. Calidad del agua
- 2.2.8. Norma COGUANOR

3. PRESENTACIÓN DE RESULTADOS

4. DISCUSIÓN DE RESULTADOS

CONCLUSIONES

RECOMENDACIONES

REFERENCIAS Y BIBLIOGRAFÍA

ANEXOS

9. METODOLOGÍA

9.1. Características del estudio

El enfoque del estudio propuesto es cuantitativo, ya que se realizará un análisis matemático y estadístico con datos históricos de la planta de tratamiento de agua y residuos sólidos.

El alcance es correlacional, dado que se busca conocer la relación e influencia que existe entre las variables del proceso para inferir el costo del proceso y medir

El diseño adoptado será observacional y retrospectivo, pues la información a utilizarse son datos históricos de la planta, que se analizarán en su estado original sin ninguna manipulación.

9.2. Unidad de análisis

La población en estudio estará constituida por los procesos de aguas residuales y desechos sólidos del Departamento de Guatemala, la cual se encuentra dividida en subpoblaciones dadas por las diferentes industrias de las cuáles procederá el agua. Partiendo de esto se construirá el análisis de regresión para explicar el costo.

9.3. Operativización de las variables

Tabla IV. Operativización de Variables

Variable	Definición teórica	Definición operativa
Costo (c)	Cantidad en dinero que cuesta realizar el proceso.	Variable dependiente para explicar con el modelo de regresión (Variable cuantitativa y continua).
Peso de los desechos (w)	Cantidad de aguas residuales y desechos sólidos a tratarse en kilogramos (Variable independiente).	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).
Cantidad de materia orgánica presente DBO (dbo)	Cantidad de oxígeno para descomponer biológicamente la muestra.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).
Cantidad de materia orgánica presente DQO (dqo)	Cantidad de oxígeno para descomponer químicamente la materia.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).

Continuación de tabla IV.

Aire (a)	Cantidad de aire a inyectar en el sistema de aireación en metros cúbicos.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).
Energía eléctrica (Kwh) (E)	Energía consumida en Kilovatios hora.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).
Desinfección (d)	Unidades de tricloro a usar en el proceso de desinfección.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).
Tiempo en el biodigestor en días (t)	Los días que debe pasar el agua residual y desechos sólidos en el biodigestor.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).

Continuación de tabla IV.

Distancia (l)	Distancia de transporte en kilómetros.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).
Pago de trabajadores (m)	Salario devengado por los trabajadores en quetzales por día.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).
Combustible (g)	Consumo de combustible en galones y su precio en quetzales.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).
Mantenimiento de vehículos (s)	Servicio de mantenimiento a los vehículos utilizados en quetzales por kilómetro recorrido.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).

Continuación de tabla IV.

Mantenimiento de maquinaria (p)	Costo de realizar el mantenimiento de los equipos de la planta en Quetzales por hora.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).
Misceláneos (v)	Utensilios, uniformes, en quetzales.	Variable independiente, que conforma la base de datos histórica del proceso de tratamiento de aguas residuales y desechos sólidos. (Variable cuantitativa y continua).

Fuente: elaboración propia.

9.4. Fases del estudio

El desarrollo de la investigación se compone de distintas fases, las cuáles se presentan a continuación:

- Fase 1: Revisión de literatura

En esta fase se revisará literatura que aporte valor a la investigación y que sustente lo que se realizará. Para ello se buscarán temas sobre la importancia y calidad del agua, así como temas de análisis estadístico para la interpretación de modelos de regresión múltiple.

- Fase 2: Gestión o recolección de la información

Para esta fase, se recopila la información histórica de la planta de tratamiento de agua para su posterior análisis.

- Fase 3: Análisis de información

Posterior a la recolección de la información, se procede a realizar el análisis de correlación para identificar las variables con mayor influencia en el costo. Dicho modelo debe cumplir con los supuestos de los residuos (normalidad, homocedasticidad, independencia). Luego, con un análisis de regresión serán detectadas las diferentes interacciones entre variables y será posible discriminar las que pueden ser eliminadas para maximizar la parsimonia.

Con esas variables serán construidos diferentes modelos, variando las formas matemáticas a través de transformaciones por anamorfosis. Los estadísticos del ajuste Coeficiente de determinación, criterios de información permitirán priorizar las opciones más consistentes de modelo.

- Fase 4: Interpretación de información

Acorde a los resultados se podrá definir si el modelo matemático es el óptimo para explicar el costo del proceso.

- Fase 5: Redacción del informe final

Con base a la literatura y los resultados, se procede a plasmar dicha información en un informe final que sea entendible y contribuya al conocimiento

ya que, no se cuenta con una estandarización al momento de asignar el costo al proceso de tratar aguas residuales.

10. TÉCNICAS DE ANÁLISIS DE INFORMACIÓN

En este apartado se pretende describir las técnicas que se emplearán para el desarrollo de la investigación:

- Análisis de correlación

El coeficiente de Pearson permitirá inferir si hay alguna relación estadística entre cada variable independiente y el costo del proceso.

- Diagrama de dispersión

Este diagrama permitirá cruzar todas las variables para detectar colinealidad o influencia y vislumbrar las funciones matemáticas que probablemente describen esas relaciones.

- Análisis de regresión

Este análisis se realizará con el fin de conocer que variables inciden en explicar el costo final del tratamiento de aguas residuales, para este estudio saber si el peso del agua y desechos sólidos, el mantenimiento de vehículos y equipo, combustible, salario de los trabajadores, por mencionar algunas variables, son incidentes en explicar el costo de tratar el agua y sus residuos para obtener agua potable. Además, será posible detectar relaciones entre variables que permitan reducir la complejidad del modelo, con pruebas de multicolinealidad.

- Análisis de residuos

Al tener las opciones más plausibles para el modelo, serán evaluados los supuestos que deben cumplir sus residuos para inferir si su ajuste es congruente. Debe evaluarse si estos se ajustan a una distribución normal, si su varianza es homogénea y si son independientes entre sí, de esa forma se descartan patrones no descritos por el modelo y cualquier inconformidad estadística. El modelo que se acerque con mayor probabilidad a cumplir con todos los requisitos será seleccionado como el mejor modelo de predicción.

11. CRONOGRAMA

Figura 7. Cronograma

FASES	TIEMPO (Semanas)																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Semana																				
Aprobación de Protocolo de Tesis																				
Desarrollo de Trabajo de Investigación																				
FASE 1: Revisión literaria																				
1.1. Revisión de antecedentes																				
1.2. Revisión de marco teórico																				
FASE 2: Recopilación de la información																				
2.1. Revisar la información histórica																				
2.2. Ordenar la información.																				
FASE 3: Análisis de la información																				
3.1. Realizar el modelo de regresión lineal																				
FASE 4: Interpretación de la información																				
4.1. Interpretación de resultados																				
FASE 5: Informe Final																				
5.2. Discusión de resultados																				
5.3. Elaboración de conclusiones																				
5.4. Elaboración de recomendaciones																				
5.5. Elaboración de anexos																				
TIEMPO TOTAL:																				

20 semanas

Fuente: elaboración propia, realizado con Microsoft Excel.

12. FACTIBILIDAD DEL ESTUDIO

El estudio a realizarse, si es factible porque se cuenta con el recurso humano, financiero y tecnológico para llevarse a cabo.

12.1. Recurso humano

El grupo de personas involucradas en el desarrollo de la investigación. Que contempla al asesor e investigador.

12.2. Recursos tecnológicos

En este apartado se consideran los programas informáticos que se emplearán en la elaboración de gráficos y análisis estadístico de la investigación, entre ellos, R Studio, Microsoft Excel e Infostat.

12.3. Recursos financieros

El estudio será financiado por el investigador, es decir, quién incurrirá con los gastos que sean necesarios para completar el proyecto.

Tabla V. **Presupuesto**

Recurso	Descripción	Cantidad	Costo Unitario	Costo Total
Humano	Inversión de tiempo del investigador por hora	360	Q 125.00	Q 45,000.00
Humano	Inversión de tiempo del asesor por hora	360	Q 0.00	Q 0.00
Transporte	Traslado para las reuniones con el asesor por kilómetro recorrido	300	Q 6.00	Q 1,800.00
Material	Útiles de oficina	1	Q 1,000.00	Q 1,000.00
Material	Internet por mes	6	Q 375.00	Q 2,250.00
Material	Electricidad por mes	6	Q 100.00	Q 600.00
Tecnológico	Equipo de cómputo	1	Q 33,000.00	Q 33,000.00
Otros	Imprevistos	1	Q 5,000.00	Q 5,000.00
TOTAL				Q 88,650.00

Fuente: elaboración propia.

12.4. Acceso a información y permisos

La información será proporcionada por la Empresa de Tratamiento de Aguas Residuales y Desechos Sólidos.

12.5. Equipo e infraestructura

Es el equipo que se emplea en la investigación, como el equipo de cómputo que está integrado por una computadora portátil y un monitor, teléfono celular y programas informáticos.

13. REFERENCIAS

1. Auge, M. (2007). *AGUA FUENTE DE VIDA*. Argentina: Universidad de Buenos Aires.
2. Camacho, N. C. C. (2011). *Tratamiento de agua para consumo humano*. Perú: Universidad Nacional Mayor de San Marcos.
3. Clara, M. R. M. (2005). *Análisis de la calidad del agua para consumo humano y percepción local de las tecnologías apropiadas para su desinfección a escala domiciliaria, en la microcuenca El Limón, San Jerónimo, Honduras*. Centro Agronómico Tropical de Investigación y Enseñanza , Costa Rica.
4. Copano, M. A. Z. (2018). *ANÁLISIS Y EVALUACIÓN DE LA CALIDAD DEL AGUA POTABLE PARA LA CIUDAD DE ANTOFAGASTA BAJO EL CONTEXTO DEL SUMINISTRO DE AGUA DESALADA*. Universidad de Chile, Chile.
5. De Los Ángeles Aguilar Moncayo, M. (s/f). *Análisis ecológico de la calidad y cantidad del recurso hídrico para consumo humano en la parroquia San Isidro de Patulú*. 2017, Universidad Técnica de Ambato.
6. Gálvez, J. J. O. (2011). *Contribuyendo al desarrollo de una Cultura del Agua y la Gestión Integral de Recurso Hídrico*. Perú: Sociedad Geográfica de Lima.

7. Garay, M. M. B. (2016). *TRATAMIENTO DE AGUAS RESIDUALES*. Guayaquil, Ecuador.
8. González, M. Á. M., Sánchez-Villegas, A., Toledo Atucha, E. A., & Fajardo, J. F. (2020). *BIOESTADÍSTICA AMIGABLE*. Elsevier.
9. Martínez Bencardino, C. (2012). *Estadística y muestreo*. Colombia: ECOE EDICIONES.
10. Milton, J. (2007). *Estadística para biología y ciencias de la salud*. Madrid, España: McGRAW-HILL.
11. Montgomery, D. C. (2004). *Diseño y análisis de experimentos*. México D.F.: LIMUSA, S.A.
12. Pérez, C. (2010). *Técnicas de muestreo estadístico*. 2010: IBERGARCETA PUBLICACIONES, S.L.
13. Quiroz, S., Menéndez, C., & Izquierdo, E. (2019). *Tratamiento de agua y aguas residuales*. Ecuador: Ediciones UTM.
14. Ramírez, C. A. S. (2011). *Calidad del agua*. Medellín, Colombia. Ediciones de la U.
15. Sáenz, A. M. G. (2018). *Determinación estadística de la influencia de los parámetros del control de calidad del agua potable de los municipios de Floridablanca, Girón y Bucaramanga*. Universidad De Santander , Bucaramanga.

16. Saldaña, J. P. C., & de los Godos Urcia, L. A. (2011). *ESTADÍSTICA PARA LA TESIS DE POSTGRADO*. Perú: LULU International.
17. Santisteban, M. L. C., & Peña, W. (2015). *Evaluación de la calidad del agua superficial con potencial para consumo humano en la cuenca alta del Sis Icán, Guatemala*. Universidad de San Carlos de Guatemala, Guatemala.
18. Santizo, M. I. D. (2015). *ANÁLISIS DE LA CALIDAD DE AGUA EN LA CUENCA MEDIA DEL RÍO MOTAGUA*. UNIVERSIDAD RAFAEL LANDÍVAR, Guatemala.
19. Solares, P. C. A. (2016). *DETERMINACIÓN DE LOS ÍNDICES DE CALIDAD DEL AGUA (ICA-NSF E ISCA) PARA CONSUMO HUMANO DE LOS RÍOS TEOCINTE Y ACATÁN, QUE ABASTECEN LA PLANTA DE TRATAMIENTO DE AGUA SANTA LUISA ZONA 16, GUATEMALA*. Universidad de San Carlos de Guatemala, Guatemala.
20. Solorzano, S. H. (2018). *ANÁLISIS DE LA PERCEPCIÓN EN LA CONTAMINACIÓN DE ARROYOS URBANOS EN LA MICROCUENCA EL RIÍTO EN TONALÁ CHIAPAS, MÉXICO*. El Colegio de la Frontera Norte, México.
21. Sotomayor, G. (2016). *Evaluación de la calidad de las aguas superficiales mediante técnicas de estadística multivariante: Un estudio de caso en la cuenca del Río Paute, al sur de Ecuador*. Universidad Nacional de La Plata, Argentina.

22. Triola, M. (2018). *Estadística*. México, Ciudad de México. PEARSON.
23. Valdes-Basto, J., Samboni-Ruiz, N. E., & Carvajal-Escobar, Y. (Eds.). (2011). *Desarrollo de un Indicador de la Calidad del Agua usando Estadística Aplicada, Caso de Estudio: Subcuenca Zanjón Oscuro*. Rev. Tecno Lógicas No. 26, ISSN 0123-7799.
24. Wackerly, D. (2010). *Estadística matemática con aplicaciones*. México D.F., México. Cengage Learning.
25. Walpole, Myers, & Myers. (2012). *Probabilidad y estadística para ingeniería y ciencias*. México: PEARSON.

14. APÉNDICE

Apéndice 1. **Matriz de Coherencia**

ELEMENTOS	PROBLEMA DE INVESTIGACIÓN	PREGUNTAS DE INVESTIGACIÓN	OBJETIVOS	METODOLOGÍA
GENERAL	Se desconoce la varianza entre los valores para la calidad del agua respecto a los parámetros de la norma COGUANOR de agua para consumo humano.	¿Cuál es la diferencia entre los valores para la calidad del agua de los diferentes métodos de filtrado respecto a los parámetros según la norma COGUANOR NGO 29 001:2010 de agua para consumo humano?	Estimar la varianza entre los parámetros de la calidad del agua respecto a los límites establecidos por la norma COGUANOR NGO 29 001:2010 de agua para consumo humano utilizando un análisis de varianza para comparar diferentes métodos de filtrado.	El enfoque del estudio propuesto es cuantitativo ya que se estimará la varianza de los parámetros de la calidad del agua respecto a lo que dita la norma COGUANOR NG 29001.
	No se conoce cuál es el mejor método de filtrado de agua para consumo humano.	¿Cuál es el mejor método de filtrado para agua para consumo humano en casa?	Comparar la presencia de E. coli en los tres distintos métodos de filtrado a través de una prueba MCB de Hsu para clasificarla como agua apta para consumo humano.	El alcance es descriptivo transversal, porque se estudiarán las características biológicas y físicas del agua con el fin de determinar si es apta para consumo humano y serán medidas en una sola ocasión, siendo muestras independientes.
ESPECÍFICOS	No se sabe que tanto se desvía la calidad del agua filtrada respecto a la norma COGUANOR.	¿Qué tanto se desvía la calidad del agua filtrada respecto a lo establecido en la norma COGUANOR NGO 29 001:2010?	Contrastar la presencia de coliformes totales en los tres distintos métodos de filtrado con el límite normado a través de pruebas de hipótesis de dos colas para catalogarla como agua apta para consumo humano.	El diseño adoptado será experimental (observacional), pues la información para determinar las características del agua será tomada a partir de las muestras de agua filtradas que serán tomadas en los municipios del departamento de Guatemala.
	No se tiene conocimiento sobre cuál es la mejor elección de agua para consumo humano.	¿Cuál es la mejor elección de agua para consumo humano?	Comparar la concentración de plomo en los tres distintos métodos de filtrado a través de un análisis de varianza para conocer si su valor es significativamente dañino para la salud.	También será un experimento de campo, porque las muestras serán tomadas del grifo de distintas casas de municipios de la Ciudad de Guatemala.

Fuente: elaboración propia.

