

# **Métodos cuantitativos aplicados a recursos naturales, medioambiente y la sostenibilidad**

**Jesús Ramón Rodríguez Apodaca / Claudia Selene  
Castro Estrada / Alma Lorena Quintero Romanillo**  
*Coordinadores*





# Métodos cuantitativos aplicados a recursos naturales, medioambiente y la sostenibilidad





# Métodos cuantitativos aplicados a recursos naturales, medioambiente y la sostenibilidad

Jesús Ramón Rodríguez Apodaca  
Claudia Selene Castro Estrada  
Alma Lorena Quintero Romanillo  
*Coordinadores*



*Métodos cuantitativos aplicados a recursos naturales, medioambiente y la sostenibilidad.* **Autores-Coordinadores:** Dr. Jesús Ramón Rodríguez Apodaca, Dra. Claudia Selene Castro Estrada y Dra. Alma Lorena Quintero Romanillo. —Universidad Autónoma Indígena de México, Sinaloa, México. 2024.

212 P. 23 cm.

*Primera edición*

D. R. © copyright 2024

ISBN: **978-84-10215-06-1**

DOI: <https://doi.org/10.61728/AE24250009>

La presente obra fue dictaminada bajo el sistema de doble ciego y cuenta con el aval de los dictámenes de pares académicos en el campo de las ciencias socioambientales en México.

Edición y corrección: **Astra ediciones**

Se prohíbe la reproducción, el registro o la transmisión parcial o total de esta obra por cualquier sistema de recuperación de información, sea mecánico, fotoquímico, electrónico, por fotocopia, cualquier otro existente o por existir; sin el permiso previo, por escrito, del titular de los derechos.

## Contenido

|                                                                                                                                                                    |            |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Prólogo .....</b>                                                                                                                                               | <b>9</b>   |
| <i>Jesús Manuel Díaz Gaxiola</i>                                                                                                                                   |            |
| <br><b>Capítulo 1</b>                                                                                                                                              |            |
| <b>Aplicaciones del muestreo simple al azar para proporciones .....</b>                                                                                            | <b>11</b>  |
| <i>Elvia Nereyda Rodríguez Saucedo</i>                                                                                                                             |            |
| <i>Raquel Rodríguez Saucedo</i>                                                                                                                                    |            |
| <i>Adán Guillermo Ramírez García</i>                                                                                                                               |            |
| <br><b>Capítulo 2</b>                                                                                                                                              |            |
| <b>Análisis univariado y multivariado para evaluación de fitoplancton y calidad de agua en estanques acuícolas en el municipio de Ahome, Sinaloa, México .....</b> | <b>37</b>  |
| <i>Jesús Manuel Díaz Gaxiola</i>                                                                                                                                   |            |
| <i>Jesús Octavio Escobedo Urquidez</i>                                                                                                                             |            |
| <br><b>Capítulo 3</b>                                                                                                                                              |            |
| <b>Análisis clúster o conglomerados: identificación de dos temporalidades ambientales en la zona marino-costera de Guasave.....</b>                                | <b>61</b>  |
| <i>Graciano-Obeso Adalid</i>                                                                                                                                       |            |
| <i>Alzate-Espinoza Juan Héctor</i>                                                                                                                                 |            |
| <i>Bojórquez-Delgado Gilberto</i>                                                                                                                                  |            |
| <br><b>Capítulo 4</b>                                                                                                                                              |            |
| <b>Regresión lineal y análisis de la serie de tiempo mediante la codificación de datos .....</b>                                                                   | <b>79</b>  |
| <i>Lennin Enrique Amador-Castro</i>                                                                                                                                |            |
| <i>Jesús Ramón Rodríguez-Apodaca</i>                                                                                                                               |            |
| <i>Esther Graciela Lizárraga-Mata</i>                                                                                                                              |            |
| <br><b>Capítulo 5</b>                                                                                                                                              |            |
| <b>Regresión múltiple desde el punto de vista de la gestión-administración con un enfoque sustentable caso ganadería Rubio .....</b>                               | <b>105</b> |
| <i>Juan Héctor Alzate Espinoza</i>                                                                                                                                 |            |
| <i>Adalid Graciano Obeso</i>                                                                                                                                       |            |
| <i>Grace Erandy Báez Hernández</i>                                                                                                                                 |            |

**Capítulo 6****Metodología para simulaciones de liberación de sustancias peligrosas empleando el programa Aloha ..... 125***Marco Arturo Arciniega Galaviz***Capítulo 7****Aplicaciones de redes neuronales recurrentes en el análisis cuantitativo de series temporales para la gestión sostenible de recursos naturales y medioambientales ..... 155***Gilberto Bojórquez-Delgado**Jesús Bojórquez-Delgado**Adalid Graciano-Obeso***Capítulo 8****Cuantificación de recursos naturales pétreos a partir de métodos aerofotogramétricos y análisis en sistemas de información geográfica ..... 179***Román Edén Parra-Galaviz**Manuel de Jesús Pérez-Valdez**Ciro de la Concepción Díaz-Hernández***Síntesis curriculares ..... 205**

## Prólogo

Es un gran honor poder escribir este prólogo para este libro lleno de mucho conocimiento, experiencias y trabajos en el medioambiente. El uso de métodos cuantitativos en la gestión de recursos naturales, medioambiente y sostenibilidad es fundamental para comprender, evaluar y abordar los desafíos que enfrenta nuestro planeta. Estos métodos proporcionan herramientas analíticas y técnicas que permiten a los investigadores y profesionales recolectar, analizar y presentar datos de una manera cuantitativa, lo que a su vez facilita la toma de decisiones informadas.

En el contexto de los recursos naturales, los métodos cuantitativos se utilizan para realizar evaluaciones de inventario, modelado de recursos, estimación de reservas y proyecciones de demanda. En el ámbito del medioambiente, estos métodos son aplicados para el monitoreo de la calidad del aire y del agua, evaluación de impacto ambiental, análisis de riesgos y modelado de ecosistemas. En cuanto a la sostenibilidad, los métodos cuantitativos son empleados para medir la huella de carbono, realizar análisis de ciclo de vida, evaluar la eficiencia energética y cuantificar el impacto de políticas y prácticas sostenibles.

Dada la complejidad y la interconexión de los sistemas naturales, la aplicación de métodos cuantitativos ofrece una ventaja al proporcionar un enfoque riguroso y basado en evidencia para abordar los desafíos ambientales y de sostenibilidad. A través de la integración de datos cuantitativos, modelos matemáticos y herramientas estadísticas, es posible obtener una comprensión más profunda de los problemas ambientales y tomar decisiones más efectivas para la conservación de los recursos naturales y la promoción de la sostenibilidad a largo plazo.

Los autores de estos ocho capítulos del libro, exponen de manera magistral y a través de escenarios reales, los beneficios del análisis cuantitativo ya sea en el mar, la tierra o el aire y de los recursos naturales que estos poseen y sus servicios ambientales que brindan, para realizar aprovechamientos sostenibles y generar conocimientos actuales y de un valor inestimable dado las afectaciones que hemos realizado como humanidad.

---

<sup>1</sup> <https://doi.org/10.61728/AE24250016>

Este libro constituye un excelente manual, que lleva al lector paso a paso en la realización de los análisis cuantitativos diversos y lo que es mejor, lo hacen de una manera amena y aplicada, considerando los aspectos cotidianos en el medioambiente y en los sistemas productivos de nuestra región, para que sigamos obteniendo los beneficios que nos brinda nuestro ecosistema, pero de una manera más inteligente y previendo escenarios futuros, relacionados con la disponibilidad de recursos naturales, los efectos del cambio climático y las consecuencias de las acciones humanas. Felicidades a los coordinadores, autores e instituciones educativas participantes y a todos los lectores que tienen en sus manos a un excelente material de información y formación.

*Dr. Jesús Manuel Díaz Gaxiola*  
Academia de Biología.  
Instituto Tecnológico de Los Mochis.  
Tecnológico Nacional de México.

# Capítulo 1

---

## **Aplicaciones del muestreo simple al azar para proporciones**

*Elvia Nereyda Rodríguez Saucedo*

*Raquel Rodríguez Saucedo*

*Adán Guillermo Ramírez García*

## Introducción

Una de las primeras preguntas que un investigador se plantea en su trabajo es de ¿cuántos individuos necesito para realizar mi estudio estadísticamente significativo? la respuesta es de acuerdo con los objetivos, los alcances, la naturaleza del trabajo de investigación y de los resultados esperados; las preguntas antes planteadas deben ser muy bien planeadas por los ejecutores del proyecto (Ross, 1999; Martínez y Martínez, 2008; Arias, 2000).

Esta pregunta de investigación sobre el número de individuos de la muestra, es muy lógica e indicada. Los objetivos de la investigación en ciencias sociales o en las ciencias ambientales son, por una parte, caracterizar las poblaciones y, por otro, conocer y contribuir al descubrimiento de las leyes que gobiernan o influyen en estas poblaciones (Cochran, 1977; Tortora, 1978; Dawson, 2002). Sea cual sea el objetivo de los estudios de investigación, todas tienen interés de conocer las características de los estimadores que representen a las poblaciones (Barón, 2005).

Debido a que es muy difícil caracterizar a las poblaciones en su totalidad, ya sea por el tiempo que se requiere, los recursos financieros (costos) que se necesitan o el personal involucrado en la toma de las muestras (Goodman, 1965; Thompson, 1978; Ávila, 2006); se requiere trabajar en esas poblaciones con grupos menores de individuos, que se les llama muestras y esto relacionado con la pregunta que se planteó anteriormente. En los trabajos de investigación aplicada en las ciencias sociales y ambientales, por lo general se trata de caracterizar situaciones de fenómenos de grupos de individuos llamado universo o población, y está compuesto por la totalidad de los individuos o unidades a estudiar (Fitzpatrick y Scott, 1978; Bromaghin, 1993; Arias et al., 2016).

Al caracterizar poblaciones por medio de un menor grupo de individuos a una escala menor se le llama “muestra”. El muestreo estadístico nos proporciona métodos y técnicas que tienen el propósito de obtener mediante la muestra representativa una caracterización que nos describa de manera razonable la población de interés (Abger, 1974; Cameron y Trivedi, 2009; Bai et al., 2013).

Con el muestreo estadístico se busca elaborar estrategias de muestreo y de medición que permitan la caracterización y la inferencia (extrapolar) hacia el total de la población (Martínez y Martínez, 2008). Los estudios de investigación que van de lo particular a lo general, pueden tener errores y entre mayor sea

el tamaño de la muestra a medir, es mayor la posibilidad de cometerlos y esto llevará a conclusiones que no representen la realidad de la población sobre la que recaerán las conclusiones generales de la inferencia estadística (Cochran, 1977; Roos, 1999; Goto et al., 2014).

Muchos estudios de investigación en ciencias sociales y del ambiente, se caracterizan por realizar inferencias sobre las poblaciones y esto se hace mediante información levantada mediante encuestas y/o entrevistas a través de la aplicación de un cuestionario o la realización de mediciones de los individuos (Thompson, 1987; Cameron y Trivedi, 2009; Hernández et al., 2006). Estas estrategias corresponden a una aproximación metodológica que permite medir un objeto de estudio como por ejemplo una persona entrevistada (Zermeño, 2011; Hund et al., 2015). Estos individuos ya sean personas entrevistadas o sujetos medidos cualitativa o cuantitativamente, deben representar al conjunto de la población objetivo a caracterizar y el muestreo estadístico permite explorar, describir o explicar las características de la población que dio origen a la muestra y analizar de manera precisa y estadísticamente significativa dicha información (Evans y Rosenthal, 2004).

Por lo anteriormente expuesto, el objetivo de este trabajo es presentar el procesamiento del cálculo del tamaño de muestra para variables cualitativas mediante el muestreo simple al azar para proporciones (Wolpole et al., 1996).

## Metodología

El muestreo estadístico es un área de las matemáticas que tiene técnicas y métodos adecuados para el diseño, cálculo del tamaño de muestra y obtención de información de las poblaciones a partir de sus muestras, que sean lo estadísticamente significativas, y que la información obtenida de los datos permita hacer inferencias sobre las características de la población de estudio, y que al hacerlo de esta manera se cometan errores que tengan las características de ser acotados y medibles (Dieterich, 1996; Hund et al., 2015; Ruiz, 2016). El muestreo simple al azar (MSA) para proporciones es una técnica de muestreo utilizada en estadísticas y análisis de datos para obtener una muestra representativa de una población con el propósito de estimar una proporción específica dentro de esa población. Este método de muestreo se utiliza comúnmente en encuestas y estudios de investigación. Para realizar el muestreo simple al azar para proporciones de deben seguir los siguientes pasos (Cochran, 1977; Roos, 1999; Arias, 2000; García, 2002; Dawson, 2002; Martínez y Martínez 2008):

### *Definición de la población*

Primero, debes definir la población que deseas estudiar. La población es el conjunto completo de elementos o individuos que estás interesado en analizar en tu estudio.

### *Tamaño de la muestra*

Debes determinar el tamaño de la muestra que deseas recopilar para tu estudio. El tamaño de la muestra se refiere al número de elementos o individuos que seleccionarás de la población. El tamaño de la muestra suele ser un subconjunto relativamente pequeño de la población total.

### *Selección aleatoria*

Para realizar el muestreo simple al azar, cada elemento o individuo en la población debe tener la misma probabilidad de ser seleccionado en la muestra. Esto significa que se utiliza un proceso de selección al azar, como el uso de números aleatorios o una tabla de números aleatorios, para elegir los elementos de la población de manera aleatoria y sin ningún sesgo.

El muestreo aleatorio es un método de selección de elementos de una población de manera que cada unidad tenga la misma probabilidad de ser incluida en la muestra. Sin embargo, antes de realizar un muestreo aleatorio, es importante tener en cuenta ciertos supuestos y criterios para garantizar la validez de los resultados. Algunos supuestos y criterios importantes son:

## **Supuestos**

### *Independencia de las unidades de muestreo*

Se asume que la selección de una unidad de muestreo no afecta la probabilidad de selección de otra. Cada unidad debe ser seleccionada de manera independiente.

### *Homogeneidad de la población*

Se supone que la población es homogénea en relación con las características que se están estudiando. Si hay subgrupos significativos, se debe considerar un muestreo estratificado.

### *Condiciones estables*

Se asume que las condiciones en la población no cambian durante el proceso de muestreo. Cambios en las condiciones podrían afectar los resultados.

## **Criterios para elegir el muestreo aleatorio**

### *Lista o marco de muestreo*

Es esencial tener una lista completa de todas las unidades en la población (marco de muestreo) para que cada unidad tenga la posibilidad de ser seleccionada.

### *Número suficiente de unidades*

La muestra debe ser lo suficientemente grande para proporcionar resultados estadísticamente significativos. La determinación del tamaño de la muestra debe basarse en consideraciones estadísticas y prácticas.

### *Uso de métodos aleatorios*

Utilizar métodos aleatorios como el uso de tablas de números aleatorios, generadores de números aleatorios computarizados o técnicas de muestreo aleatorio sistemático.

### *Ausencia de sesgo en la selección*

Se debe evitar cualquier sesgo en el proceso de selección. Cada unidad debe tener una probabilidad igual de ser seleccionada.

### *Conservación de la aleatoriedad*

Una vez que una unidad ha sido seleccionada, debe ser eliminada del grupo de posibles selecciones para conservar la aleatoriedad en la elección de las unidades restantes.

### *Registro y documentación*

Es fundamental registrar y documentar todo el proceso de muestreo, incluidos los métodos utilizados y cualquier desviación de los procedimientos planeados.

### *Análisis y verificación*

Después de completar el muestreo, es crucial realizar análisis estadísticos para verificar la representatividad de la muestra en comparación con la población.

### *Recopilación de datos*

Una vez que hayas seleccionado aleatoriamente los elementos de la población que formarán parte de tu muestra, recopilarás los datos necesarios de cada elemento en la muestra. En el contexto de proporciones, podrías estar interesado en, por ejemplo, la proporción de individuos con cierta característica o comportamiento en la población.

### *Análisis de datos*

Después de recopilar los datos de la muestra, puedes utilizar estos datos para estimar la proporción en la población completa. Dependiendo de tus objetivos de investigación, puedes realizar análisis estadísticos para calcular estimaciones puntuales y obtener intervalos de confianza para la proporción.

Cuando se realiza un muestreo aleatorio para obtener una muestra de una población y se busca hacer estimaciones puntuales sobre parámetros de interés en la población, se pueden realizar diversos análisis estadísticos. Los pasos y técnicas comunes para calcular estimaciones puntuales en muestreo son:

## Calcular la estimación puntual

La estimación puntual es el valor único que se elige para representar la característica de interés en la población. Por ejemplo, si se está estimando la media de una variable, la estimación puntual sería simplemente la media de la muestra.

## Intervalo de confianza

Junto con la estimación puntual, es común calcular un intervalo de confianza. Este intervalo proporciona una medida de la precisión de la estimación puntual y generalmente se expresa como un rango dentro del cual se espera que esté el verdadero parámetro poblacional con cierto nivel de confianza.

## Desviación estándar y error estándar

Calcular la desviación estándar de la muestra ( $s$ ) es esencial para estimar la variabilidad de la población. El error estándar (SE) se calcula dividiendo la desviación estándar por la raíz cuadrada del tamaño de la muestra ( $n$ ). El SE, se utiliza en la construcción de intervalos de confianza y pruebas de hipótesis.

$$SE = \frac{s}{\sqrt{n}}$$

## Tamaño de la muestra

El tamaño de la muestra ( $n$ ) es crítico para la precisión de las estimaciones. Un tamaño de muestra más grande generalmente conduce a estimaciones más precisas.

## Pruebas de hipótesis

Se pueden realizar pruebas de hipótesis para determinar si hay evidencia suficiente para rechazar una afirmación sobre la población. Esto implica comparar la estimación puntual con un valor hipotético y calcular un valor  $p$  para evaluar la significancia estadística.

### **Coeficiente de variación (CV)**

El coeficiente de variación es otra medida de la variabilidad relativa de una muestra y se calcula dividiendo la desviación estándar por la media. Es especialmente útil cuando se comparan características de poblaciones con escalas diferentes.

$$CV = \frac{s}{\bar{x}} \times 100$$

### **Análisis de regresión (si es aplicable)**

Si la relación entre dos variables es de interés, se puede realizar un análisis de regresión para modelar la relación y hacer predicciones.

### **Comparaciones múltiples (si es aplicable)**

Si se están realizando múltiples comparaciones, se deben tener en cuenta los ajustes para controlar el error tipo I, como el método de Bonferroni.

### **Visualización de datos**

Utilizar gráficos, como gráficos de dispersión, histogramas o diagramas de caja, para visualizar la distribución de la muestra y cualquier patrón que pueda ser relevante.

### **Revisión de supuestos**

Revisar los supuestos del método estadístico utilizado y asegurarse de que se cumplan para interpretar correctamente los resultados.

La elección de la técnica estadística dependerá de la naturaleza de tus datos y de la pregunta de investigación específica que estás abordando. Además, siempre es recomendable consultar con un estadístico o experto en la materia para garantizar la validez y la interpretación correcta de los resultados. El muestreo simple al azar es una técnica fundamental en estadísticas y proporciona una forma eficaz de obtener estimaciones representativas de una población sin introducir sesgos significativos en la selección de la muestra. Esto es especialmente útil

cuando la población es grande y es costoso o impracticable recopilar datos de todos sus elementos (Cameron y Trivetti, 2009; Arias, 2000; Hund et al., 2015).

### **Cálculo del tamaño de muestra**

El tamaño de muestra para muestreo cualitativo se calcula con la siguiente ecuación (Cochran, 1977; Ross, 1999; Dawson, 2002; Hund et al., 2015):

$$n = \frac{N Z^2_{\alpha/2} p_n q_n}{N d^2 + Z^2_{\alpha/2} p_n q_n}$$

donde:

$d$                       Precisión

$Z_{\alpha/2}$                   Confiabilidad

$N$                       Tamaño de la población

$p_n$                     Proporción con la característica de interés

$q_n$                     Proporción sin la característica de interés

### **Ejemplos de aplicación**

#### **1. Percepción efectiva sobre la educación en línea ante la contingencia sanitaria por el covid-19**

El presente estudio fue de nivel exploratorio-descriptivo, pues como afirman Hernández y Sampieri (2018) citado por Ruiz (2019) pretende examinar un tema o problema de investigación poco estudiado e identificar sus propiedades, características y elementos importantes. La metodología a utilizarse fue para un estudio de caso, que, según Hernández Sampieri (2018) citado Zermeño (2011) consiste en:

(...) una investigación que, mediante los procesos cuantitativo, cualitativo y/o mixto, se analiza profundamente una unidad integral para responder al planteamiento del problema, probar hipótesis y

desarrollar alguna teoría; y citan a Mertens (2005) quien define el estudio de caso como una investigación sobre un individuo, grupo organización, proceso, etc., que es visto y analizado como una entidad. Gerring, (2004), señala que el estudio de caso consiste en la investigación intensiva de un fenómeno con la finalidad de comprenderlo en profundidad. Dicho fenómeno se convierte en la unidad y puede ser una persona, organización, revolución, etcétera.

Por lo que, a pesar de que se utilizará un cuestionario, la sistematización de los datos pretende tener una generalización estadística y también, ya que las muestras son estadísticamente representativas, como lo afirma Ruiz (2019), con la expansión a otros casos de una teoría o de un modelo que ha permitido analizar (exitosamente) un caso concreto. Se considerará cada institución analizada, como un caso emblemático, del cual se podrían inferir características comunes, a otras instituciones similares a la analizada (Hernández, 2006; Hund et al., 2015).

### **Objetivo general**

El objetivo de este trabajo fue generar un artículo donde se aplicarán métodos estadísticos aplicados al área de los recursos naturales y la sostenibilidad, y por lo tanto, se decidió explorar la percepción de los estudiantes de educación media Superior con el aprendizaje/enseñanza en línea en medio de la crisis sanitaria por el COVID-19 en instituciones educativas de México desde la perspectiva de los estudiantes.

#### *Instrumento*

El cuestionario es una adaptación de “Effects of Communication Mode on Social Presence, Virtual Presence, and Performance in Collaborative Virtual Environment” por Eva-Lotta Sallnäs (2006) citado por Argandoña et al. (2020). Los cuestionarios finales, utilizados para encuestar a los estudiantes fueron desarrollados por el equipo de investigación que elaboró la propuesta general. Los instrumentos fueron validados por dos expertos para validar la terminología utilizada y su claridad (Dieterich, 1996; Hund et al., 2015).

El cuestionario tiene un total de 18 ítems (opción múltiple), organizados en las siguientes seis secciones:

1. Datos demográficos.

2. Recursos y condiciones de trabajo para el aprendizaje en línea.
3. Comparación experiencia de aprendizaje presencial con su experiencia de enseñanza en línea.
4. Efectos de la enseñanza en línea en la comunicación interpersonal e interacción social.
5. Impacto en aspectos subjetivos.

## **Operativización de términos, variables y categorías analíticas**

En el presente estudio se definieron los términos centrales de la investigación como se detalla en las siguientes definiciones:

### **Percepción**

Con respecto a la noción de percepción, se entenderá como la definición de Allport 1974; Cohen, 1973; Coren y Ward, 1979; Ardila, 1980; Day, 1981a; Rock, 1985, compendiada por Vargas (2020) que la define como: "(...) proceso cognitivo de la conciencia que consiste en el reconocimiento, interpretación y significación para la elaboración de juicios en torno a las sensaciones obtenidas del ambiente físico y social."

### *Aprendizaje y enseñanza en línea*

Se tomará como referencia, para definir el aprendizaje y la enseñanza en línea como (Abreu, 2020):

#### *Enseñanza en línea:*

“Acercamiento con los estudiantes para facilitar actividades escolares, procesos de inducción a las plataformas digitales y acompañamiento en el proceso de integración dentro de la comunidad educativa”.

#### *Aprendizaje en línea:*

“Actividades pedagógicas de refuerzo de aprendizajes previos y espacios de reflexión por medios virtuales y acceso y uso de la Caja de Herramientas con recursos educativos digitales abiertas”.

### *Recogida y sistematización de datos*

Los datos fueron recabados aplicando el cuestionario a estudiantes de educación media superior de la UACH y del CBTA. La sistematización de los ítems cuantitativos, se realizaron en Excel. La sistematización de los ítems cualitativos, se realizará manualmente, usando una codificación de categorías de todas las respuestas, para luego ser contabilizadas en Excel (Hernández, 2006).

En este enfoque, cada elemento de la población tiene una probabilidad igual de ser seleccionado. Esto se logró utilizando números aleatorios computarizados donde cada uno de los estudiantes tenían la misma probabilidad de ser seleccionados. Este método garantiza que la selección de datos sea aleatoria y libre de sesgo, lo que es esencial para obtener estimaciones precisas y generalizables sobre la población de los estudiantes (Hund et al., 2015).

### *Diseño de muestreo*

Se empleó un diseño de muestreo complementario al azar para proporciones (Cochran, 1977; Hund et al., 2015). En la presente investigación se utilizó un enfoque cuantitativo, ya que la intención fue conocer la percepción de los estudiantes de nivel medio superior que tuvieron que migrar de los estudios presenciales a la educación a distancia, las ventajas y desventajas del mismo y el impacto que perciben en sus logros académicos.

El enfoque es no experimental ya que se estudió el fenómeno de manera natural, no existiendo manipulación deliberada de las variables, solo se observan las percepciones de los estudiantes en el cambio de modalidad (Zermeño, 2011; Hernández, 2006; Goto et al., 2014).

Por otro lado, se desarrolló de manera transversal ya que solo se recolectó la información en un momento único para describir las variables a analizar y el diseño es correlacional– causal para identificar la percepción de los estudiantes en el cambio de estudios de la modalidad presencial a la virtual (Martínez y Martínez, 2008; Bai et al., 2013).

### *Recolección de los datos*

El instrumento utilizado para la recolección de los datos fue una encuesta, la cual permitió obtener información cuantitativa, es confiable y con validez (Dawson, 2002; Zermeño, 2011; Goto et al., 2014) ya que fue posible aplicarla en diferentes momentos obteniendo una información similar.

La recolección se llevó a cabo a través de la aplicación de cuestionarios, posteriormente se recopilieron las respuestas de cada uno de los encuestados, se revisó la información. Posteriormente se recopiló la información en el programa Excel para obtener las gráficas y los análisis pertinentes de la información.

### *Análisis de datos*

Posteriormente se creó una base de datos con los resultados más significativos. La información se analizó con estadística descriptiva utilizando el programa de Excel 365 Pro plus (Hund et al., 2015).

### *Cálculo del tamaño de muestra*

La muestra se determinó de acuerdo con los datos de los alumnos matriculados en la Universidad Autónoma Chapingo y Centro de Bachillerato Tecnológico Agropecuario de una población de 10,000 estudiantes y fue de 261.9 y se aproxima al inmediato superior: 262 estudiantes el tamaño de la muestra (Evans y Rosenthal, 2004; Goto et al., 2014):

|                                               |        |
|-----------------------------------------------|--------|
| N =                                           | 10,000 |
| $Z_{\alpha/2}$ = Nivel de confianza (alpha) = | 0.95   |
| d (error) =                                   | 0.05   |
| $p_n$ =                                       | 0.50   |
| $q_n$ =                                       | 0.50   |

Donde N es el tamaño de la población, la precisión o error ( $d = 0.05$ ) es el máximo alejamiento —que el investigador decide que debe tener— entre el estimador (media muestral) y el parámetro (media poblacional), la probabilidad de error Z (es el error que estamos dispuestos a cometer o tolerar),  $p_n$  es la proporcional de la variable de interés que encontramos en la población de estudios y  $q_n$  es la proporción de la variable que complementa el 100 de la población (Hernández et al., 2006; Goto et al., 2014).

## MUESTREO CUALITATIVO

### TAMAÑO DE MUESTRA

$$n = \frac{N Z^2_{\alpha/2} p_n q_n}{N d^2 + Z^2_{\alpha/2} p_n q_n}$$

|                  |                                                          |                        |       |
|------------------|----------------------------------------------------------|------------------------|-------|
| N                | Tamaño de la Población                                   | Tamaño de la Población | 10000 |
| d                | Precisión                                                | Precisión              | 0.05  |
| Z <sub>α/2</sub> | Confiabilidad. Valor de Z (distribución normal estándar) | Confiabilidad          | 1.64  |
| p <sub>n</sub>   |                                                          | p <sub>n</sub>         | 0.5   |
| q <sub>n</sub>   | Varianza                                                 | q <sub>n</sub>         | 0.5   |

### Tamaño de la muestra

$$n = 261.9$$

El nivel de confiabilidad (alpha) o de probabilidad representa el porcentaje con el que se realizó la estimación del parámetro por medio de un estadístico muestral, en este caso del 95 % y la probabilidad de error en la investigación fue de un 5 %, con la aplicación de la fórmula estadística para la determinación de la muestra (Evans y Rosenthal, 2004; Hund et al., 2015).

Se realizaron 262 entrevistas, aplicándose cuestionarios a estudiantes de educación media superior de la UACH y del CBTA, de forma aleatoria y que cambiaron de la modalidad presencial a la modalidad en línea por la situación de salud en México.

## Resumen de las conclusiones del trabajo de investigación

A partir de esta investigación se observó que será interesante realizar otras investigaciones para conocer de manera cuantitativa y cualitativa la calidad en la educación en línea en las instituciones que migraron de lo presencial a lo virtual, es decir identificar realmente si los estudiantes están alcanzando los objetivos de aprendizaje esperados, y en caso de no ser así contar con información valiosa que permita a las instituciones adecuar sus programas de estudio, estrategias de aprendizaje y herramientas para lograr que el proceso de aprendizaje sea el adecuado para los estudiantes.

De acuerdo con los resultados obtenidos de los cuestionarios se observa que en general la percepción de los estudiantes a nivel media superior sobre la educación virtual ha sido positiva, aunque la mayoría de los encuestados ya

conocían esta modalidad al experimentarla al 100 % en sus estudios universitarios, ha tenido un impacto más importante.

## **2. Seguridad alimentaria en la comunidad indígena forestal mayo-yoreme en el norte de Sinaloa**

El presente estudio se ubicó en el municipio de El Fuerte Sinaloa, México, en la comunidad forestal indígena mayo-yoreme de Charay, un asentamiento que data desde 1550 situado sobre el margen izquierdo del Río Fuerte a 60 kilómetros del mar de Cortés. Viven alrededor de 3,041 personas, cuenta con una iglesia y una plazuela principal con arquitectura moderna. Es el pueblo más poblado en la posición número 6 de todo el municipio del Fuerte. La comunidad indígena Charay está a 26 metros de altitud. Presenta un índice de marginación alta. Las actividades económicas se basan en actividades principalmente forestales, seguidas de la agricultura y la pesca (INEGI, 2020).

El presente estudio fue de nivel exploratorio-descriptivo, pues como afirman Hernández y Sampieri (2018) citado por Ruiz (2019) pretende examinar un tema o problema de investigación poco estudiado e identificar sus propiedades, características y elementos importantes. La metodología a utilizarse fue para un estudio de caso, que, según Hernández Sampieri (2018) citado por Zermeño (2011) consiste en:

(...) una investigación que, mediante los procesos cuantitativo, cualitativo y/o mixto, se analiza profundamente una unidad integral para responder al planteamiento del problema, probar hipótesis y desarrollar alguna teoría; y citan a Mertens (2005) quien define el estudio de caso como una investigación sobre un individuo, grupo organización, proceso, etc., que es visto y analizado como una entidad. Gerring, (2004), señala que el estudio de caso consiste en la investigación intensiva de un fenómeno con la finalidad de comprenderlo en profundidad. Dicho fenómeno se convierte en la unidad y puede ser una persona, organización, revolución, etcétera.

Por lo que, a pesar de que se utilizará un cuestionario, la sistematización de los datos pretenden tener una generalización estadística y analítica, tendrán que ver, como lo afirma Ruiz (2019), con la expansión a otros casos de una teoría o de un modelo que ha permitido analizar (exitosamente) un caso concreto. Se considerará la población de Charay Municipio de El Fuerte Sinaloa, como un

caso emblemático, del cual se podrían inferir características comunes, a otras comunidades indígenas con problemas de inseguridad alimentaria similares a la analizada.

### *Objetivo*

El objetivo de este trabajo fue generar un artículo donde se aplicarán métodos estadísticos aplicados al área de los recursos naturales y en este caso fue conocer cómo se encuentra la seguridad alimentaria en la comunidad mayo-yoreme en el norte de Sinaloa. Se analizaron estudios relacionados con la salud y nutrición de la población en 2022 a comunidades indígenas forestales en el Municipio del Fuerte.

### *Instrumento*

El equipo de investigación elabora un cuestionario como propuesta general. El instrumento fue validado por tres expertos para validar la terminología utilizada y su claridad empleada. El Cuestionario tuvo un total de 30 ítems (opción múltiple) organizados en tres secciones:

1. Socioeconómicos.
2. Antropométricos.
3. Disponibilidad de alimentos y diversidad dietética.

### *Operativización de términos, variables y categorías analíticas*

En el presente estudio se entenderán los términos centrales de la investigación como se detalla en las siguientes definiciones:

### **Seguridad alimentaria**

En la Cumbre Mundial de la Alimentación (CMA) de 1996, organizada por la FAO (1996), se acuerda que existe seguridad alimentaria y nutricional en la población de cualquier país: “Cuando todas las personas, en todo momento, tienen acceso físico y económico a alimentos suficientes, seguros y nutritivos, para satisfacer sus necesidades alimenticias y sus preferencias, con el objeto de llevar una vida activa y sana” (FAO, 1996).

Los elementos cualitativos y cuantitativos básicos para tener seguridad alimentaria y nutricional

Los elementos básicos para que las personas tengan seguridad alimentaria son (FAO, 2016; FAO, 2020):

- 1) La disponibilidad de alimentos a nivel local o nacional.
- 2) La estabilidad, que se refiere a tener control en los ciclos agrícolas de los cultivos, también se debe contar con silos y almacenes para posibles contingencias en épocas donde falte alimentos de la canasta básica.
- 3) El acceso y control sobre los medios de producción, como la tierra, agua, insumos, tecnología, conocimiento y a los alimentos disponibles en el mercado; y por último.
- 4) El consumo y utilización biológica, que se refiera a la existencia e inocuidad de los alimentos, además de contar con condiciones higiénicas y una distribución equitativa de éstos, dentro de los hogares (FAO, 2020).

### *Inseguridad alimentaria*

Si alguna de cualquiera de estas características anteriormente mencionadas no se tiene, podemos afirmar que hay inseguridad alimentaria y nutricional en las personas, hogares, grupos de población, municipios, estados o países (FAO, 2016, FAO, 2019).

### *Recogida y sistematización de datos*

Se analizaron los resultados realizados en 2022 relacionados con la Seguridad alimentaria (SA) de la comunidad de Charay. El cuadro 1 muestra el tipo de estudio, el grupo de edad y las variables evaluadas. Las herramientas utilizadas para evaluar la SA de la comunidad se muestran en el cuadro 2.

En este enfoque, cada elemento de la población tiene una probabilidad igual de ser seleccionado. Esto se logró utilizando números aleatorios computarizados donde cada uno de los sujetos tenían la misma probabilidad de ser seleccionados. Este método garantiza que la selección de datos sea aleatoria y libre de sesgo, lo que es esencial para obtener estimaciones precisas y generalizables sobre la población de los estudiantes.

Cuadro 1. Descripción de los estudios analizados.

| Estudio                                                                         | Tipo de estudio | Individuos                                                 | Variables                                                                                                                                                                       |
|---------------------------------------------------------------------------------|-----------------|------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Vulnerabilidad por inseguridad alimentaria                                      | Transversal     | 199 mujeres entre 18 y 91 años                             | Acceso económico a los alimentos                                                                                                                                                |
| Nivel socioeconómico                                                            | Transversal     | 199 mujeres entre 18 y 91 años                             | Bienes y servicios básicos en el hogar                                                                                                                                          |
| Transición nutricional en adultos                                               | Transversal     | 64 mujeres y 11 hombres                                    | Antropométricas:<br>Peso, talla, IMC, % masa/grasa<br>Socioeconómicas: ocupación del jefe del hogar<br>Diversidad dietética<br>Fuente de alimentos                              |
| Intervención en la dieta de niños menores de cinco años                         | Longitudinal    | 45 niños al inicio y 27 al final del estudio de 1 a 5 años | Antropométricas: anemia (puntuación Z OMS), talla/edad, IMC/edad<br>Socioeconómicas: ocupación del jefe del hogar, ingreso mensual, gastos en alimentos<br>Diversidad dietética |
| Índice de alimentación saludable en población pediátrica                        | Transversal     | 135 niños y adolescentes de 2 a 16 años                    | Diagnostico por puntuación Z OMS para la relación talla/edad y el IMC/edad<br>Diversidad dietética                                                                              |
| Validación de un instrumento para medir seguridad alimentaria familiar indígena | Transversal     | 98 mujeres y 2 hombres de entre 22 y 71 años               | Escolaridad del jefe del hogar, ocupación, abastecimiento de alimento, fuentes de agua<br>Diversidad dietética<br>Saneamiento básico                                            |

Cuadro 2. Herramientas utilizadas para evaluar la seguridad alimentaria en población indígena de Charay, municipio del Fuerte Sinaloa.

| Herramientas                                                                                         | Descripción                                                                                                                                                                                                                                                |
|------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Escala del Componente de Acceso de la Inseguridad Alimentaria en el Hogar (HFIAS)                    | Mide el acceso físico y económico de los alimentos. Tiene una escala de cuatro categorías.                                                                                                                                                                 |
| Índice Nivel Socioeconómico de la Asociación Mexicana de Agentes de Investigación del Mercado (AMAI) | Clasifica a los hogares por su capacidad de satisfacer las necesidades de sus integrantes. La clasificación va desde el nivel más alto “A/B”, hogar que invierte más en la educación; hasta el nivel más bajo “E”, donde el mayor gasto es para alimentos. |
| Patrones de alimentación y frecuencia de alimentos                                                   | Mide patrones de alimentación, aspectos culturales, laborales y estilo de vida. Consumo de alimentos locales e industrializados.                                                                                                                           |
| Recordatorio de 24 horas                                                                             | Analiza la calidad de la dieta con el consumo de calorías, de micro y macronutrientes.                                                                                                                                                                     |
| Índice de Alimentación Saludable (IAS)                                                               | Evalúa la calidad de la dieta el ambiente alimentario. Se adoptó a los factores sociales, económicos, culturales y gastronómicos de la comunidad indígena.                                                                                                 |
| Cuestionario de Seguridad Alimentaria Familiar Indígena (CUSAFI)                                     | Mide acceso, utilización y la disponibilidad de los alimentos. Tiene cuatro categorías: SA, IA leve, IA moderada, IA severa.                                                                                                                               |

## Recolección de los datos

El instrumento utilizado para la recolección de los datos fue una encuesta, la cual permitió obtener información cuantitativa, es confiable y con validez (Zermeno, 2011; Goto et al., 2014) ya que fue posible aplicarla en diferentes momentos obteniendo una información similar.

La recolección se llevó a cabo a través de la aplicación de cuestionarios, posteriormente se recopilaron las respuestas de cada uno de los encuestados, se

revisó la información. Posteriormente se recopiló la información en el programa Excel para obtener las gráficas y los análisis pertinentes de la información (Goto et al., 2014)

### *Análisis de datos*

Posteriormente se creó una base de datos con los resultados más significativos. La información se analizó con estadística descriptiva utilizando el programa de Excel 365 Pro plus.

### *Diseño de muestreo*

Se empleó un diseño de muestreo Complemente al Azar para Proporciones (Angers, 1974, Cochran, 1977; Goto et al., 2014). En la presente investigación se utilizó un enfoque cuantitativo, ya que la intención fue conocer la inseguridad alimentaria en la población de Charay El Fuerte Sinaloa.

El enfoque es no experimental ya que se estudió el fenómeno de manera natural, no existiendo manipulación deliberada de las variables, solo se observan las percepciones de los estudiantes en el cambio de modalidad (Zermeño, 2011; Goto et al., 2014).

Por otro lado, se desarrolló de manera transversal ya que solo se recolectó la información en un momento único para describir las variables a analizar y el diseño es correlacional-causal (Martínez y Martínez, 2008) para identificar la inseguridad alimentaria de los pobladores de Charay, El fuerte Sinaloa.

### *Cálculo del tamaño de la muestra*

La muestra se determinó de acuerdo con los datos de los pobladores de la comunidad de Charay El Fuerte Sinaloa de una población de 3041 (Evans y Rosenthal, 2004; Goto et al., 2014):

|                           |       |
|---------------------------|-------|
| N =                       | 3047  |
| Nivel de certeza (alpha): | 0 .95 |
| Error:                    | 0.05  |
| p <sub>n</sub> =          | 0.50  |
| q <sub>n</sub> =          | 0.50  |

## MUESTREO CUALITATIVO

### TAMAÑO DE MUESTRA

$$n = \frac{N Z^2 \frac{p_n q_n}{d^2}}{N Z^2 \frac{p_n q_n}{d^2} + Z^2 \frac{p_n q_n}{d^2}}$$

|                |                                                          |                        |      |
|----------------|----------------------------------------------------------|------------------------|------|
| N              | Tamaño de la Población                                   | Tamaño de la Población | 3041 |
| d              | Precisión                                                | Precisión              | 0.05 |
| Z%             | Confiabilidad. Valor de Z (distribución normal estándar) | Confiabilidad          | 1.64 |
| p <sub>n</sub> |                                                          | p <sub>n</sub>         | 0.5  |
| q <sub>n</sub> | Varianza                                                 | q <sub>n</sub>         | 0.5  |

### Tamaño de la muestra

247.1

El tamaño de muestra calculado fue de 247.1 y se aproxima al valor entero superior; serian entonces 248 individuos de Charay.

El nivel de certeza (alpha) o de probabilidad representa el porcentaje con el que se realizó la estimación del parámetro por medio de un estadístico muestral, en este caso del 95 % y la probabilidad de error en la investigación fue de un 5 %, con la aplicación de la fórmula estadística para la determinación de la muestra (Evans y Rosenthal, 2004; Goto et al., 2014):

Se realizaron 247 entrevistas, aplicándose cuestionarios a pobladores que de la comunidad de Charay, El Fuerte, Sinaloa, de forma aleatoria.

## Resumen de las conclusiones del trabajo de investigación

Se identificó la participación de 248 hogares, las madres se dedican al hogar y los padres son jornaleros, el ingreso promedio mensual por familia es de \$1,522.70 y el gasto en alimentos de \$ 804.3. Todos los grupos de edad presentan malnutrición, tienen dieta poco diversa alta en carbohidratos. Presentan estrategias para obtener alimentos como regalos de terceros, intercambio o pedir dinero prestado. Los hogares se encuentran en inseguridad alimentaria, sufren de experiencias de hambre y ayunos prolongados. La comunidad indígena es vulnerable a presentar inseguridad alimentaria y nutricional por su bajo ingreso económico.

Es importante garantizar la autosuficiencia alimentaria familiar y promover en consumo de dietas variadas basadas en la producción de alimentos locales.

Se concluyó de manera general que en la población indígena forestal de Charay se encuentra en inseguridad alimentaria debido a su bajo ingreso económico lo que se refleja en la carencia alimentaria y la pobre diversidad dietética; carecen de agua potable y dependen de apoyos gubernamentales para su subsistencia, afectando principalmente a mujeres en edad reproductiva y ancianos.

En la población indígena forestal de Charay se encuentra Inseguridad Alimentaria debido a su bajo ingreso económico lo que se refleja en la carencia alimentaria y la pobre diversidad dietética; carecen de agua potable y dependen de apoyos gubernamentales para su subsistencia, afectando principalmente a mujeres en edad reproductiva y ancianos. Las estrategias de intervención en seguridad alimentaria y nutricional en comunidades indígenas deben asegurar la autosuficiencia alimentaria familiar y promover el consumo de dietas diversas basadas en alimentos producidos de manera local.

Debe ser prioritario incorporar estrategias preventivas contextualizadas en la realidad de las comunidades indígenas para disminuir la presencia de sobrepeso y obesidad como parte de la política de desarrollo social de México, ya que la política asistencialista no ha sido suficiente como estrategia preventiva.

## **Conclusiones generales**

El muestreo completamente al azar para proporciones, o muestreo aleatorio simple, es una técnica fundamental en las investigaciones de las disciplinas sociales o del ambiente, y tiene varias ventajas y una gran importancia en diversas aplicaciones. Las razones clave de su importancia para su utilización son:

**Representatividad:** El muestreo completamente al azar garantiza que cada elemento de la población tenga la misma probabilidad de ser seleccionado en la muestra. Esto hace que la muestra sea representativa de la población en su conjunto, lo que es esencial para hacer inferencias precisas sobre la población.

**Control de sesgo:** Al eliminar cualquier sesgo en la selección de la muestra, el muestreo aleatorio simple evita la introducción de prejuicios sistemáticos que podrían distorsionar los resultados de la investigación. Esto es crucial para obtener estimaciones imparciales de las proporciones en la población.

**Validez de las inferencias:** Cuando se utiliza un muestreo completamente al azar, las inferencias basadas en la muestra son más válidas y generalizables a la población completa. Esto permite a los investigadores tomar decisiones informadas y generalizar los resultados de la muestra a la población de interés.

**Simplicidad:** El muestreo aleatorio simple es un método relativamente sencillo de implementar y comprender, lo que lo hace accesible para una amplia

variedad de investigadores y profesionales. No requiere suposiciones complicadas y se puede aplicar en una variedad de contextos.

**Eficiencia:** A menudo, el muestreo aleatorio simple es más eficiente en términos de costos y tiempo en comparación con otros métodos de muestreo más complejos, como el muestreo estratificado o el muestreo por conglomerados. Esto es especialmente importante cuando se trabaja con grandes poblaciones.

**Fiabilidad de los resultados:** Al eliminar el sesgo y garantizar que la muestra sea representativa, el muestreo completamente al azar contribuye a la fiabilidad de los resultados del estudio. Esto significa que los resultados son más confiables y reproducibles.

**Cumplimiento de supuestos estadísticos:** En muchos análisis estadísticos y pruebas de hipótesis, se asume que los datos provienen de una muestra aleatoria simple para poder aplicar métodos estadísticos apropiados. El incumplimiento de este supuesto podría afectar negativamente la validez de los resultados.

El muestreo completamente al azar para proporciones es una herramienta esencial en la investigación debido a su capacidad para proporcionar muestras representativas y sin sesgo de una población, lo que, a su vez, garantiza que las inferencias basadas en la muestra sean confiables y válidas. Esto es fundamental en la toma de decisiones informadas, la investigación científica y muchas otras aplicaciones en las que se necesita estimar proporciones o características de una población.

## Referencias

- Abreu J. (2020). Tiempos de Coronavirus: La Educación en Línea como Respuesta a la Crisis. *Daena: International Journal of Good Conscience*, 15(1), 1-15. [http://www.spentamexico.org/v15-n1/A1.15\(1\)1-15.pdf](http://www.spentamexico.org/v15-n1/A1.15(1)1-15.pdf)
- Angers, C. (1974), "A Graphical Method to Evaluate Sample Sizes for the Multinomial Distribution," *Technometrics*, 16, 469-471.
- Arias, Fidias (2000). *El Proyecto de Investigación. Introducción a la Metodología Científica*. 4ta Edición. Editorial EPISTEME. Caracas – Venezuela.
- Arias-Gómez, J.; Villasís-Keever, M. Á. & Miranda-Navales, M. G. The research protocol III. Study population. *Rev. Alerg. Mex.*, 63(2):201-6, 2016.
- Ávila Baray, H. L. *Introducción a la Metodología de la Investigación*. Edición electrónica. Cuauhtémoc (Chihuahua), Instituto Tecnológico de Cd. Cuauhtémoc, 2006 Disponible en: <http://www.eumed.net/libros-gratis/2006c/203/index.htm>
- Bai, X.; Tsiatis, A. A. & O'Brien, S. M. Doubly-robust estimators of treatment-specific survival distributions in observational studies with stratified sampling. *Biometrics*, 69(4):830-9, 2013.
- Barón F. L (2005). *Apuntes de bioestadística: tercer ciclo en ciencias de la salud y medicina*: Universidad de Málaga España.
- Bromaghin, J.F. (1993). Sample size determination for interval estimation of multinomial parameters. *The American Statistician*, 47, 203-206.
- Cameron, C., y Trivedi, P. (2009). *Microeconometrics using Stata*. A Stata Press Publication. 732 pp.
- Cochran, WG. (1977). *Sampling Technics*. Third Edition. John Wiley y Sons, Inc. 442 pp.
- Dawson B. 2002. *Bioestadística médica*. 2002. 3era edición, Edit. Manual Moderno México Geoffrey R. 2001- Norman, David L. Streiner. *Bioestadística*: Edit. Harcourt Brace España.
- Dieterich, H. 1996. *Nueva Guía para la Investigación Científica*. Ciudad de México, Editorial Planeta Mexicana, 1996. Disponible en: [http:// www. ceuarkos.com/heinz.pdf](http://www.ceuarkos.com/heinz.pdf)
- Evans, M. y Rosenthal, J. (2004). *Probabilidad y estadística*. Editorial Reverte, S.A. 489 p.
- FAO. 1996. *Cumbre Mundial de Alimentación*. FAO, Roma. 1996.
- FAO. 2016. *Trabajo estratégico de la FAO para contribuir a la erradicación del hambre y la mala nutrición*. FAO, Roma. 2016. Documento: FAO.
- FAO. 2019. *Diagnóstico del sector rural y pesquero: identificación de la problemática del sector agropecuario y pesquero de México* [documento en

- internet]. México: FAO, Sagarpa. Disponible en: <http://www.sagarpa.gob.mx/programas2/evaluacionesExternas/Lists/Otras%20Evaluaciones/Attachments/22/Procampo%20Diagnostico%20FINAL%202022.03.212.p>
- FAO. 2020. El estado mundial de la agricultura y la alimentación. 2020. Superar los desafíos relacionados con el agua en la agricultura. Roma. <https://doi.org/10.4060/cb1447es>
- Fitzpatrick, S., y Scott, A. (1987). Quick simultaneous confidence intervals for multinomial proportions. *Journal of American Statistics Association*, 82, 875-878.
- García, A. (2002). La educación a distancia. De la teoría a la práctica. Barcelona: Ed. Ariel.
- Guillen, F. (2018, diciembre). Investigación cualitativa: Método fenomenológico hermenéutico. [http://www.scielo.org.pe/scielo.php?script=sci\\_arttext&pid=S2307-79992019000100010](http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S2307-79992019000100010).
- Goodman, L. (1965). On Simultaneous Confidence Intervals for Multinomial Proportions. *Technometrics*, vol. 7, No. 2, 247-254.
- Goto, R.; Arai, K.; Kitada, H.; Ogoshi, K. & Hamashima, C. 2014. Labor resource use for endoscopic gastric cancer screening in Japanese primary care settings: A work sampling study. *PLoS One*, 9(2):e88113.
- Hernández Sampieri, R.; Fernández-Collado, C. & Baptista Lucio, P. 2006. Metodología de la Investigación. 4ª ed. Ciudad de México, McGraw-Hill, 2006. Disponible en: [https://competenciashg.files.wordpress.com/2012/10/sampieri-et-al-metodologia-de-la-investigacion-4ta-edicionsampieri-2006\\_ocr.pdf](https://competenciashg.files.wordpress.com/2012/10/sampieri-et-al-metodologia-de-la-investigacion-4ta-edicionsampieri-2006_ocr.pdf).
- Hund, L.; Bedrick, E. J. & Pagano, M. 2015. Choosing a cluster sampling design for lot quality assurance sampling surveys. *PLoS One*, 10(6):e0129564, 2015.
- INEGI. 2020. Población total según tamaño de la localidad. Censo de Población y Vivienda 2020.
- Martínez, JA., y Martínez, L. (2008). Determinación de la máxima varianza para el cálculo del factor de Imprecisión sobre la Escala de Medida, y extensión a diferentes tipos de muestreo. *Psicothema*, Vol 20 n°2, 311-316.
- Ross, S. (1999). Simulación. Prentice Hall. México. 281 pp.
- Ruiz O., J. I. (2019). Metodología de la investigación cualitativa (5a. ed.). Publicaciones de la Universidad de Deusto. 567 p.
- Thompson, S. (1987). Sample Size for Estimating Proportions. *The American Statistician*, vol. 41, No.1, 42-46.
- Tortora, R. (1978). A Note on Sample Size estimation for Multinomial Populations. *The American Statistician*, vol. 32, No. 3, 100-102.
- Vargas, K. (2020). Enseñanza aprendizaje virtual en tiempos de pandemia. <http://repositorio.unesum.edu.ec/bitstream/53000/2554/1/Ense%C3%B1anza%20aprendizaje%20virtual%20en%20tiempos%20de%20pandemia.pdf>.

Walpole, R. E. & Myers, R. H. 1996. Probabilidad y Estadística. 4ª. ed. Ciudad de México, McGraw-Hill, 1996.

Zermeño, S. (2011). Métodos de investigación. Investigación para la vida. Pearson. 678 p.

## Capítulo 2

---

### **Análisis univariado y multivariado para evaluación de fitoplancton y calidad de agua en estanques acuícolas en el municipio de Ahome, Sinaloa, México**

*Jesús Manuel Díaz Gaxiola  
Jesús Octavio Escobedo Urquidez*

## Introducción

En el ámbito profesional, cada día hay que resolver los problemas que se presentan de la mejor manera y sin duda, el uso de la estadística es fundamental para comprender y tomar decisiones más acertadas. ¿Qué es la estadística? Estadística es la parte de las matemáticas que estudia las características de las muestras y de las poblaciones de las que provienen (Corral-Dávalos, 2019), por su parte Milton (2007) la define como el arte de la decisión frente a la incertidumbre. La ciencia basada en un proceso analítico y crítico produce el conocimiento que ha permitido una mejor comprensión de la realidad circundante. Asimismo, ha facultado al hombre para penetrar en los secretos más profundos del mundo, incluido el ser del hombre mismo (Pérez-Tejada, 2008).

En este capítulo, se muestra el análisis que se realiza a los datos obtenidos de una investigación sobre las microalgas y los parámetros físico-químicos del agua utilizada para el cultivo de camarón. El proceso de análisis se lleva a cabo por medio del software estadístico PAST (Hammer, Harper y Ryan, 2001) versión 4.15.

El análisis de una base de datos siempre partirá de técnicas simples de resumen de los datos y presentación de los resultados. A partir de estos resultados iniciales, y en función del diseño del estudio y de las hipótesis preestablecidas, se aplicarán las técnicas de inferencia estadística que permitirán obtener conclusiones acerca de las relaciones estructurales entre las variables estudiadas. Las técnicas de estadística descriptiva no precisan de asunciones para su interpretación, pero en contrapartida la información que proporcionan no es fácilmente generalizable. La estadística inferencial permite esta generalización, pero requiere ciertas asunciones que deben verificarse para tener un grado razonable de seguridad en las inferencias (Pastor-Barriuso, 2012).

En los estudios descriptivos, el investigador se limita a medir la presencia, características o distribución de un fenómeno en una población en un momento de corte en el tiempo, tal sería el caso de estudios que describen la presencia de un determinado factor ambiental, una determinada enfermedad, mortalidad en la población, etc., pero siempre referido a un momento concreto y sobre todo, limitándose a describir uno o varios fenómenos sin intención de establecer relaciones causales con otros factores (Veiga-de Cabo, De la Fuente-Díez y Zimmermann-Verdejo, 2008).

Al observar muchas variables sobre una muestra es presumible que una parte de la información recogida pueda ser redundante o que sea excesiva, en cuyo caso los métodos multivariantes de reducción de la dimensión (análisis en componentes principales, factorial, correspondencias, escalamiento óptimo, homogeneidades, análisis conjunto, etc.) tratan de eliminarla. Estos métodos combinan muchas variables observadas para obtener pocas variables ficticias que las representen (Pérez-López, 2004).

El objetivo del Análisis de Componentes Principales (ACP) es reducir la dimensionalidad de un conjunto de datos multivariante mediante la sustitución de las variables originales por transformaciones lineales de las mismas, obteniendo componentes incorreladas y con poca pérdida de información. De este modo, el método permite hacer una representación gráfica que podamos interpretar y también adecuar los datos para realizar cualquier otro tipo de análisis. Esta técnica, iniciada por Pearson en 1901 y desarrollada por Hotelling en 1933, no requiere supuesto de normalidad; solo que exista el vector de medias y la matriz de varianzas y covarianzas del vector aleatorio objeto de estudio (Muñoz-Armayones, 2016).

Por otro lado, los individuos pueden presentar ciertas características comunes en sus respuestas, que permitan intentar su clasificación en grupos de cierta homogeneidad. Los métodos de clasificación (análisis clúster, análisis discriminante, árboles de decisión, etc.) buscan analizar las relaciones entre variables para ver si se pueden separar los individuos en agrupaciones a posteriori (Pérez-López, 2004).

El análisis de conglomerados busca particionar un conjunto de objetos en grupos, de tal forma que los objetos de un mismo grupo sean similares y los objetos de grupos diferentes sean disímiles. Así, el análisis de conglomerados tiene como objetivo principal definir la estructura de los datos colocando las observaciones más parecidas en grupos (Díaz-Monroy, 2007).

La acuicultura representa un importante sector de la producción alimentaria mundial y constituye una elemental fuente de proteínas, empleo e ingresos, siendo la base del sustento de una gran parte de la población mundial. En concreto, el camarón es un producto de alto valor, que se produce principalmente en Asia y América Latina. En particular, en Latinoamérica, donde *Penaeus vannamei* (Familia Penaeidae) es la principal especie cultivada (FAO, 2004).

El objetivo principal es utilizar el programa estadístico PAST®, con la finalidad de realizar un análisis de fitoplancton y parámetros físico-químicos en muestras de agua de una granja camaronícola, para evidenciar el uso de los diversos menús del *software* y su aplicación.

## Metodología

El presente trabajo se realizó durante los meses de agosto a octubre de 2023, con dos muestreos mensuales para un total de 42 muestras. El estudio se llevó a cabo en los estanques de cultivo de camarón, canal de entrada y salida en la granja acuícola “Desembocadero Vecinos” S.A de C.V, que se localiza en el poblado Las Grullas Margen Izquierda, en el norte del estado de Sinaloa, México, en las coordenadas 25°49′48 N 109°23′13 W (figura 1).



Figura 1. Localización de la granja “Desembocadero Vecinos”, en el estado de Sinaloa, México (Imágenes satelitales de Google Earth, 2023).

El análisis de los datos se realiza en el programa PAST© (Hammer, Harper y Ryan, 2001), el cual puede ser obtenido de manera gratuita de la siguiente dirección electrónica: <https://www.nhm.uio.no/english/research/resources/past/>. Este programa no requiere instalación. Una vez descargado y abierto, muestra los 13 menús que despliega para diversos análisis (figura 2) y también muestra la zona de trabajo (zona central de columnas e hileras en blanco) la cual acepta matrices de Excel© para su examen. A todos los datos se les aplicó la prueba de normalidad de Shapiro-Wilk, posteriormente a los valores de parámetros físico-químicos se les realizó una estadística descriptiva para ver la variación de los datos. Finalmente, a la matriz de fitoplancton se le realizó un análisis de varianza para buscar diferencias en los muestreos o en los estanques, y también se les realizó un análisis de componentes principales y análisis clúster.

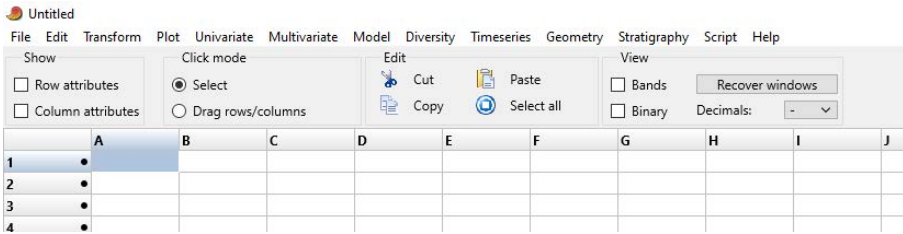


Figura 2. Opciones de análisis en el programa PAST.

El menú “Plot” será de utilidad para presentar los valores en gráficos de diversos tipos, de acuerdo con las necesidades (figura 3).

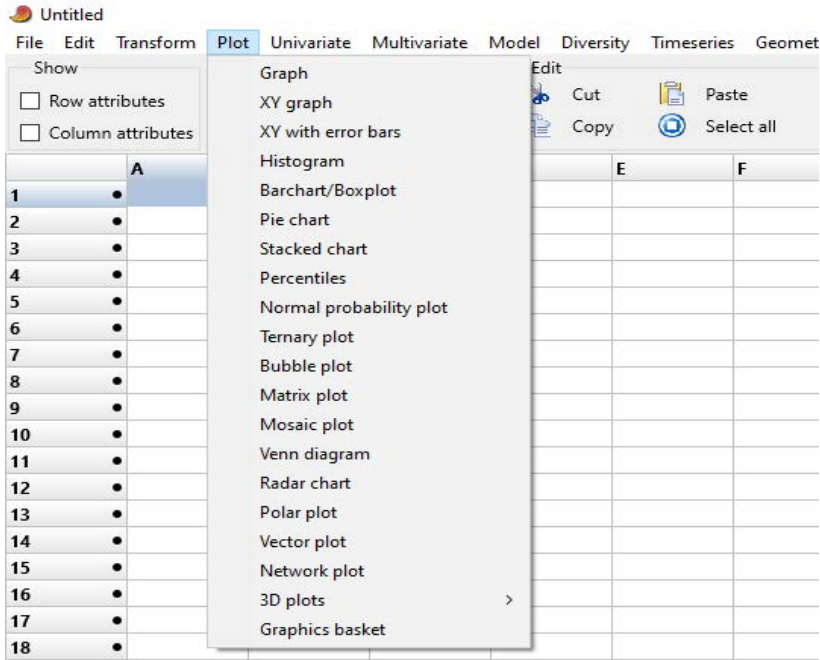


Figura 3. Menú Plot y sus diferentes opciones.

El siguiente menú es el de métodos univariados (Figura 4), el cual contiene varias opciones para análisis: estadística descriptiva, análisis de normalidad de datos, análisis de varianza (ANOVA), etcétera.

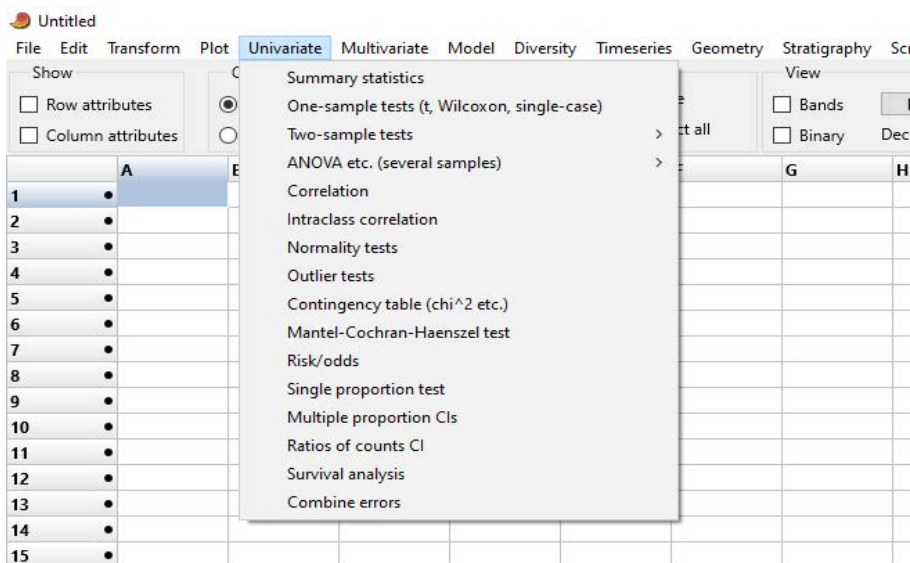


Figura 4. Menú Univariados y sus diferentes opciones.

**Normality Test:** Se utilizan para determinar si un conjunto de datos está bien modelado por una distribución normal y para calcular la probabilidad de que una variable aleatoria subyacente al conjunto de datos tenga una distribución normal (Shukla, 2015). Así mismo, se realizó el análisis de varianza (ANOVA) para los datos de parámetros físico-químicos en los estanques por muestreo. En la figura 5 se muestra el módulo de Univariados que contiene la instrucción para el cálculo de la normalidad de los datos para procesarse.

**Summary Statistic:** Las estadísticas de resumen (figura 6) La estadística descriptiva está formada por procedimientos empleados para resumir y describir las características importantes de un conjunto de mediciones (Mendenhall, Beaver y Beaver, 2006). Por otro lado, se tiene ahí mismo la opción para la realización del análisis de varianza (ANOVA) (figura 7).

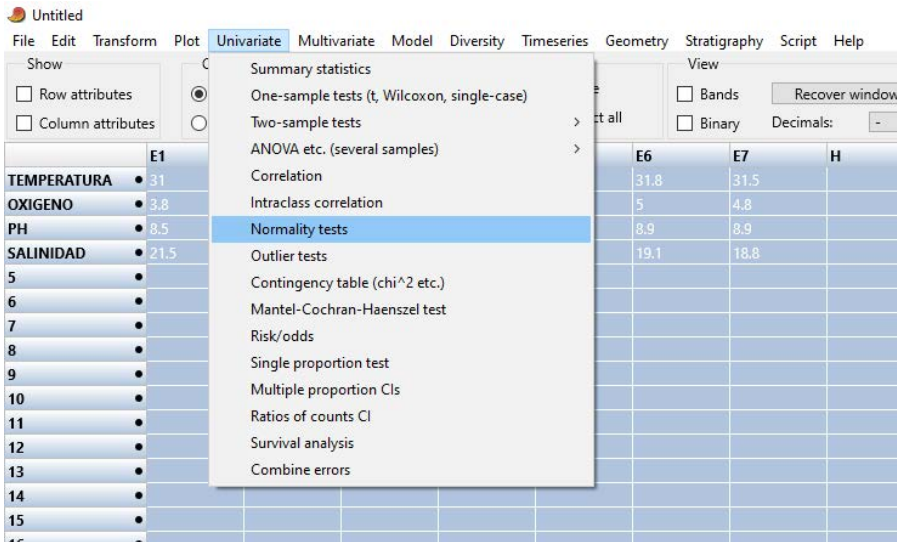


Figura 5. Módulo de Univariados y la opción de prueba de normalidad.

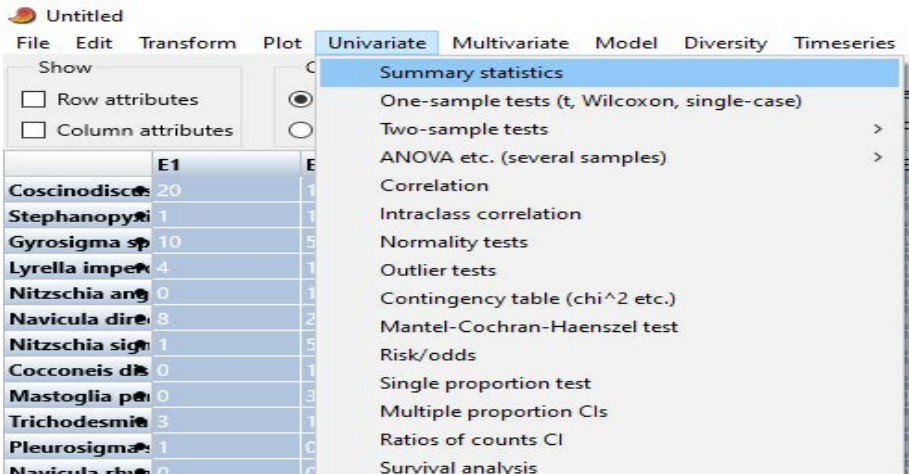


Figura 6. Módulo Univariado y la opción de summary statistics.

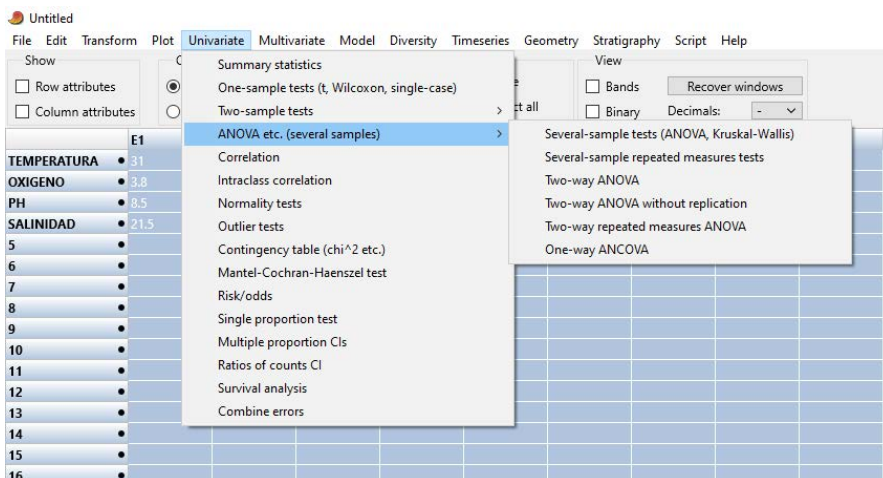


Figura 7. Módulo Univariado y la opción de análisis de varianza.

El módulo de Multivariados (figura 8) muestra las opciones (de ordenación, clúster, etc.) para realizar el análisis de datos. Una vez colocada la matriz de datos en el campo correspondiente, se elige la opción de Ordenación y se selecciona la opción de Principal Components (PCA) (figura 9), mostrando el resultado con todos los análisis realizados (figura 10), como los componentes principales (CP), los eigenvalores y el porcentaje de la varianza explicada del fenómeno, la gráfica obtenida con los vectores de mayor peso y los scores de los ítems con mayor valor. Se eligen generalmente los 2 primeros que explican el mayor porcentaje de la variación del fenómeno modelado. Así el primer eje (CP) contiene a la mayor varianza y el segundo eje contiene a la segunda varianza más grande y así sucesivamente. La otra opción dentro de multivariados lo constituye el Análisis Clúster, que en este caso fue realizado en la modalidad Classical clustering (figura 11) por el método de Mahalanobis, ya que esta es invariante a los cambios de escala (Correa-Henao, 2021).

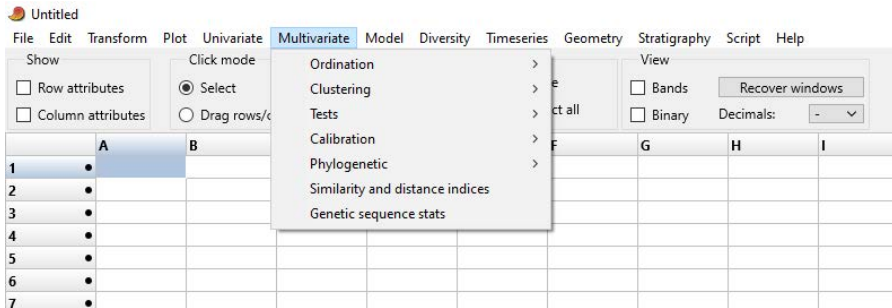


Figura 8. Menú multivariados y sus múltiples opciones.

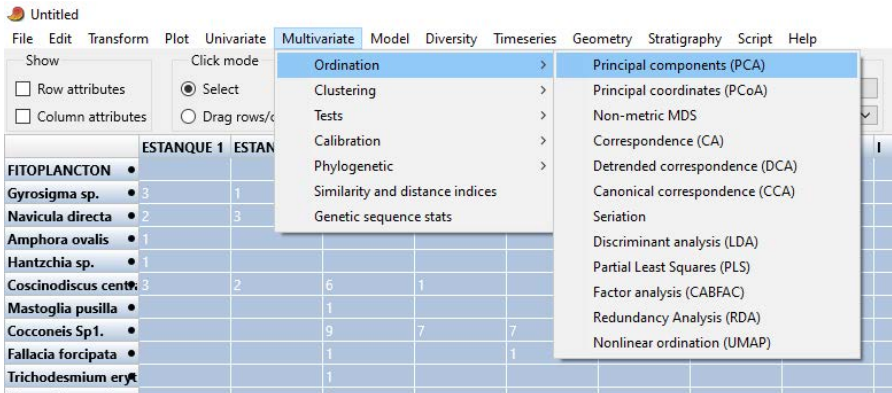


Figura 9. Opción de PCA en ordenación dentro del menú multivariado.

The screenshot shows the 'Principal components analysis' results panel. The panel displays a table of eigenvalues and variances for the first seven principal components. The table has columns: PC, Eigenvalue, and % variance. The data is as follows:

| PC | Eigenvalue | % variance |
|----|------------|------------|
| 1  | 12.5128    | 75.279     |
| 2  | 2.23428    | 13.442     |
| 3  | 1.07552    | 6.4706     |
| 4  | 0.532507   | 3.2037     |
| 5  | 0.192862   | 1.1603     |
| 6  | 0.04096    | 0.24643    |
| 7  | 0.0329195  | 0.19805    |

The panel also includes a 'Matrix' dropdown set to 'Variance-covariance', 'Groups' set to 'Disregard', 'Missing values' set to 'Mean value imputation', 'Suppl. vars' set to 0, and 'Bootstrap N' set to 0. A 'Recompute' button is at the bottom right.

Figura 10. Panel con los resultados del análisis de componentes principales.

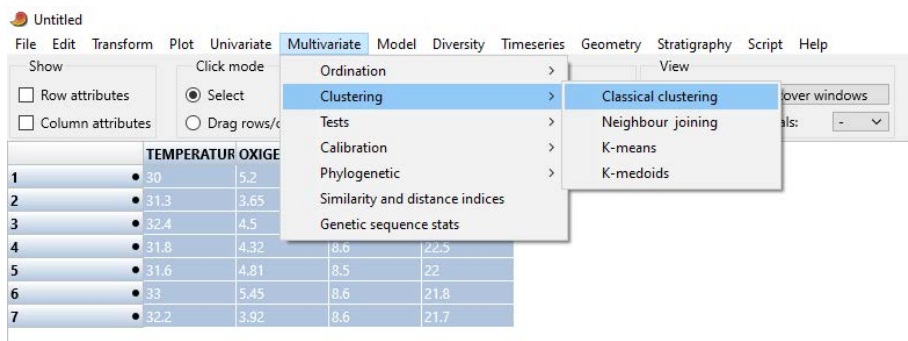


Figura 11. Menú multivariados, opción clúster: classical clustering.

Se utilizó el módulo Diversity (Figura 12) para analizar el comportamiento de los índices de diversidad biológica (métodos de diversidad alfa, los cuales se dividen en los que miden riqueza específica y la estructura de la comunidad), basados en Moreno (2001).

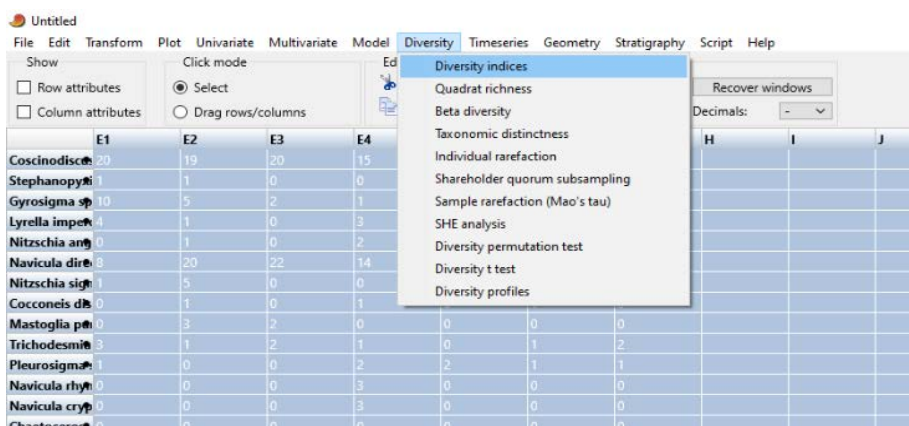


Figura 12. Módulo Diversity y sus opciones.

## Resultados y discusión

A continuación, se establece el análisis paso a paso de cada uno de los módulos señalados anteriormente, utilizando para ello los datos del monitoreo del riqueza y abundancia de fitoplancton y de los parámetros físico-químicos del agua de 7 estanques acuícolas. En este caso, para el primer muestreo de los 7 estanques, se aprecia que el valor de P de la prueba de Shapiro-Wilk es mayor al alfa (figura 13) y la regla de decisión indica que en ese caso si existe distribución normal de los datos.

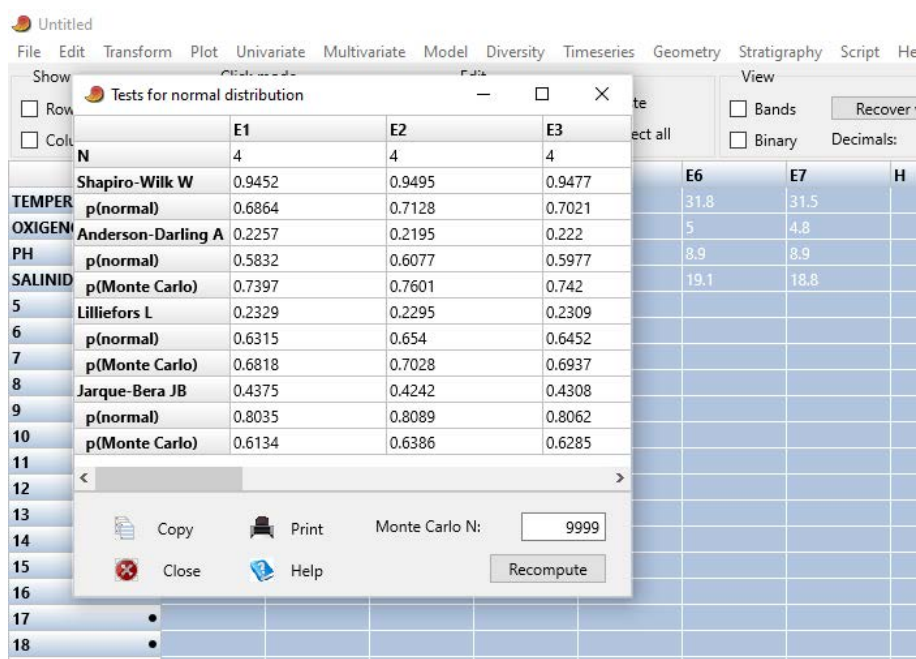


Figura 13. Se muestran los valores de P de la prueba de Shapiro-Wilk.

## Caracterización ambiental por muestreo

Para llevar a cabo esta caracterización, se realizó un análisis en la opción summary statistics del menú Univariado. Mostrando los resultados como se observan en la figura 14. Despliega el tamaño de datos analizados, los valores máximos y mínimos, la media, desviación estándar y error estándar, varianza, mediana, curtosis, entre otra información importante.

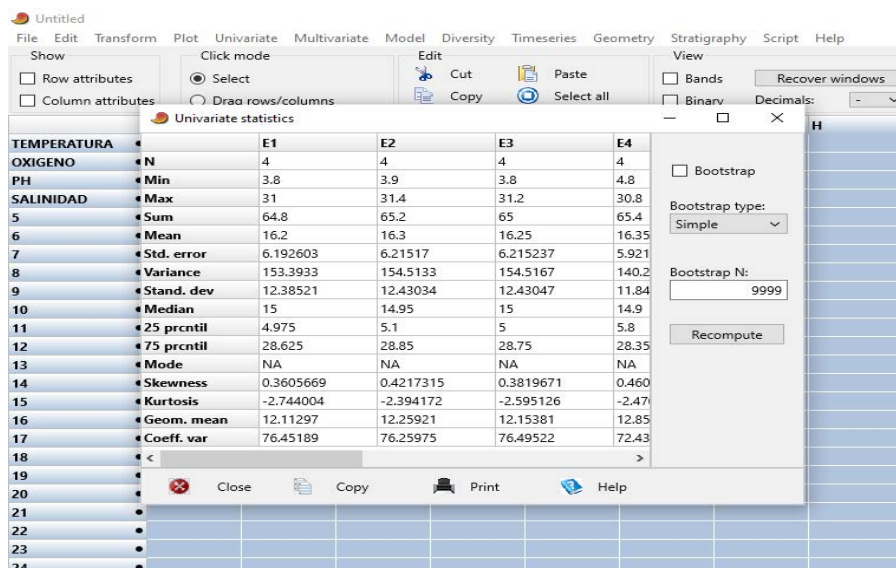


Figura 14. Resultados de la estadística descriptiva de los datos.

La temperatura tuvo un promedio de 31.77°C durante todo el estudio, con una mínima de 30°C en el estanque 2 del muestreo 1, mientras que la máxima fue de 33.3°C en el estanque 3 durante el muestreo 3. Al aumentar la temperatura, disminuye la cantidad de oxígeno disuelto en el agua. Cuando el agua contiene todo el oxígeno disuelto a una temperatura dada, se dice que está 100 por cien saturada de oxígeno. El agua puede estar sobresaturada de oxígeno bajo ciertas condiciones (Boyd, 1989). El oxígeno disuelto presentó un promedio de 4.55 mg/L en toda la investigación, con un valor mínimo de 3.7 mg/L en el estanque 4 durante el muestreo 7 y el valor máximo fue de 6.0 mg/L en el estanque 4 durante el muestreo 5, se indica que este parámetro ambiental puede variar mucho, ya sea subiendo por evaporación o bajar su cantidad con la lluvia, también dice que el rango esencial para la sobrevivencia del camarón, debe de ser entre 0 % hasta 50 % ppm donde el valor obtenido en el análisis realizado sobrepasa el rango de tolerancia del camarón (LABOMERSA, 2021). Los valores de pH muestran un promedio general de 8.76, con un valor mínimo de 8.5 registrado en los estanques 1, 2, 4 y 6, y el valor máximo fue de 9.1 encontrado en el estanque 3 durante los muestreos 1 y 5. La salinidad tuvo un promedio de 21.62 ‰, con un valor mínimo de 18.8 ‰ en los estanques 1 y 6 del muestreo 7, mientras que el valor máximo se presentó en el estanque 4 del muestreo 2. Según Gao (2022), en cuanto a la calidad del agua de los sistemas

de producción, los parámetros más importantes son la temperatura, el oxígeno disuelto, el pH y la salinidad. Los camarones no pueden regular su temperatura interna y el ambiente la determina, lo que afecta varios procesos fisiológicos, incluidas las tasas metabólicas y la alimentación (figura 15).

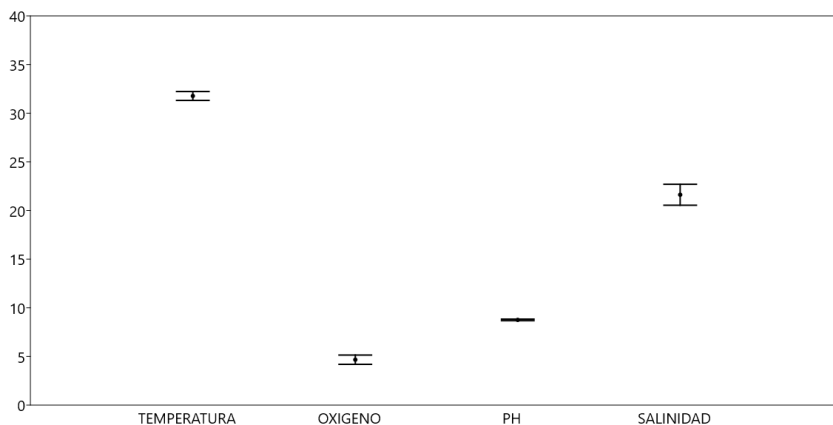


Figura 15. Box plot de los parámetros monitoreados.

La siguiente prueba realizada fue el ANOVA, la cual se aplicó para los valores de los resultados de los parámetros físico-químicos de cada estanque en cada muestreo, con la finalidad de buscar diferencias significativas entre estos valores. Aquí se utilizó el módulo Univariado en la opción de ANOVA (Figura 16), y enseguida muestra los resultados de dicha prueba a los valores de salinidad del agua en los 7 estanques durante el estudio (figura 17).

Al realizar una comparación mediante ANOVA con los valores de temperatura en los estanques, durante todos los muestreos se observa que no existieron diferencias significativas ( $P\text{value} = 0.5745 > 0.05$ ). Para el caso del oxígeno disuelto tampoco se observaron diferencias significativas ( $P\text{value} = 0.2724 > 0.05$ ). Con respecto al pH se obtuvo un  $P\text{value} = 0.7189 > 0.05$ , por lo que tampoco hubo diferencias significativas entre los estanques. La salinidad, por el contrario, mostró diferencias significativas entre los estanques, al obtenerse un  $P\text{value} = 0.01966 < 0.05$ .

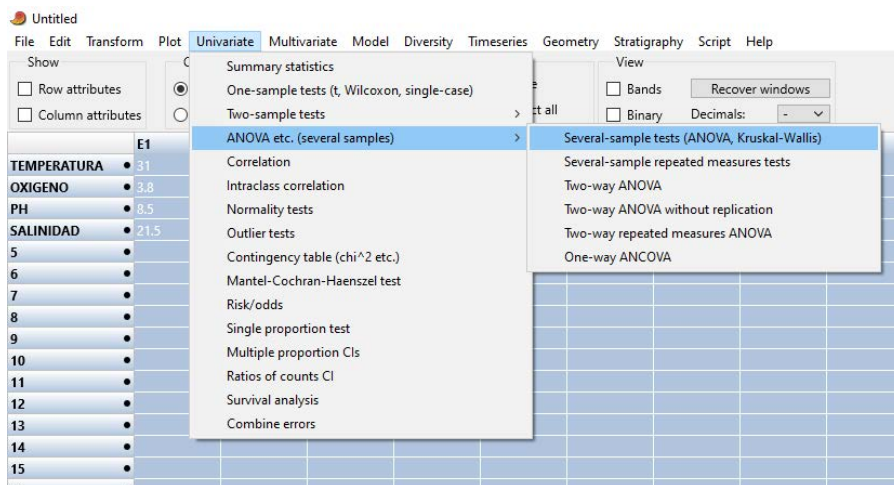


Figura 16. Menú desplegado del Análisis de Varianza.

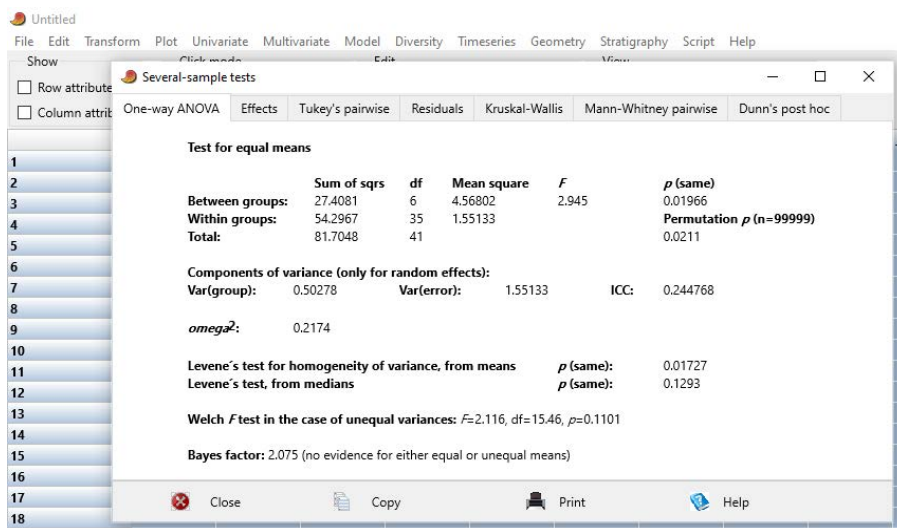


Figura 17. Resultados del ANOVA, con el valor P.

### Caracterización ambiental general

De acuerdo con los datos obtenidos a través de los diferentes monitoreos se realizaron promedios generales para la comparación del comportamiento de los parámetros ambientales en los meses de estudio. El valor máximo de temperatura lo obtuvo el mes de septiembre con 32.3°C, mientras que octubre fue el mes con menor registro con 31.5 °C. En cuanto al oxígeno el mes de septiembre contó con el registro más alto con 5.2142 mg/L, mientras que octubre con 4.3785 mg/L fue el registro más bajo. Por su parte el pH en septiembre tuvo el valor más alto con 8.85, mientras que agosto con 8.69 tuvo el registro más bajo. Finalizando con salinidad, con el mes de septiembre como el registro más alto con 22.7857 ‰ y el mes de octubre con el registro más bajo con 20.6642 ‰.

### Presencia y abundancia de fitoplancton

Con las especies identificadas se realizó un catálogo general relacionado a la presencia y distribución de especies fitoplanctónicas, donde destacan las especies *Coscinodiscus centralis*, *Gyrosigma sp*, *Navicula directa* y *Cocconeis sp 1*, las cuales se colectaron en la totalidad de estaciones, mientras que especies como *Stephanopyxis turris*, *Navicula cryptocephala*, *Navicula rhynchocephala*, *Fragilaria brevistriata*, *Bacillaria paxilifera*, *Prorocentrum gracile*, *Pleurosigma sp. 3*, *Lithodesmium undulatum* y *Gyrosigma sp.1*, fueron las menos predominantes, colectándose únicamente en una estación. En los sistemas de cultivo de camarón se encuentran microalgas provenientes de los cuerpos de agua de abastecimiento, ya sea un estero, laguna costera o del mar. La composición y la abundancia del fitoplancton del agua de abastecimiento de la granja es modificada en el ambiente de los estanques (Páez-Osuna, 2001).

De igual manera se realizó el análisis de la estadística descriptiva, siguiendo las indicaciones señaladas anteriormente en el software, para los datos de la abundancia de especies por estanques y por muestreo (figura 18). El estanque 1 presentó un total de 21 especies con 104 individuos, mientras que el estanque 2 tuvo 20 especies y 93 individuos. En el caso del estanque 3 tuvo el registro más bajo de especies con 14 solamente y 89 individuos, para el estanque 4 se registraron 23 especies y 87 individuos y el estanque 5 registró 15 especies con 87 individuos. Finalmente, el estanque 6 tuvo 24 especies y el valor más bajo de individuos con 64.

En la figura 19 se muestra la estadística descriptiva que se obtiene en el software para las especies de fitoplancton por muestreo, tras la identificación

y conteo de los ejemplares de fitoplancton, se evidencia una diversidad de 45 especies con una abundancia total de 597 ejemplares. Destacando una clara dominancia de la especie *Navicula directa*, con un total de 128 ejemplares, seguida de cerca por *Coscinodiscus centralis* con 115 ejemplares. Entre las especies menos abundantes, se encuentran *Thalassionema nitzschioides*, *Fragilaria brevistriata*, *Navicula regularis*, *Amphora proteus*, *Bacillaria paxillifera*, *Prorocentrum gracile*, *Pleurosigma sp. 3*, *Coscinodiscus centralis* y *Lithodesmium undulatum* con un solo ejemplar de cada uno. El fitoplancton de los sistemas acuícolas está formado por casi todos los grupos taxonómicos de algas, donde el grupo de las diatomeas es normalmente el grupo dominante de la comunidad fitoplanctónica en aguas templadas (Martínez-Córdova, et al, 2004,), lo cual coincide con lo encontrado en este estudio.

Para el mes de agosto se registró una riqueza específica de 34 especies, con una abundancia total de 189 individuos. Donde la especie con mayor abundancia fue *Navicula directa* con 56 individuos y *Trichodesmium erythraeum* con únicamente 2 ejemplares. En septiembre se registró una riqueza específica de 21 especies, con una abundancia total de 194 individuos. Donde la especie con mayor abundancia fue *Coscinodiscus centralis* con 56 individuos y *Climacophenia monilifera* con únicamente 1 ejemplar. Durante el mes de octubre se registró una riqueza específica de 25 especies, con una abundancia total de 198 individuos. Donde la especie con mayor abundancia fue *Cocconeis sp1* con 41 individuos y *Achnanthes fimbriata* con únicamente 1 ejemplar.

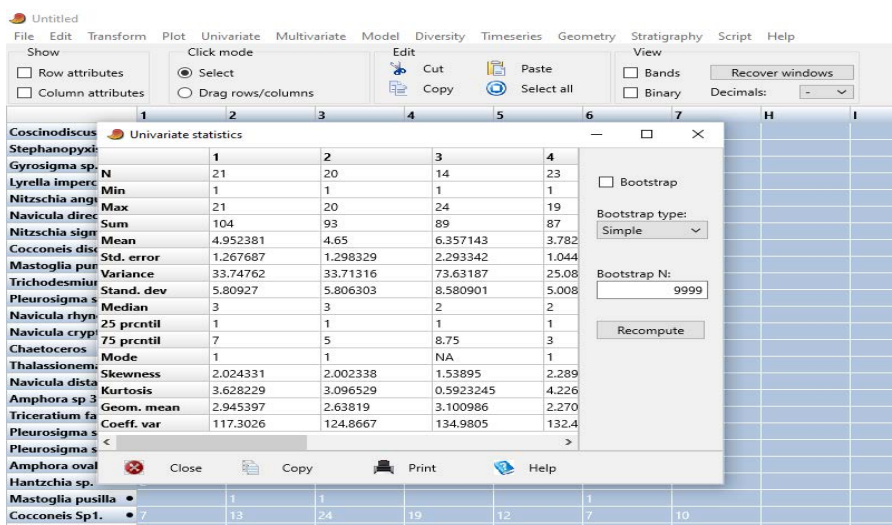


Figura 18. Resultados de estadística descriptiva por estanques.

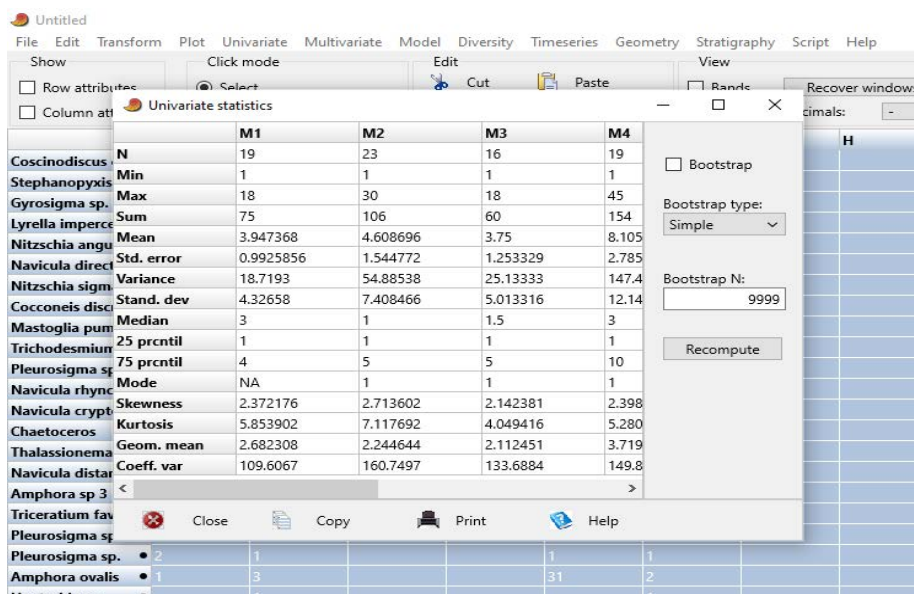


Figura 19. Resultados de estadística descriptiva para fitoplancton por muestreo.

## Análisis multivariado por muestreos

### Análisis de componentes principales

En este análisis se puede apreciar que la mayor explicación de la varianza del fenómeno se concentra en los componentes principales 1 y 2, con valores de 66.94 % y 20.47 % respectivamente. Los vectores representan la magnitud o fuerza de la variable muestreo. En la figura 20 se muestra a las especies de fitoplancton que estuvieron con mayor abundancia y asociadas a los diferentes estanques monitoreados. Por ejemplo, se aprecia a *Navicula directa* y *Coscinodiscus centralis* fuertemente ligadas al muestreo 4 y a *Amphora ovalis* al muestreo 5. Para el caso del análisis en referencia a los estanques, se aprecia que destacan tres especies de fitoplancton: *Navicula directa*, *Coscinodiscus centralis*, *Cocconeis sp1* y *Amphora ovalis* (Figura 21) y los primeros 2 componentes principales describen la mayor variabilidad del problema con 72.50 % y 13.18 % respectivamente. Las especies aquí identificadas coinciden con aquellas encontradas por Valenzuela-Sánchez, (2014) para zonas impactadas y no por actividades acuícolas en Sonora.

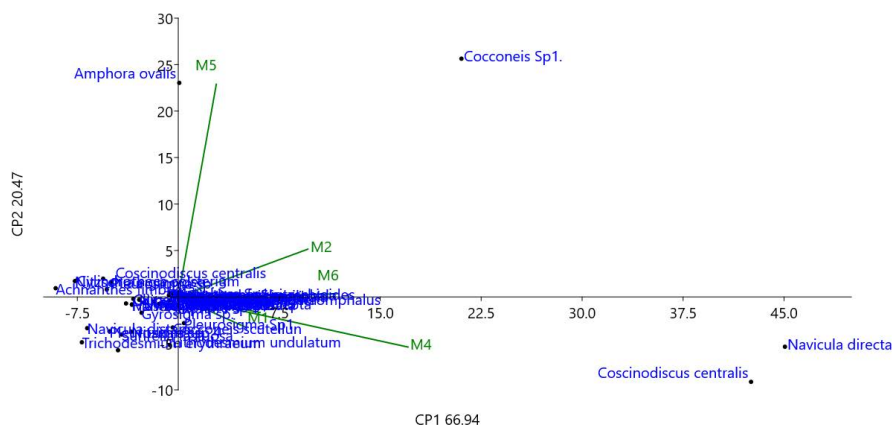


Figura 20. Análisis de componentes principales por muestreos.

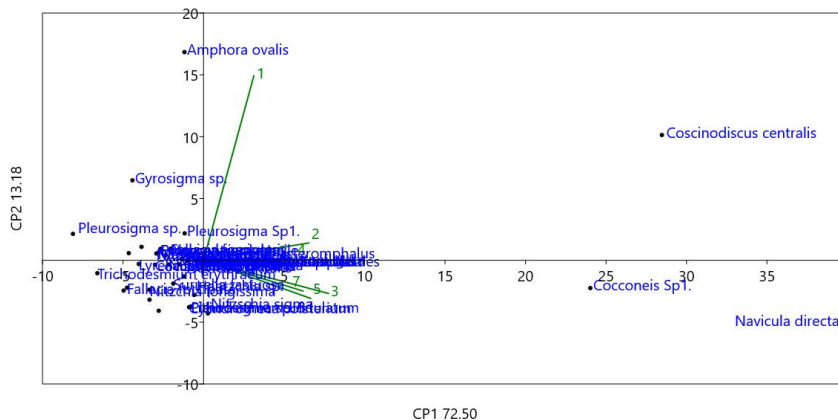


Figura 21. Análisis de componentes principales por estación.

## Análisis clúster

Al analizar los parámetros físico-químicos del muestreo 1, mediante el índice de similitud de Mahalanobis, se aprecia que se forman 2 conglomerados o clústeres, en uno de ellos presente solamente el estanque 1 (con una distancia

de separación de 1.55 del otro grupo), debido a sus condiciones de temperatura, oxígeno disuelto, pH y salinidad relativamente más altas que el resto de los estanques. En el segundo conglomerado se observan a los estanques 2 al 7 formando subconglomerados: el 7 separado en este segundo conglomerado por una distancia de 1.5, los estanques 2 y 4 con una similitud entre ellos de 0.7, y a su vez similares al 5 con una similitud de 1.1. Los estanques 3 y 6 similares entre ellos con 1.5 y separados del resto con una distancia de 1.31 (figura 22).

Para el muestreo 2 se observaron 2 conglomerados formados por el estanque 1 y el resto formando el otro grupo. Durante el muestreo 3 se formaron 2 conglomerados integrados por el estanque 7 en uno y el resto de las estaciones en el segundo conglomerado. Para el muestreo 4 también hubo 2 conglomerados siendo el 5 diferente del resto. En el quinto muestreo destacó de nuevo el estanque 1 separado del resto de los demás estanques y finalmente en el sexto muestreo el estanque 4 se separó de los demás estanques.

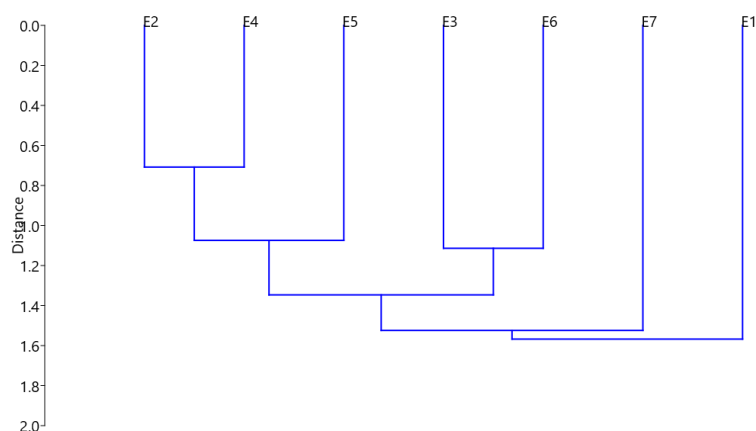


Figura 22. Análisis clúster para parámetros físico-químicos en muestreo 1.

## Comportamiento biológico

### Índices ecológicos por estación

Se observa que el índice de dominancia muestra valores bajos, con lo cual se interpreta que no hay una dominancia de un taxón en específico. El índice de Simpson por su parte, mide la igualdad en la comunidad presentando valores que fluctúan desde 0.81 hasta 0.90, lo que indica una uniformidad. Por otro lado, el índice de Shannon mide la diversidad y considera a individuos y taxa en cada estación, para este estudio los valores son superiores a 2 y menores a 3, por lo que se considera valor de diversidad media-baja. Con respecto al índice de Margalef se obtuvieron resultados entre 2.90 (baja diversidad) hasta 5.51 (diversidad media) (cuadro 1).

|                | E1    | E2    | E3    | E4    | E5    | E6    | E7    |
|----------------|-------|-------|-------|-------|-------|-------|-------|
| Taxa_S         | 21    | 20    | 14    | 23    | 15    | 24    | 15    |
| Individuals    | 104   | 93    | 89    | 87    | 87    | 65    | 70    |
| Dominance_D    | 0.10  | 0.11  | 0.18  | 0.11  | 0.15  | 0.11  | 0.19  |
| Simpson_1-D    | 0.90  | 0.89  | 0.82  | 0.89  | 0.85  | 0.89  | 0.81  |
| Shannon_H      | 2.64  | 2.55  | 2.04  | 2.70  | 2.22  | 2.80  | 2.11  |
| Evenness_e^H/S | 0.67  | 0.64  | 0.55  | 0.64  | 0.61  | 0.69  | 0.55  |
| Brillouin      | 2.27  | 2.17  | 1.76  | 2.24  | 1.91  | 2.21  | 1.75  |
| Menhinick      | 2.06  | 2.07  | 1.48  | 2.47  | 1.61  | 2.98  | 1.79  |
| Margalef       | 4.31  | 4.19  | 2.90  | 4.93  | 3.14  | 5.51  | 3.30  |
| Equitability_J | 0.87  | 0.85  | 0.77  | 0.86  | 0.82  | 0.88  | 0.78  |
| Fisher_alpha   | 7.93  | 7.82  | 4.67  | 10.20 | 5.23  | 13.75 | 5.86  |
| Berger-Parker  | 0.20  | 0.22  | 0.27  | 0.22  | 0.30  | 0.28  | 0.37  |
| Chao-1         | 26.20 | 55.61 | 15.19 | 28.93 | 22.41 | 49.85 | 24.20 |
| iChao-1        | 30.97 | 58.88 | 16.02 | 31.01 | 34.12 | 62.62 | 32.45 |
| ACE            | 27.83 | 28.87 | 16.57 | 29.62 | 21.38 | 58.88 | 35.37 |

Cuadro 1. Principales índices ecológicos.

## Conclusiones

Los métodos empleados para el análisis resultan muy eficientes para organizar la información. Los métodos univariados permitieron conocer los comportamientos de los parámetros a lo largo del tiempo y a través de las diferentes zonas de muestreo. Las variaciones en parámetros físico-químicos y biológicos se presentaron de manera similar en todos los estanques, presentando oscilaciones y variaciones en cuanto a temperatura, pH y oxígeno, variando los niveles de salinidad entre los estanques a lo largo del estudio.

El módulo de Diversidad permitió conocer que la riqueza y la abundancia por muestreo fue: 1 con 19 especies y 75 ejemplares; 2 con 23 y 106; 3 con 16 y 60; 4, 19 y 154; 5, 16 y 104; y 6, 19 especies y 98 ejemplares. Las especies más abundantes por estanque fueron: *Amphora ovalis* para el 1; *Navicula directa* para 2, 5, 6, y 7; *Cocconeis sp1* para 3 y 4.

En el método de componentes principales se aprecia que, las especies de fitoplancton estuvieron cambiando su nivel de importancia en el curso del tiempo y espacio. Se identificaron 47 especies con una abundancia de 597 ejemplares, destacando el estanque 4 con 154 individuos y la menor fue la estación 6 con 65 individuos. El grupo predominante fueron las diatomeas, con la especie *Navicula directa* con una abundancia de 134 ejemplares, los cuales estuvieron presentes en todos los muestreos.

## Referencias

- Boyd, C. E. (1989). Water quality management and aeration in shrimp farming. Fisheries and allied aquacultures department series No. 2. Alabama. Agricultural experiment station, Auburn University, Alabama, 70 p.
- Corral-Dávalos, L. (2019). Estadística y técnicas experimentales para la investigación biológica. Universidad Politécnica Salesiana. Cuenca, Ecuador. 522 p.
- Correa-Henao, M. (2021). Análisis de clúster automático. Tesis. Universidad Nacional de Colombia. 74 p.
- Díaz Monroy, L.G. (2007). Estadística multivariada: inferencia y métodos. Bogotá: Universidad Nacional de Colombia. Facultad de Ciencias. XVII, 595 p.
- FAO. (2004). El estado mundial de la agricultura y la alimentación. 241 p.
- Hammer, Ø., Harper, D.A.T., Ryan, P.D. (2001). PAST: Paleontological Statistics software package for education and data analysis. *Paleontologia Electronica*. 4(1): 9 pp.
- Labomersa (2021) Importancia de la Calidad del Agua en los cultivos de camarón, Labomersa. Available at: <https://labomersa.com/2021/05/04/importancia-de-la-calidad-del-agua-en-los-cultivos-de-camaron/> (Accessed: 14 December 2023).
- Martínez Córdova, L. R., Campaña Torres, A. y Martínez Porchas, M. (2004). Manejo de la Productividad Natural en el Cultivo del Camarón. In: Cruz Suárez, L.E., Ricque Marie, D., Nieto López, M.G., Villarreal, D., Scholz, U. y González, M. 2004. Avances en Nutrición Acuícola VII. Memorias del VII Simposium Internacional de Nutrición Acuícola. 16-19 noviembre, 2004. Hermosillo, Sonora, México.
- Mendenhall, W., R.J. Beaver y B. M. Beaver. (2006). Introducción a la probabilidad y estadística. Décima tercera edición. Cengage Learning Editores, S.A. de C.V., 780 pp.
- Milton, J.S. (2007). Estadística para biología y ciencias de la salud. McGraw-Hill Interamericana. Tercera edición ampliada. 744 p.
- Moreno, C. E. (2001). Métodos para medir la biodiversidad. M&T-Manuales y Tesis SEA, vol. 1. Zaragoza, 84 pp.
- Muñoz-Armayones, S. (2016) Técnicas multivariantes para el análisis de datos ómicos. Trabajo de fin de grado. Facultad de Matemáticas. Departamento de estadística e investigación operativa. Universidad de Sevilla. 123 p.
- Páez-Osuna, F., 2001. Camaronicultura y medio ambiente. Mazatlán, Sinaloa: Instituto Ciencias del Mar y Limnología, Programa Universitario de Alimentos. El colegio de Sinaloa.

- Pastor-Barriuso R. (2012). Bioestadística. Madrid: Centro Nacional de Epidemiología, Instituto de Salud Carlos III. 262 p.
- Pérez-López, C. (2004). Técnicas de análisis multivariante de datos. Aplicaciones con SPSS. Pearson Educación, Madrid. 665 p.
- Pérez-Tejada, H.E. (2008). Estadística para las ciencias sociales, del comportamiento y de la salud.
- Shukla, S. (2015). Normality Test. Conference paper. 8 p.
- Valenzuela-Sánchez, C.G. (2014). Composición y abundancia de fitoplancton en áreas impactadas y no impactadas por descargas de granjas camaronícolas en la costa del estado de Sonora. Tesis Profesional. Universidad de Sonora. 66 p.
- Veiga de Cabo, J., De la Fuente Díez, E. y Marta Zimmerman-Verdejo, M. (2008). Modelos de estudios en investigación aplicada: Conceptos y criterios para el diseño. Med. Segur Trab. Vol LIV N° 210: 81-88



# Capítulo 3

---

## **Análisis clúster o conglomerados: identificación de dos temporalidades ambientales en la zona marino-costera de Guasave**

*Graciano-Obeso Adalid  
Alzate-Espinoza Juan Héctor  
Bojórquez-Delgado Gilberto*

## Introducción

Los métodos multivariados se aplican en procesos medioambientales desde principios del siglo XX, pero han tenido una enorme difusión en los últimos años, debido a la gran cantidad de información acumulada en las bases de datos y al enorme progreso de la tecnología computacional que comenzó en la década de 1960 (Palacio et al., 2020). Específicamente para la conformación de distintos conglomerados o clústeres, se aplica la técnica estadística dentro del análisis multivariado llamada “análisis de clúster o conglomerados”, este método permite la conformación de grupos homogéneos hacia el interior y lo más heterogéneos entre los distintos grupos (Gutiérrez y Ciancio, 2023).

Cuando los investigadores están interesados en identificar grupos de individuos, la técnica más adecuada del análisis multivariado es el análisis de conglomerados o análisis de clúster, esta técnica se basa en el análisis y la interpretación de la asociación observada entre los individuos, de modo que el cálculo de su distancia o proximidad sirve para conformar grupos homogéneos en relación con las características seleccionadas que, a la vez, sean tan heterogéneos entre ellos como sea posible (Meneses, 2019; Rodríguez et al., 2019). En esta técnica se considera a cada unidad de análisis como un conglomerado, y posteriormente va uniendo los conglomerados de acuerdo con su homogeneidad hasta que queda un conglomerado, entendiendo por grupo homogéneo aquel cuyos miembros difieren significativamente de los de cualquier otro (Starstedt y Mooi, 2014; Tussel, 2023).

Por otro lado, esta técnica ha sido utilizada en diversas investigaciones en el mundo para agrupar elementos con características similares o dividir por zonas (Burbano et al., 2018; Price et al., 2006; Pérez et al., 2004). Además, se ha demostrado que el análisis jerárquico, facilita determinar el nivel de correlacionamiento entre las diferentes variables, permitiendo generar grupos de asociación, lo cual redundaría en una mejor segmentación de los elementos que influyen el aprovechamiento de los activos medioambientales (Blanco et al., 2017).

El Análisis Clúster (AC) es un método estándar del análisis multivariado que puede reducir una compleja cantidad de información en pequeños grupos o clústers, donde los miembros de cada uno de ellos comparten características similares (Lin y Chen, 2006). El AC se considera una técnica eminentemente exploratoria que no utiliza ningún tipo de modelo estadístico para llevar a cabo el proceso de clasificación (Hair et al., 1999) y, por ello, se le podría calificar como una técnica de aprendizaje no supervisado, es decir, una técnica muy

adecuada para extraer información de un conjunto de datos sin imponer restricciones previas en forma de modelos estadísticos (Peterson, 2002).

El AC tiene por objeto formar grupos o clústers homogéneos en función de las similitudes o similaridades entre ellos (Peña, 2002). Los grupos se forman de tal manera que cada objeto es parecido a los que hay dentro del clúster con respecto a algún criterio de selección predeterminado (Rao y Srinivas, 2006; Hair et al.). Las técnicas de agrupamiento en el AC se pueden clasificar en dos categorías: el clúster jerárquico y el no jerárquico. Los procedimientos jerárquicos consisten en la construcción de una estructura en forma de árbol.

Existen dos tipos de procedimientos de obtención de clústers jerárquicos: los de aglomeración y los divisivos. Dentro de los métodos jerárquicos aglomerativos se tienen: (i) método de encadenamiento simple, (ii) métodos de encadenamiento completo, (iii) método de encadenamiento medio, (iv) método de Ward, y (v) método del centroide (Hair et al.). Estos procedimientos difieren en la forma como se calcula la distancia entre los conglomerados, entre los que se encuentran la DEC, Manhattan, coeficiente de correlación de Pearson, Chebichev y Cosine.

El clúster por medio de técnicas no jerárquicas no requiere de procesos de construcción de árboles; en su lugar, asignan los objetos a clústers una vez que el número de grupos a formar esté especificado. Los procedimientos de aglomeración no jerárquicos se denominan frecuentemente agrupaciones de *k*-medias, *k*-medianas y *k*-modas. Una desventaja con respecto a la técnica jerárquica consiste en que debe conocerse apriori el número de clústers a obtener, lo que implica un grado de subjetividad en el proceso (Peterson, 2002).

A pesar de lo anterior, se considera un método dinámico en el sentido en que los objetos dentro de los clústers se pueden mover de un clúster a otro, minimizando la distancia entre objetos dentro de un mismo clúster (Rao y Srinivas, 2006). Pese a las ventajas del método de aglomeración no jerárquico, en este artículo se presenta la aplicación del método jerárquico dado el interés de no querer asignar a priori el número de grupos a formar. A continuación, se describen las técnicas empleadas en el análisis clúster y el método de similitud utilizado.

- Encadenamiento medio entre grupos. Mide la proximidad entre dos grupos calculando la media de las distancias entre objetos de ambos grupos o las medias de las similitudes entre objetos de ambos grupos. Algunos autores, como Hair et al., afirman que el método está sesgado a formar conglomerados con aproximadamente la misma varianza.

- Método de Ward. Este proceso de aglomeración tiene como objetivo establecer grupos de tal forma que la suma de los cuadrados de las desviaciones con respecto a la media de cada variable es mínima para todas las estaciones al mismo tiempo. Este procedimiento tiende a combinar los conglomerados con un número reducido de observaciones y a formar grupos con aproximadamente el mismo número de grupos (Rao y Srinivas, 2006).
- Método del centroide. En este método la distancia entre los grupos se define como la distancia entre sus centroides. El centroide de cada grupo o clúster es a su vez el promedio de las posiciones de todos los puntos dentro del clúster. En este método, cada vez que se agrupa a los individuos se calcula nuevamente el centroide; así el centroide cambia a medida que se fusionan los grupos (Hair et al.).
- La distancia euclídea al cuadrado (DEC). Ese cuadrado de la suma de las diferencias al cuadrado de dos elementos en la variable o variables consideradas (Castellarin et al., 2001).

Con el objetivo analizar con técnicas estadísticas multivariadas por conglomerados o cluster las temporalidades de la zona marino-costera adyacente al litoral del municipio de Guasave, y semejanzas de la variable de temperatura superficial del mar (tsm), así como identificar la influencia de las surgencias de la zona marino-costera de Guasave.

## Metodología

### *Área de estudio*

El golfo de California está ubicado en el noroeste de México y está delimitado por la península de Baja California y la costa continental de los estados de Sonora, Sinaloa y Nayarit. Se localiza entre los 20° y 32° de latitud norte y los 105.5° y 114.5° de longitud oeste en el Pacífico Oriental, y se encuentra orientado en una dirección Noroeste (NO) – Sureste (SE) (Fig. 1).

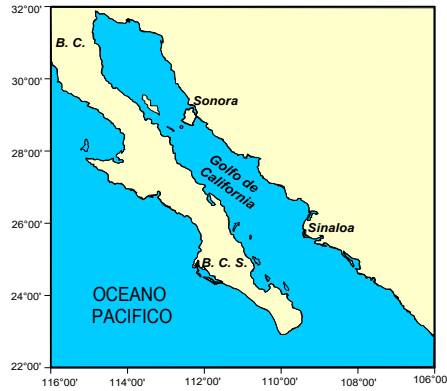


Figura 1. Ubicación geográfica del Golfo de California, México.

La presente investigación se realizó en la zona marino-costera adyacente al litoral del municipio de Guasave, Sinaloa, donde se realizaron ocho recorridos en una red de 15 estaciones ubicadas 50 km mar adentro (Fig. 2) durante febrero, abril, julio, octubre y diciembre de 2012 y febrero, marzo y abril de 2013. Para realizar los recorridos se consultaron las condiciones ambientales, específicamente del viento, el cual es un factor determinante para poder realizar los recorridos.

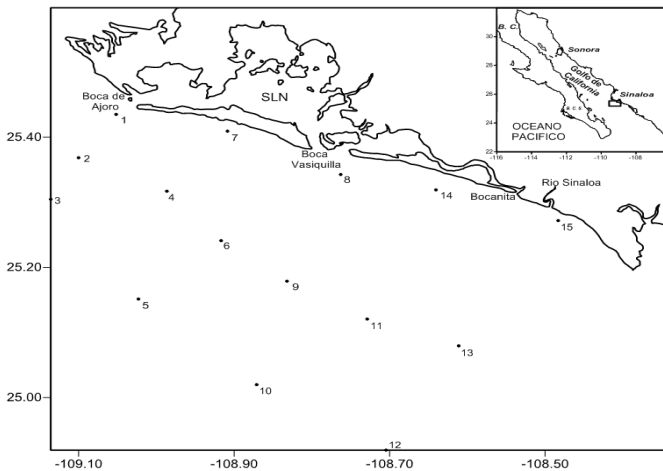
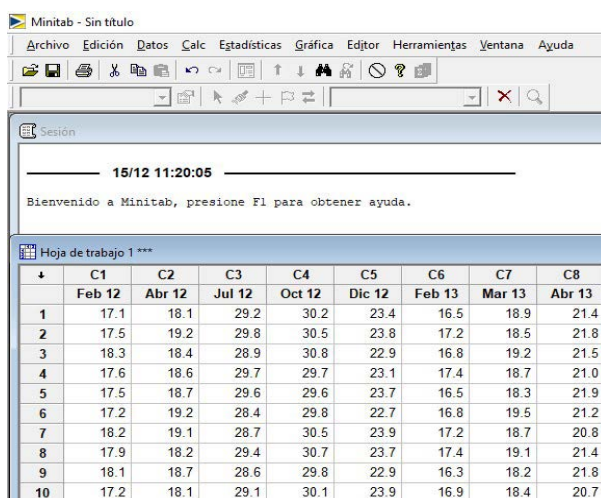


Figura 2. Delimitación del área de estudio y ubicación de las estaciones de muestreo en la zona marino-costera del municipio de Guasave, Sinaloa, México.

En cada estación se registró in situ la temperatura superficial del mar (tsm), donde se realizaron 10 mediciones por punto de muestreo.

## **Análisis clúster o de conglomerados con el Software estadístico Minitab versión 15**

*Ingresar los datos*



|    | C1     | C2     | C3     | C4     | C5     | C6     | C7     | C8     |
|----|--------|--------|--------|--------|--------|--------|--------|--------|
|    | Feb 12 | Abr 12 | Jul 12 | Oct 12 | Dic 12 | Feb 13 | Mar 13 | Abr 13 |
| 1  | 17.1   | 18.1   | 29.2   | 30.2   | 23.4   | 16.5   | 18.9   | 21.4   |
| 2  | 17.5   | 19.2   | 29.8   | 30.5   | 23.8   | 17.2   | 18.5   | 21.8   |
| 3  | 18.3   | 18.4   | 28.9   | 30.8   | 22.9   | 16.8   | 19.2   | 21.5   |
| 4  | 17.6   | 18.6   | 29.7   | 29.7   | 23.1   | 17.4   | 18.7   | 21.0   |
| 5  | 17.5   | 18.7   | 29.6   | 29.6   | 23.7   | 16.5   | 18.3   | 21.9   |
| 6  | 17.2   | 19.2   | 28.4   | 29.8   | 22.7   | 16.8   | 19.5   | 21.2   |
| 7  | 18.2   | 19.1   | 28.7   | 30.5   | 23.9   | 17.2   | 18.7   | 20.8   |
| 8  | 17.9   | 18.2   | 29.4   | 30.7   | 23.7   | 17.4   | 19.1   | 21.4   |
| 9  | 18.1   | 18.7   | 28.6   | 29.8   | 22.9   | 16.3   | 18.2   | 21.8   |
| 10 | 17.2   | 18.1   | 29.1   | 30.1   | 23.9   | 16.9   | 18.4   | 20.7   |

Figura 3. Captura de datos de las mediciones en software minitab.

Se ingresaron los datos de las mediciones en la hoja de trabajo de la pantalla principal del software, para este caso práctico fueron las temperaturas que se tomaron durante los meses de febrero, abril, julio, octubre y diciembre de 2012, así como los meses de febrero, marzo y abril del año 2013 (Fig. 3).

## Estandarizar los datos

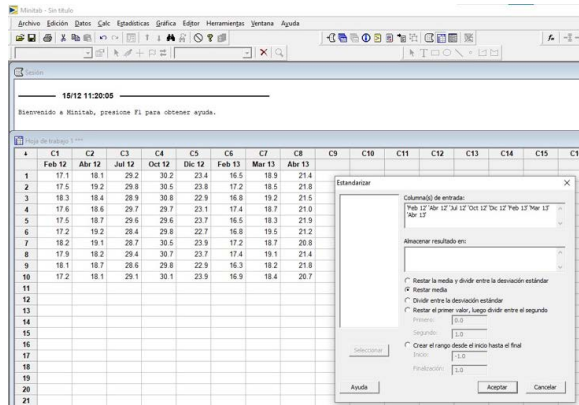


Figura 4. Estandarizar los datos de los resultados obtenidos en los puntos de muestreo.

Una vez que tenemos registrados los datos, se estandarizan los datos con el objetivo de que el software Minitab pondere todas las variables de igual manera, en la barra de herramientas seleccionamos Calc y posteriormente estandarizar (Fig. 4).

## Especificar el método de enlace

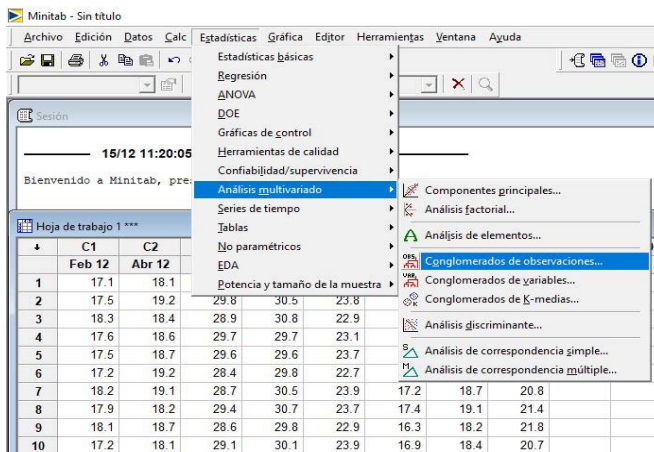


Figura 5. Elegir el método de enlace entre los conglomerados.

Una vez que se tienen los datos están estandarizados por el software, seleccionamos el método de enlace para especificar cómo se define la distancia entre los conglomerados, para este caso práctico elegimos el método completo, también conocido como el método del vecino más lejano (Fig. 5).

*Especificar el método de vinculación, medición de la distancia y mostrar dendograma*

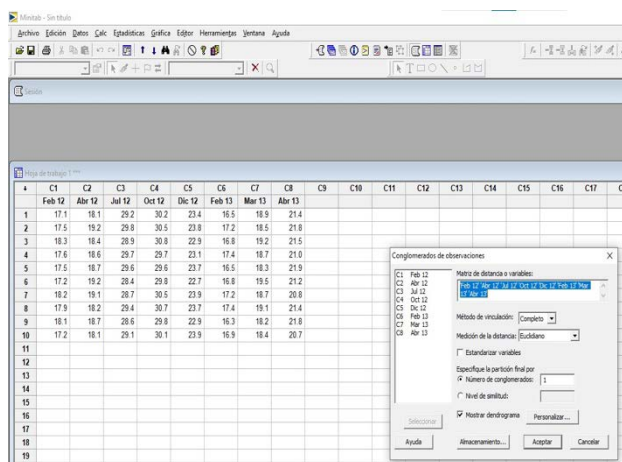


Figura 6. Seleccionar el método de medición de distancia entre los conglomerados.

Por último, ya que se tiene definido el método de enlace, es necesario seleccionar el método de la distancia, para el caso práctico de la presente investigación se eligió el método euclidiano, el cual es la medida de distancia más común utilizado para el análisis clúster o de conglomerados, mismo que calcula la raíz cuadrada de la suma de las diferencias al cuadrado. Una vez definido el método de enlace y la medición de la distancia, seleccionamos mostrar dendograma (Fig. 6).

### *Análisis de imágenes satelitales*

Con la finalidad de identificar las temporalidades con una mayor resolución, así como fenómenos oceanográficos que debido al diseño de los muestreos podrían no identificarse, como es el caso de los eventos de surgencias, se obtuvieron composiciones mensuales globales de la temperatura superficial del mar (tsm) del sensor ModisAqua disponibles en el sitio web <http://oceandata.sci.gsfc.nasa.gov/MODISA/Mapped/8Day/4km/SST/>, con una resolución espacial de 4x4 km por pixel (utilizando los compuestos de noche para evitar la contaminación por reflectancia). La serie de los compuestos obtenidos de tsm comprende el periodo de octubre de 2011 a enero de 2013. Para recortar el área de estudio se utilizó el paquete computacional MatLab (Matrix Laboratory).

### *Análisis estadístico*

En el presente trabajo los análisis estadísticos fueron realizados mediante el manejo del paquete Statistica v8.0. Se aplicó una prueba de normalidad y debido a que no mostraron una distribución normal, se realizó la transformación logarítmica de la base de datos para normalizar los datos y aplicarles el estadístico correspondiente. Con la finalidad de identificar las temporalidades del área de estudio, se aplicó un análisis de grupos (Clúster) utilizando el método de unión completa y distancias euclidianas a la tsm.

Con los datos obtenidos *in situ* de TSM, se realizaron mapas de distribución espacial por fecha de muestreo, así como por temporada fría y cálida en la zona marino-costera perteneciente al área de estudio con ayuda del programa de interpolación bidimensional Surfer® Ver. 9.0, por último, la representación temporal de variables en estudio se graficó con el programa SigmaPlot 10.0.

## **Resultados y discusión**

### *Identificación de temporalidades respecto a la temperatura superficial del mar (TSM)*

Durante el período de estudio, los valores de tsm fluctuaron entre 15.8 a 30.9 °C, el valor promedio mensual mínimo en febrero de 2013 (17.21°C) y el máximo en octubre de 2012 (29.76°C) (Tabla 1; Fig. 7).

|                   |          | Temporada fría | Temporada cálida |
|-------------------|----------|----------------|------------------|
| Temperatura<br>°C | Máxima   | 23.20          | 30.90            |
|                   | Mínima   | 15.80          | 27.60            |
|                   | Promedio | 19.20          | 29.40            |

Tabla 1. Valores promedio, máximos y mínimos de temperatura superficial del mar.

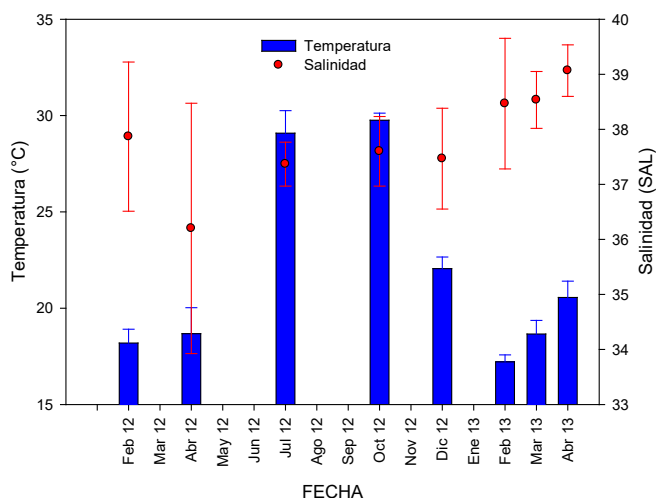


Figura 7, Distribución temporal de tsm (°C) y Salinidad (Sal), en la zona marino-costera del municipio de Guasave, Sinaloa, México, durante un periodo de febrero de 2012 a abril de 2013.

Del análisis de grupos (Clúster) aplicado a la tsm se identificaron 2 grupos a un nivel de corte de 10 unidades de distancia aplicado a la temperatura (Fig. 8), y se eligió el criterio de la temperatura para realizar el análisis de los resultados de las variables analizadas, debido a que separa claramente los resultados en dos temporadas climáticas: fría y cálida, correspondiendo a la primera el periodo de diciembre a abril con temperaturas de 19.20° C en promedio, y de junio a octubre a la cálida con un promedio de 29.40° C.

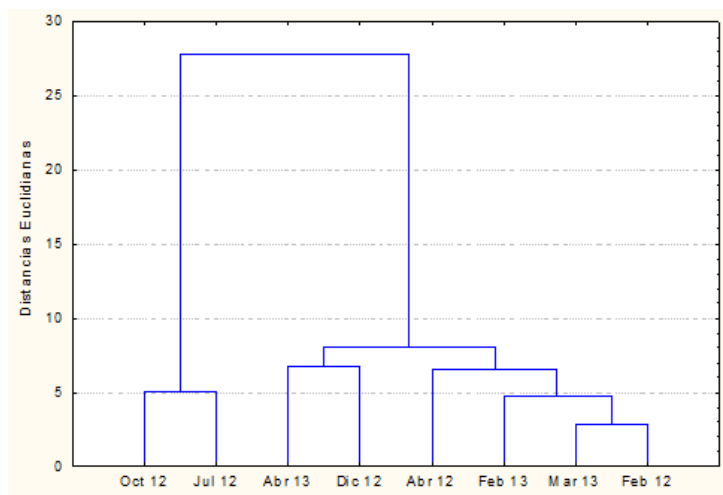


Figura. 8. Análisis de grupos de tsm en la zona marino-costera del municipio de Guasave, Sinaloa, México.

Respecto a la distribución espacial, los mayores valores de temperatura se asociaron a las estaciones más cercanas a la línea de costa, específicamente en las bocas del sistema lagunar y frente al Río Sinaloa (Fig. 9).

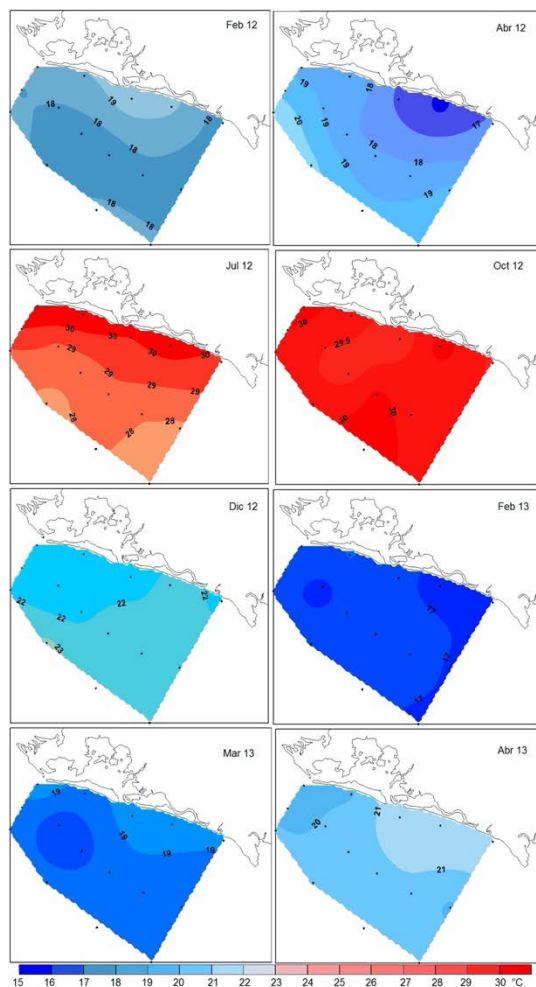


Figura 9. Distribución espacial de temperatura superficial del mar (°C) en la zona marino-costera del municipio de Guasave, Sinaloa, México, durante un periodo de febrero de 2012 a abril de 2013.

Al realizar el análisis por épocas climáticas, durante la temporada fría se obtuvieron temperaturas promedio de 15.80 a 23.20 °C. Espacialmente se observó una zona de menor temperatura hacia el norte del área de estudio en el transecto central respecto a la línea de costa (Fig. 10a). En la temporada cálida el rango promedio fue de 27.60 a 30.90 °C, con los valores menores en la zona oceánica y máximos en las estaciones más cercanas a la línea de costa (Fig. 10b).

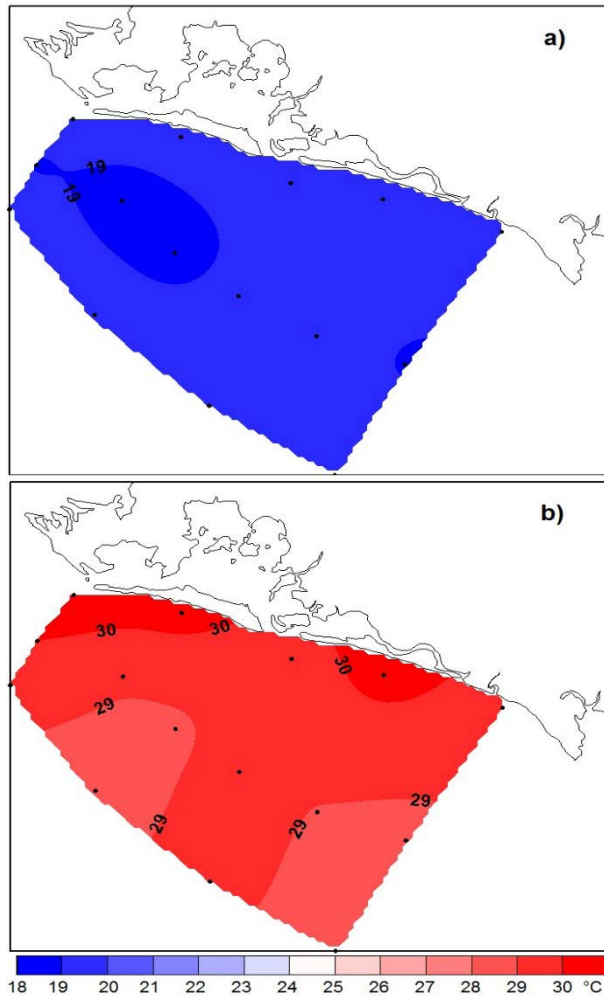


Figura 10. Distribución espacial promedio de tsm (°C) durante: a) temporada fría y b) temporada cálida en la zona marino-costera del municipio de Guasave, Sinaloa, México.

El patrón general de las temperaturas en el área de estudio presenta una clara influencia de la señal estacional en la que pueden agruparse en temporada fría y cálida, dicha agrupación fue posible con el análisis clúster o de conglomerados aplicado en el presente estudio, además, dicha estacionalidad se corrobora con las imágenes satelitales.

Durante la temporada fría, el promedio de temperaturas mostró un claro proceso de enfriamiento en la zona de estudio típico de los eventos de surgen-

cias costeras, con aguas más cálidas hacia la región sur del golfo de California (Fig. 11).



Figura 11. Distribución espacial de Temperatura Superficial del Mar (TSM) en temporada fría en la zona marino-costera del municipio de Guasave, Sinaloa, México.

Para la temporada cálida por otro lado, las mayores temperaturas se ubicaron en la zona costera, con un claro gradiente de mayores a menores temperaturas hacia la zona oceánica (Fig. 12).

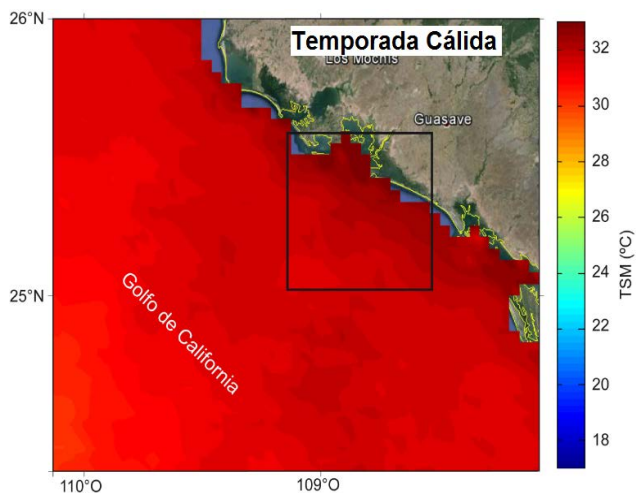


Figura 12. Distribución espacial de Temperatura Superficial del Mar (TSM) en temporada cálida en la zona marino-costera del municipio de Guasave, Sinaloa, México.

Durante la temporada fría, se registraron las menores temperaturas, con los valores mínimos ubicados en la zona norte del área de estudio, lo cual concuerda con lo reportado por Ulloa-Pérez (2005) y Valencia-Martínez (2012), quienes describen este mismo patrón de temperaturas. Esta distribución de temperaturas puede estar relacionado con la presencia de aguas oceánicas de menor temperatura que las costeras y que se internan en el Sistema Lagunar Navachiste durante el intercambio por mareas debido a fenómeno descrito para la zona por Escobedo-Urías, (2010).

Hay numerosos trabajos que analizan distintas variables del agua aplicando métodos de conglomerados, como es el caso de (Díaz y Mormeneos, 2002) quienes determinan que por la estructura de los datos en su investigación este método superó otros siete métodos jerárquicos o difusos por la capacidad de estudio aun con un grupo reducido de objetos. Robles (2018) aplico el método para el análisis de la Temperatura Superficial del Mar (TSM) en la zona costera de sonora para la caracterización de procesos físicos y biológicos, y puede ser afectada por procesos en escalas espacio temporales muy diversas.

Por otro lado, al contrastar los valores promedios de temperatura para el área con los trabajos anteriores, se observó que el promedio obtenido en este estudio (19.20 °C) es menor al reportado por Ulloa-Pérez (2005) quien obtuvo un promedio de 20°C, y que se debió a que en los trabajos mencionados tuvieron una cobertura espacial más reducida con estaciones costeras, debido a lo cual registraron aguas más cálidas que en este estudio. Sin embargo, al comparar este valor promedio con el obtenido por Valencia-Martínez (2012), (TSM: 18.45 °C), el promedio en este trabajo fue mayor, lo cual pudo deberse a que este monitoreo tuvo una cobertura temporal más amplia (febrero de 2012 a abril 2013) a la del trabajo citado (febrero-octubre 2012), además que el monitoreo de Valencia-Martínez (2012) se realizó en condiciones La Niña (Fig. 13), fenómeno oceanográfico caracterizado por presentar valores de temperatura más bajos que en condiciones normales (Hayward et al, 1999), por lo cual aunque también se obtuvieron datos en la misma época, al promediarlos con los valores medidos en 2013 cuando el efecto del evento había finalizado, se obtuvieron temperaturas mayores.

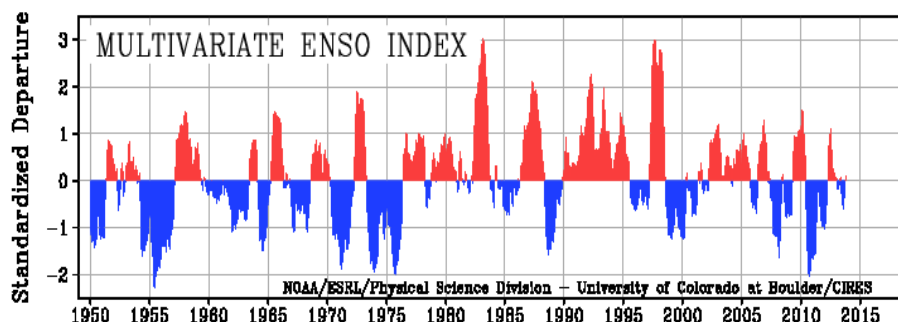


Figura 13. Índice multivariado del ENOS (El Niño Oscilación del Sur) para el periodo de 1950 a 2015.

## Conclusiones

- El análisis clúster o análisis de conglomerados es una técnica estadística multivariante que agrupa individuos con el fin de lograr la máxima homogeneidad en cada grupo, en la presente investigación fue posible identificar dos temporadas en el área de estudio de acuerdo a los resultados que arrojó el análisis clúster.
- Los diferentes patrones espaciales y temporales de la temperatura superficial del mar en el área de estudio, permiten identificar claramente dos temporadas climáticas: una fría durante los meses de diciembre a abril y una cálida de julio a octubre.
- Las imágenes de satélite de temperatura superficial del mar comprobaron la temporalidad de los eventos de surgencias costeras durante la temporada fría con inicio en diciembre y terminó en abril, la mayor intensidad fue en enero, febrero y marzo de 2012.

## Referencias

- Blanco, A., Muñoz, I., E., M., de Miguel, C., & Casas, E. (2017). Aplicaciones de la segmentación jerárquica en medición y evaluación de programas educativos. Ejemplos con un programa de educación financiera., 235-257.
- Burbano, V., Edy, L., & Moreno, E. (2018). Análisis de Conglomerados del Norte del valle del cauca. Caso estudio Cartago, Zarzal y la Unión. *Ingeniería Industrial*, 1, 78-91.
- Castellarin, A., Burn, D., & Brath, A. (2001). Assessing the effectiveness of hydrological similarity measures for flood frequency analysis. *Journal of Hydrology*, 241, 270-285.
- Escobedo\_Urías, D. (2010). Diagnóstico y descripción del proceso de eutrofización en lagunas costeras del Norte de Sinaloa. Tesis de Doctorado. CICIMAR-IPN, 298. La Paz, Baja California Sur, México.
- Gutiérrez, A., & Ciancio, A. (2023). Análisis del destino de la producción en la economía argentina: una aplicación de análisis de clúster. *Revista de la Facultad de Ciencias Económicas*, 20-40.
- Hair, J., Anderson, R., Tatham, R., & Black, W. (1999). Análisis multivariante. Madrid: Prentice Hall.
- Hayward, T., Baumgartne, T., C. D., R., D., G., G.-C., Hyrenbach, K., M., T. (1999). The State of the California Current, 1998-1999: transition to cool-water conditions. California Cooperative Oceanic Fisheries Investigations. 29-62: Report 40.
- Lin, G., & Chen, L. (2006). Identification of homogeneous regions for regional frequency analysis using the self-organizing map. *Journal of Hydrology*, 324, 1-9.
- Meneses, J. (2019). Introducción al análisis multivariante. Primera Edición. Universidad Oberta de Catalunya,. Barcelona, 39-43.
- Palacio, F., Apodaca, M., & Crisci, J. (2020). Análisis multivariado para datos biológicos: Teoría y su aplicación utilizando el lenguaje R. Fundación Natural de Félix de Azara, 268.
- Peña, D. (2002). Análisis de Datos Multivariantes. España: McGraw-Hill.
- Pérez, S., Montes, J., & Vázquez, J. (2004). Managing knowledge: the link between culture and organizational learning. *Journal of knowledge management*, 8, 93-104.
- Peterson, L. (2002). CLUSFAVOR 5.0: hierarchical cluster and principal-component analysis of microarray based transcriptional profiles. (B. C. Medicine, Ed.) Texas, One Baylor Plaza, ST-924, USA. Obtenido de <http://genome-biology.com/2002/3/7/software/0002>.

- Price, A., Patterson, N., Plenge, R., Meinblatt, M., Shadick, N., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38, 904.
- Rao, A., & Srinivas, V. (2006). Regionalization of watersheds by hybrid-cluster analysis. *Journal of Hydrology*, 318, 37-56.
- Robles, C. (2018). Variabilidad oceanográfica de la zona costera del estado de Sonora. (Tesis de Postgrado). Universidad de Sonora.
- Rodriguez, M., Comin, C., Casanoca, D., Bruno, O., Amancio, D., Costa, L., & Rodriguez, F. (2019). Clustering algorithms: A comparative approach. *PLoS ONE*. e0210236, 14(1).
- Starstedt, M., & Mooi, E. (2014). Cluster Analysis. In *a Concise Guide to Market Research*. Berlin, Heidelberg: Springer.
- Tussel, C. (2023). Calidad de la democracia en América Latina, 2013-2018: una clasificación con observaciones de conglomerado y dendrograma. *Estado & comunes, revista de políticas y problemas públicos.*, 2, 39-56.
- Ulloa-Pérez, A. (2005). Influencia de la disponibilidad de nutrientes sobre los cambios espacio-temporales de la comunidad de fitoplancton en el litoral del municipio de Guasave, Sinaloa. Tesis de Maestría, CIIDIR-IPN, 100. Unidad Sinaloa.
- Valencia-Martínez, S. (2012). Caracterización del área de alimentación de tortugas marinas en la zona marino-costera del complejo insular San Ignacio-Navachiste-Macapule, Sinaloa, Golfo de California. Tesis de Maestría, CIIDIR-IPN, 86. Unidad Sinaloa.

# Capítulo 4

---

## **Regresión lineal y análisis de la serie de tiempo mediante la codificación de datos**

*Lennin Enrique Amador-Castro  
Jesús Ramón Rodríguez-Apodaca  
Esther Graciela Lizárraga-Mata*

## I. Introducción

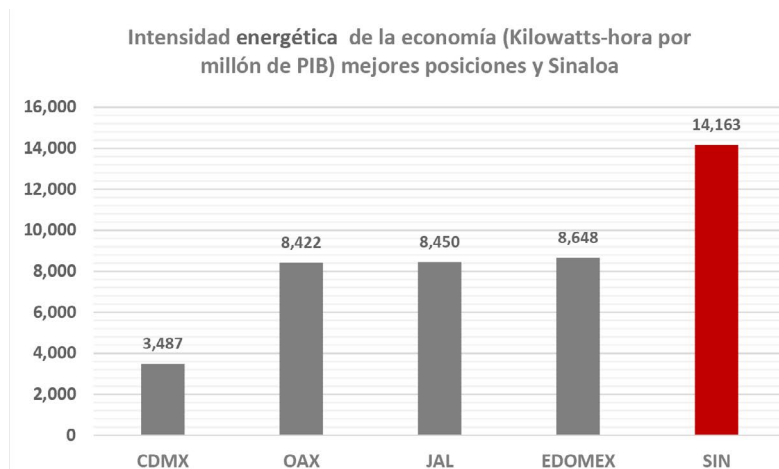
Los métodos cuantitativos son un conjunto de técnicas y herramientas de investigación basadas en cifras que permiten examinar y verificar los datos de estudio, con el fin de establecer la relación existente entre las variables que intervienen en la investigación, su importancia en los datos históricos, así como también determinar la dependencia entre las variables cuya particularidad sea representable por algún modelo matemático. La utilización de los métodos cuantitativos permitirá estimar el comportamiento de los datos que interactúan en el problema de estudio, por lo que los métodos más utilizados se encuentran los modelos matemáticos y los modelos causales.

En cuanto a los modelos matemáticos, estos son técnicas que permiten que todas las conductas o alternativas se puedan representar por medio de ecuaciones matemáticas cuyas variables están previamente establecidas de acuerdo con la investigación que se pretende realizar, los cuales a su vez existen tres modelos matemáticos empleados de manera más frecuente en el análisis para el pronóstico de ventas: método de incrementos absolutos, método de incremento porcentual y método de tendencia (Moreno, 2019). Por otro lado, los modelos causales son técnicas cuantitativas que utilizan información histórica para establecer pronósticos futuros. En su aplicación se requiere cumplir con tres requisitos: es necesario disponer de información del pasado, que la información que se obtenga sea cuantitativa y que el comportamiento de estos valores se repita en el futuro.

Es por esto, que ante la creciente necesidad de contar con información estadística de calidad y comparable sobre la situación ambiental, los recursos naturales y sus principales tendencias y amenazas de la región, así como de los países de respaldar con evidencias las políticas ambientales y de cambio climático, la producción de trabajos de estadísticas e indicadores comenzaron a desarrollarse a partir del año 2000 en la Comisión Económica para América Latina y el Caribe (CEPAL, 2021). A partir de esta idea, el presente capítulo realiza un análisis descriptivo e inferencial de las estadísticas ambientales con el fin de describir la situación y tendencia en materia de consumo energía. Por lo tanto, realizar una estimación de las variables permitirá establecer un comportamiento estacional de las mismas. Sin embargo, aunque existen fenómenos

como los meteorológicos que cambian constantemente, es posible aproximar su impacto.

Gráfica 1. Intensidad energética de la economía (Kilowatts-hora por millón de PIB) mejores posiciones y Sinaloa.



Fuente: (Gobierno del Estado de Sinaloa, 2022).

En el rubro energético, Gobierno del Estado (2022) en su Plan Estatal de Desarrollo (PED) menciona que Sinaloa se encuentra en el lugar 28 a escala federal en cuanto a intensidad energética de su economía, lo que significa que es de los estados que mayor energía necesitan para generar valor económico (Gráfica 1). Así mismo, es el segundo estado con mayor consumo de energía por hogar en el país, solo por debajo del estado de Sonora, y con base a datos de la Comisión Nacional para el Uso Eficiente de la Energía (CONUEE), Sinaloa tiene la tasa de mayor uso de aires acondicionados en el país con un 80 %. Esto es debido a las condiciones climáticas adversas y a las altas temperaturas que afectan a la población, generando una mayor demanda de energía e incrementando los costos por facturación de la energía eléctrica, elevando la contaminación del aire, los gases de efecto invernadero (GEI) y en algunos casos produciendo enfermedades y mortandad por el efecto de las islas de calor (Gobierno del Estado de Sinaloa, 2022).

En resumen, la mayoría de los municipios de Sinaloa se encuentran entre un intervalo mediano-alto a las afectaciones del cambio climático, por lo tanto, es necesario transitar en una economía que sustente el cuidado ambiental que ayude a mitigar los efectos adversos del cambio climático, pues de no hacerlo las consecuencias serán irreversibles (Gobierno del Estado de Sinaloa, 2018).

Por lo anterior, es necesario establecer estrategias que contrarresten los efectos adversos del cambio climático que permitan reducir los daños económicos y ambientales ocasionados por los diferentes meteoros que impactan al estado. Por lo tanto, una contribución directa será conocer el comportamiento en la demanda de energía y los factores que influyen en ella con el fin de disminuir las emisiones de GEI, estableciendo estadísticas energéticas para proporcionar información sustentada en el análisis y lo más fiable posible sobre la forma en que se consume la energía (Amador, Parra, y Rodríguez, 2022).

Entonces, el uso de los métodos cuantitativos permitirá predecir el comportamiento de un grupo de datos con resultados que pueden ser generalizados con el resto de la población, por lo que en este capítulo se desarrollarán los pasos para llevar a cabo un análisis de regresión lineal simple, así como también el análisis de series de tiempo (AST) de datos estadísticos relacionados a la demanda de energía.

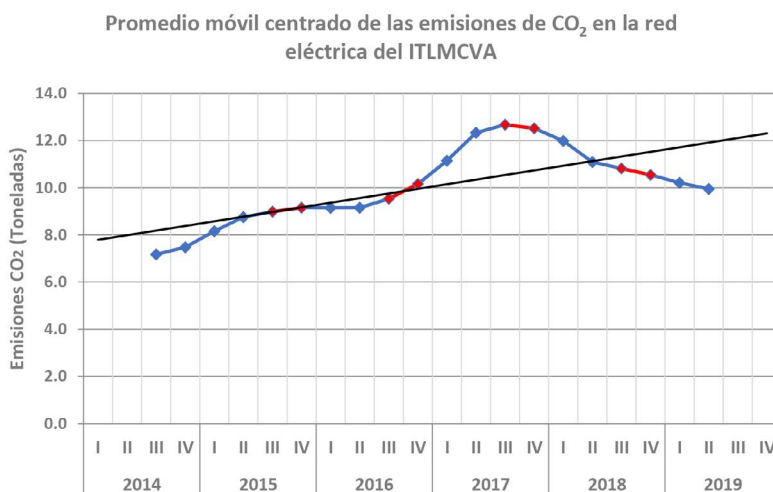
## **1.1. Antecedentes**

La necesidad por determinar las emisiones de GEI no solo se debe a los múltiples efectos originados en el medio ambiente, sino a la oportunidad que ofrece a todas las organizaciones el ahorro financiero y la competitividad. Los estudios de inventario de GEI en instituciones educativas de diferentes países han logrado identificar los impactos relacionados con el consumo de energía y recursos, factores que contribuyen al calentamiento global, mostrando en ciertos casos un bajo compromiso institucional (Amador, Parra, y Rodríguez, 2022).

En estudios realizados para el cálculo de GEI se tienen trabajos en instituciones de educación superior, donde se analizan las emisiones de CO<sub>2</sub> derivadas del consumo de energía eléctrica en el Instituto Tecnológico de Los Mochis Campus Villa de Ahome (ITLMCVA) (Amador, Parra, y Castro, 2021).

La modelación del comportamiento estacional de la serie de tiempo de la gráfica 2 establece las bases para predecir las estimaciones a futuro, determinando un modelo representativo que ayude a establecer acciones que contrarresten los efectos dañinos de los gases de efecto invernadero.

Gráfica 2. Promedio móvil centrado de las emisiones de CO<sub>2</sub> en la red eléctrica del ITLMCVA.



Fuente: (Amador, Parra, y Castro, 2021)

## 1.2. Justificación

La necesidad de encontrar un modelo que relacione los valores de un conjunto de variables, reconociendo que no es posible encontrar un método que permita pronosticar exactamente los resultados, ya que, en la mayoría de los casos, existe la posibilidad de un error en la estimación, porque los procedimientos matemáticos realizan una síntesis de la información, el presente trabajo describe el modelo matemático y los pasos necesarios para llevar a cabo un análisis de regresión lineal simple, basado en una serie de tiempo con el propósito de explicar el comportamiento de una variable en su relación con un grupo de variables que son conocidas y fáciles de medir. En este contexto, esto permitirá realizar un análisis histórico de los datos, visualizando los patrones de cambio de estos, permitiendo replicarlo en cualquier problema de aplicación.

## 1.3. Objetivo general

Aplicar el modelo de regresión simple y el análisis de series de tiempo en el pronóstico de los datos de estudio, con el fin de identificar el comportamiento de los patrones de cambio de las variables involucradas.

### **1.4. Objetivos específicos**

- Determinar los pasos que comprende aplicar el análisis de la regresión simple a través del método de los mínimos cuadrados.
- Identificar los componentes de una serie de tiempo para determinar los promedios móviles en los datos históricos de estudio.
- Aplicar en la tendencia secular una función lineal o curvilínea, realizando ajustes cíclicos y estacionales en los pronósticos.  
Analizar la línea de tendencia y la serie de tiempo utilizando Excel.

## **II. Materiales y métodos**

En este capítulo se presentan las técnicas utilizadas para realizar un análisis descriptivo de la demanda de energía eléctrica, estudiar datos históricos y predecir tendencias futuras, contribuyendo a la reducción de las emisiones indirectas de CO<sub>2</sub> con el fin de mantener la sostenibilidad ambiental y la mejora de la eficiencia energética.

El modelo de predicción utilizado para el análisis descriptivo en la demanda de energía eléctrica estacional fue la serie de tiempo para problemas de aplicación, basados en los métodos de mínimos cuadrados para la obtención de la línea de tendencia de mejor ajuste a través de la regresión lineal simple. Se determinó el error estándar de la estimación (, el coeficiente de determinación ( y el coeficiente de correlación , con el fin de estimar la variabilidad de los valores observados alrededor de la recta de regresión, así como también describir el grado en que una variable está linealmente relacionada con otra (Chapra y Canale, 2015).

### **2.1. Estimación mediante la recta de regresión lineal simple**

Los modelos de regresión son técnicas estadísticas que investigan la relación histórica entre una o más variables para establecer pronósticos futuros. Se trata de encontrar un modelo que relacione los valores de un conjunto de variables, reconociendo que no es posible encontrar un método que permita pronosticar exactamente los resultados, que siempre existe un error en la estimación, porque el procedimiento matemático usado es una síntesis de la información que no incluye todos los factores que influyen en el comportamiento real de la variable (Moreno, 2019).

El análisis de regresión se clasifica de acuerdo con el número de variables independientes a analizar. En el caso de la regresión simple, corresponde cuando la variable dependiente dependen únicamente de una sola variable independiente, sin embargo, cuando la variable depende de muchas variables independientes el análisis se denomina regresión múltiple.

Por lo tanto, el análisis de regresión simple es una técnica que se utiliza para realizar la “curva” (función lineal) que mejor se ajuste a los datos históricos observados. Su aplicación es análoga al modelo de tendencia descrito, y su diferencia radica en las variables examinadas, por lo que el modelo de tendencia considera como variable independiente el tiempo, y cuando la variable dependiente es la demanda de energía se basa en la hipótesis de los efectos compensatorios y admite que el comportamiento de las variables que afectan a la demanda cambia en el periodo proyectado, reconociendo que el cambio en promedio afectará la demanda en forma similar a como lo ha hecho históricamente.

### 2.1.1. Ajustes de la línea de tendencia con el método de mínimos cuadrados

Para determinar la línea de regresión se emplea una ecuación que relaciona las dos variables matemáticamente. La ecuación para una línea recta donde los valores estimados están en función con la variable independiente es:

$$\hat{Y} = a + bX \quad (\text{ec. 1})$$

$X$  = Valores de la variable independiente.

$a$  = Ordenada Y.

$\hat{Y}$  = Valores estimados de la regresión.

$b$  = Pendiente de la recta.

Sin embargo, para estimar la línea de tendencia, se ajusta la recta con el método de mínimos cuadrados de acuerdo con las siguientes ecuaciones (Levin y Rubin, 2004):

$$a = \bar{Y} - b\bar{X} \quad (\text{ec. 2})$$

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2} \quad (\text{ec. 3})$$

Donde:

$\bar{X}$  = Media de los valores de la variable independiente.

$n$  = Número de datos.

$\bar{Y}$  = Media de los valores de la variable dependiente.

$\sum X^2$  = Sumatoria de la variable  $X^2$ .

$\sum XY$  = Suma de productos de las variables  $X$  y  $Y$ .

Por otro lado, se ha desarrollado un parámetro de medición llamado error estándar de la estimación para medir la confiabilidad de la ecuación de estimación. Este parámetro determina la variabilidad, o dispersión, de los valores observados alrededor de la recta de regresión, el cual se puede calcular por medio de la siguiente expresión (Chapra y Canale, 2015):

$$s_e = \sqrt{\frac{\sum Y^2 - a \sum Y - b \sum XY}{n - 2}} \quad (\text{ec. 4})$$

Donde:

$X$  = Valores de la variable independiente.

$a$  = Ordenada Y de la ec. (2).

$Y$  = Valores de la variable dependiente.

$b$  = Pendiente de la recta de la ec. (3).

$n$  = Número de datos.

Es importante establecer que el coeficiente de determinación es la principal forma en que se puede medir el grado, o fuerza, de la asociación que existe entre dos variables,  $X$  y  $Y$  (Levin y Rubin, 2004). Debido a que la muestra de puntos es empleada para desarrollar rectas de regresión, normalmente, esta medida se le conoce como el coeficiente de determinación muestral .

$$r^2 = \frac{a \sum Y + b \sum XY - n \bar{Y}^2}{\sum Y^2 - n \bar{Y}^2} \quad (\text{ec. 5})$$

Por último, el coeficiente de correlación es la segunda medida que se puede utilizar para describir qué tan bien explica una variable a otra, también llamado coeficiente de correlación de la muestra y se determina por la raíz cuadrada del coeficiente de determinación de muestra (Levin y Rubin, 2004):

$$r = \sqrt{\frac{a \sum Y + b \sum XY - n \bar{Y}^2}{\sum Y^2 - n \bar{Y}^2}} \quad (\text{ec. 6}) \quad r = \sqrt{r^2} \quad (\text{ec. 7})$$

Cuando la pendiente de la ecuación de estimación es positiva,  $r$  es la raíz cuadrada positiva, pero si  $b$  es negativa,  $r$  es la raíz cuadrada negativa. Entonces, el signo de  $r$  indica la dirección de la relación entre las dos variables  $X$  y  $Y$ . Si

existe una relación inversa, es decir, si  $Y$  disminuye al aumentar  $X$ , entonces  $r$  caerá entre 0 y -1. De manera similar, si  $Y$  aumenta al incrementar  $X$ , entonces  $r$  será un valor en el intervalo de 0 a 1.

## 2.2 Análisis de Series de Tiempo (AST)

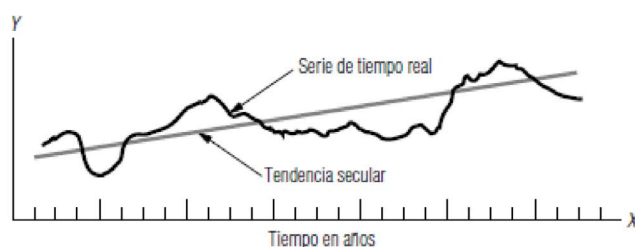
Con el fin de medir la estacionalidad, es decir, la periodicidad de los datos a lo largo de la serie de tiempo es necesario la eliminación del componente estacional (datos desestacionalizados), el cual permite modelar la variación cíclica a lo largo de la línea de tendencia (Amador, Parra y Castro, 2021).

El análisis de la serie de tiempo es utilizado para la visualización de patrones de cambio en intervalos regulares de una información estadística, permitiendo su proyección en la estimación de los datos en el futuro. Este método cuantitativo determinará el comportamiento de los patrones de los datos recolectados en la investigación a través del tiempo.

De acuerdo con Levin y Rubin (2004) el término serie de tiempo es utilizado para hacer referencia a cualquier tipo de información estadística la cual es acumulado en intervalos de tiempos, por lo que existen cuatro tipos de variaciones implicados en el análisis de la serie de tiempo:

a). *Tendencia secular*: En este caso, el valor de la variable tiende a aumentar o disminuir en un periodo de tiempo muy largo.

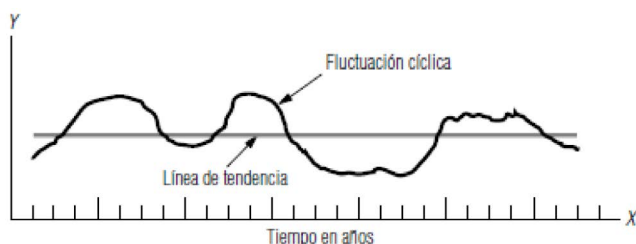
Figura 1. Tendencia secular en una serie de tiempo.



Fuente: (Levin y Rubin, 2004).

b). *Fluctuación cíclica*: En este tipo de análisis el ciclo de variación es común encontrarla en un pico arriba y abajo de la línea de tendencia en un periodo de tiempo sin seguir ningún patrón regular, sino más bien actúa de una manera impredecible.

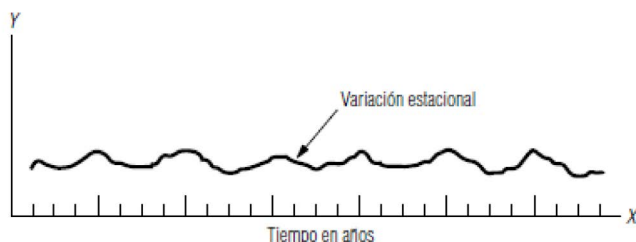
Figura 2. Fluctuación cíclica en una serie de tiempo.



Fuente: (Levin y Rubin, 2004).

c). *Variación estacional*: Implica patrones de cambio en un periodo de un año los cuales tienden a repetirse anualmente, tal como se muestra en la Figura 3. Permite comparar dos intervalos de tiempo que de otro modo serían bastante diferentes, así mismo, ayuda a la toma de decisiones para la proyección de patrones, por lo que una vez establecido el comportamiento estacional permitirá calcular la variación cíclica que se lleva cada año (Levin y Rubin, 2004).

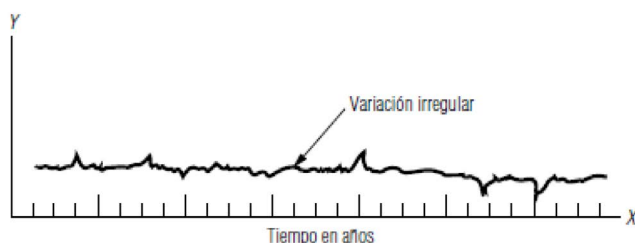
Figura 3. Variación estacional en una serie de tiempo.



Fuente: (Levin y Rubin, 2004).

d). *Variación irregular*: En este análisis el valor de una variable puede ser completamente impredecible cambiando de manera aleatoria, describiendo variaciones irregulares.

Figura 4. Variación irregular en una serie de tiempo.



Fuente: (Levin y Rubin, 2004).

### 2.2.1. Método de promedios móviles y codificación de datos

Con el fin de ajustar los datos por mes y/o trimestres es utilizado el método de promedio móvil, el cual proporciona un parámetro fundamental para describir el grado de la variación estacional. Este dato llamado índice estacional representa el porcentaje con el grado de estacionalidad medido por las variaciones respecto a la base, que de acuerdo con Levin y Rubin, (2004) se toman en cuenta los siguientes pasos para su cálculo:

- Se determina el total móvil de los cuatro trimestres para la serie de tiempo sumando el total de los valores para los trimestres en el primer año asociándolo con el dato que ocupa el valor medio del conjunto de valores del cual fue calculado.
- Se determina el promedio móvil de los cuatro trimestres dividiendo entre cuatro cada uno de los totales.
- Se centra el promedio móvil de los cuatro trimestres asociando a cada trimestre el promedio de los dos promedios móviles de cuatro trimestres que caen justo arriba y abajo de este.
- Se calcula el porcentaje del valor real con respecto al valor del promedio móvil para cada trimestre de la serie de tiempo que tenga un elemento de promedio móvil de 4 trimestres, permitiendo recuperar la componente estacional para los trimestres. Este porcentaje se determina dividiendo cada uno de los valores trimestrales reales entre los valores correspondientes del promedio móvil centrado para ser multiplicado por 100.

$$\frac{\text{Real}}{\text{Promedio móvil}} \times 100 \quad (\text{ec. 8})$$

- e). Calcular la media modificada para cada trimestre descartando los valores más alto y bajo de cada trimestre para promediar los valores restantes, con esto se reducen las variaciones cíclicas e irregulares extremas.
- f). Finalmente se hace un ligero ajuste de la media modificada, por lo que para corregir esto, se multiplica cada uno de los índices trimestrales por su constante de ajuste correspondiente.

Dado que la variable independiente tiempo es expresada en términos de semanas, meses o años para este tipo de análisis, a menudo se convierten estas medidas tradicionales de tiempo a una forma que simplifica los cálculos, la cual se le llama codificación de los datos. Para utilizar la codificación en este caso, se determina el tiempo medio y luego restamos ese valor de cada uno de los tiempos de la muestra.

Tabla 1. Codificación de valores en el tiempo.

|      | Trimestre | Deman-<br>da<br>energía<br>(kWh) | Tiempo<br>c o d i -<br>ficado<br>(x) |      | Trimes-<br>tre | Deman-<br>da<br>energía<br>(kWh) | Tiempo<br>codifi-<br>cado (x) |
|------|-----------|----------------------------------|--------------------------------------|------|----------------|----------------------------------|-------------------------------|
| Año  | X         |                                  |                                      | Año  | X              |                                  |                               |
|      | I         | 696                              | -7                                   |      | I              | 696                              | -13                           |
| 2020 | II        | 1,360                            | -6                                   | 2020 | II             | 1,360                            | -11                           |
|      | III       | 2,840                            | -4                                   |      | III            | 2,840                            | -9                            |
|      | IV        | 1,439                            | -4                                   |      | IV             | 1,439                            | -7                            |
|      | I         | 735                              | -3                                   |      | I              | 735                              | -5                            |
| 2021 | II        | 1,317                            | -2                                   | 2021 | II             | 1,317                            | -3                            |
|      | III       | 2,556                            | <b>-1</b>                            |      | III            | 2,556                            | <b>-1</b>                     |
|      | IV        | 1187                             | <b>0</b>                             |      | IV             | 1,187                            | <b>1</b>                      |
|      | I         | 785                              | <b>1</b>                             |      | I              | 785                              | <b>3</b>                      |
| 2022 | II        | 1,193                            | <b>2</b>                             | 2022 | II             | 1,193                            | <b>5</b>                      |
|      | III       | 2,474                            | <b>3</b>                             |      | III            | 2,474                            | <b>7</b>                      |
|      | IV        | 1,234                            | <b>4</b>                             |      | IV             | 1,234                            | <b>9</b>                      |
|      | I         | 638                              | <b>5</b>                             |      | I              | 638                              | <b>11</b>                     |
| 2023 | II        | 1,191                            | <b>6</b>                             | 2023 | II             | 1,191                            | <b>13</b>                     |
|      | III       | 2,822                            | <b>7</b>                             |      |                |                                  |                               |

Fuente: Elaboración propia.

La Tabla 1 a) muestra la codificación de los datos en la demanda de energía cuando el número de trimestres es impar. Si se contabiliza el total de los trimestres desde el año 2020 al 2023 da como resultado 15 trimestres, por lo tanto, el valor central se encuentra en el trimestre IV del año 2021. Entonces, la codificación de los datos se establece con pasos de 1 a partir de la media hacia el valor del último trimestre de análisis, mientras que de manera descendente con pasos de -1 a partir de la media hasta el dato del primer trimestre de análisis.

En cambio, la Tabla 1 b) representa la codificación de los datos en la demanda de energía cuando el número de trimestres es par, obteniéndose un total de 14 desde el primer trimestre de 2020 al segundo trimestre de 2023 con un valor

central entre el tercer y cuarto trimestre de los datos del año 2021. Por lo tanto, la codificación se dará con pasos de 2 a partir del valor central iniciando en 1 hasta el último trimestre, mientras que de manera descendente con pasos de -2 a partir de la media hasta el dato del primer trimestre de análisis.

### III. Análisis de resultados y discusión

#### 3.1. Variación estacional

En este capítulo se definen los pasos para calcular la componente estacional de una serie de tiempo en el análisis de datos estadísticos. Así mismo, se determina la línea de tendencia a través del método de mínimos cuadrados calculando el coeficiente de determinación, el coeficiente de correlación y el error estándar de la estimación, buscando contribuir en el ámbito de la predicción en la demanda de energía eléctrica que ayuden a la mejora continua en este tipo de modelos. En el caso de la regresión lineal, representa un método muy potente al momento de evaluar como influyen las variables en la evolución de la respuesta.

Tabla 2. Consumo de energía eléctrica.

| Meses      | Consumo total (kWh) |       |       |       |       |
|------------|---------------------|-------|-------|-------|-------|
|            | 2023                | 2022  | 2021  | 2020  | 2019  |
| Enero      | 355                 | 381   | 357   | 356   | 364   |
| Febrero    | 288                 | 286   | 276   | 294   | 344   |
| Marzo      | 329                 | 309   | 293   | 278   | 276   |
| Abril      | 345                 | 392   | 351   | 241   | 324   |
| Mayo       | 466                 | 583   | 382   | 510   | 404   |
| Junio      | 1,056               | 1,140 | 915   | 543   | 577   |
| Julio      | 1,108               | 1,333 | 1,275 | 1,075 | 969   |
| Agosto     | 1,731               | 1,346 | 1,399 | 1,091 | 1,302 |
| Septiembre | 1,626               | 1,352 | 1,343 | 1,142 | 1,130 |
| Octubre    | 1,266               | 1,409 | 1,294 | 1,200 | 1,159 |
| Noviembre  | 913                 | 922   | 954   | 864   | 910   |
| Diciembre  | 432                 | 413   | 507   | 385   | 421   |

Fuente: Elaboración propia.

La variación estacional en la variable de estudio se presenta por la influencia de las estaciones del año, por lo que representan los movimientos de la serie que se repiten año tras año en los mismos meses, en los mismos trimestres o en los mismos semestres del año y con la misma intensidad. La Tabla 2 muestra los consumos de energía eléctrica históricos de una familia de 4 integrantes durante el periodo de 2019-2023, estos datos serán utilizados para aplicar el análisis estacional de la serie de tiempo.

### 3.2. Pasos para el Análisis de Series de Tiempo (AST)

*I. Determinar los promedios trimestrales.* Cada año se divide en trimestres con el objetivo de identificar el índice estacional asociado a cada mes o trimestre del año, determinando un conjunto de índices mensuales que consiste en 12 índices en caso de que los periodos sean divididos en meses, o cuatro índices si se trata de trimestres. En la tabla 3 se muestra la relación del consumo de energía por trimestre para cada periodo de análisis.

Tabla 3. Consumo de energía eléctrica trimestrales.

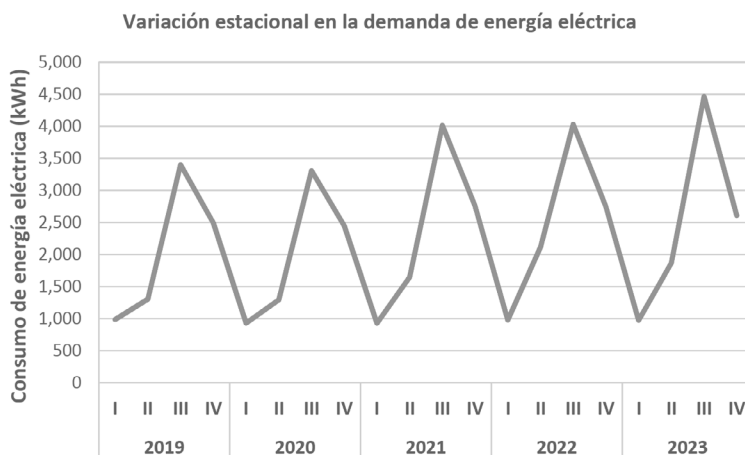
| Año  | Trimestre I | Trimestre II | Trimestre III | Trimestre IV |
|------|-------------|--------------|---------------|--------------|
| 2019 | 984         | 1,305        | 3,401         | 2,490        |
| 2020 | 928         | 1,294        | 3,308         | 2,449        |
| 2021 | 926         | 1,648        | 4,017         | 2,755        |
| 2022 | 976         | 2,115        | 4,031         | 2,744        |
| 2023 | 972         | 1,867        | 4,465         | 2,610        |

Fuente: Elaboración propia.

*II. Graficar el comportamiento de los datos trimestrales.* Con el fin de mostrar el comportamiento del consumo de energía en todo el periodo de análisis, se grafican los datos trimestrales de la tabla 3.

Se puede observar en la gráfica 3 que el mayor consumo de energía se da durante el tercer trimestre, mientras que el menor consumo es durante el primer trimestre de cada año, por lo que en términos generales los valores en los consumos de energía muestran una tendencia creciente.

Gráfica 3. Variación estacional trimestral en la demanda de energía eléctrica.



Fuente: Elaboración propia.

III. Obtención de los promedios móviles trimestrales. Como primer paso en la columna (1) de la Tabla 4 se colocan los años considerados como datos históricos, el cual para este caso corresponden de 2019 a 2023. En la columna (2) se enumeran los trimestres de cada año considerados en el análisis, mientras que en la tercera se coloca el consumo de energía que corresponde a cada trimestre. En la columna (4), se calcula el total móvil de 4 trimestres para la serie de tiempo, para hacerlo, se determina el total de los valores para los trimestres durante el primer año, por ejemplo:  $984 + 1,305 + 3,401 + 2,490 = 8,180$ . Un total móvil se asocia con el dato que ocupa el lugar medio del conjunto de valores del cual fue calculado, por lo que, para el primer total de 8,180 que se determinó a partir de cuatro datos, se colocará frente al punto medio de esos trimestres dentro de la columna (4) de la Tabla 4, entre los renglones 2019-II y 2019-III. Por lo tanto, el siguiente total móvil se calcula eliminando el valor 2019-I, por ejemplo:  $1,305 + 3,401 + 2,490 + 928 = 8,124$ . De esta manera, el segundo total de 8,124 se colocará entre los renglones 2019-III y 2019-IV de la Tabla 4.

Tabla 4. Promedios móviles centrados de los cuatros trimestres.

| Año<br>(1) | Trimes-<br>tre<br>(2) | kWh<br>(3) | Total móvil<br>4 trimestres<br>(4) | Prom. móvil<br>4 trimestres<br>(5) | Prom.<br>móvil<br>Centrado<br>(6) | Valor<br>estacio-<br>nal espe-<br>cífico<br>(7) |
|------------|-----------------------|------------|------------------------------------|------------------------------------|-----------------------------------|-------------------------------------------------|
|            | I                     | 984        |                                    |                                    |                                   |                                                 |
|            |                       |            |                                    |                                    |                                   |                                                 |
|            | II                    | 1,305      |                                    |                                    |                                   |                                                 |
| 2019       |                       |            | 8,180                              | 2,045                              |                                   |                                                 |
|            | III                   | 3,401      |                                    |                                    | 2,038                             | 166.88%                                         |
|            |                       |            | 8,124                              | 2,031                              |                                   |                                                 |
|            | IV                    | 2,490      |                                    |                                    | 2,030                             | 122.68%                                         |
|            |                       |            | 8,113                              | 2,028                              |                                   |                                                 |
|            | I                     | 928        |                                    |                                    | 2,017                             | 46.02%                                          |
|            |                       |            | 8,020                              | 2,005                              |                                   |                                                 |
|            | II                    | 1,294      |                                    |                                    | 2,000                             | 64.70%                                          |
| 2020       |                       |            | 7,979                              | 1,995                              |                                   |                                                 |
|            | III                   | 3,308      |                                    |                                    | 1,995                             | 165.86%                                         |
|            |                       |            | 7,977                              | 1,994                              |                                   |                                                 |
|            | IV                    | 2,449      |                                    |                                    | 2,039                             | 120.14%                                         |
|            |                       |            | 8,331                              | 2,083                              |                                   |                                                 |
|            | I                     | 926        |                                    |                                    | 2,171                             | 42.65%                                          |
|            |                       |            | 9,040                              | 2,260                              |                                   |                                                 |
|            | II                    | 1,648      |                                    |                                    | 2,298                             | 71.71%                                          |
| 2021       |                       |            | 9,346                              | 2,337                              |                                   |                                                 |
|            | III                   | 4,017      |                                    |                                    | 2,343                             | 171.47%                                         |
|            |                       |            | 9,396                              | 2,349                              |                                   |                                                 |
|            | IV                    | 2,755      |                                    |                                    | 2,407                             | 114.44%                                         |
|            |                       |            | 9,863                              | 2,466                              |                                   |                                                 |
|            | I                     | 976        |                                    |                                    | 2,468                             | 39.55%                                          |
|            |                       |            | 9,877                              | 2,469                              |                                   |                                                 |
|            | II                    | 2,115      |                                    |                                    | 2,468                             | 85.70%                                          |

| Año<br>(1) | Trimes-<br>tre<br>(2) | kWh<br>(3) | Total móvil<br>4 trimestres<br>(4) | Prom. móvil<br>4 trimestres<br>(5) | Prom.<br>móvil<br>Centrado<br>(6) | Valor<br>estacio-<br>nal espe-<br>cífico<br>(7) |
|------------|-----------------------|------------|------------------------------------|------------------------------------|-----------------------------------|-------------------------------------------------|
| 2022       |                       |            | 9,866                              | 2,467                              |                                   |                                                 |
|            | III                   | 4,031      |                                    |                                    | 2,466                             | 163.46%                                         |
|            |                       |            | 9,862                              | 2,466                              |                                   |                                                 |
|            | IV                    | 2,744      |                                    |                                    | 2,435                             | 112.71%                                         |
|            |                       |            | 9,614                              | 2,404                              |                                   |                                                 |
|            | I                     | 972        |                                    |                                    | 2,458                             | 39.55%                                          |
|            |                       |            | 10,048                             | 2,512                              |                                   |                                                 |
|            | II                    | 1,867      |                                    |                                    | 2,495                             | 74.82%                                          |
| 2023       |                       |            | 9,914                              | 2,478                              |                                   |                                                 |
|            | III                   | 4,465      |                                    |                                    |                                   |                                                 |
|            |                       |            |                                    |                                    |                                   |                                                 |
|            | IV                    | 2,610      |                                    |                                    |                                   |                                                 |

Fuente: Elaboración propia.

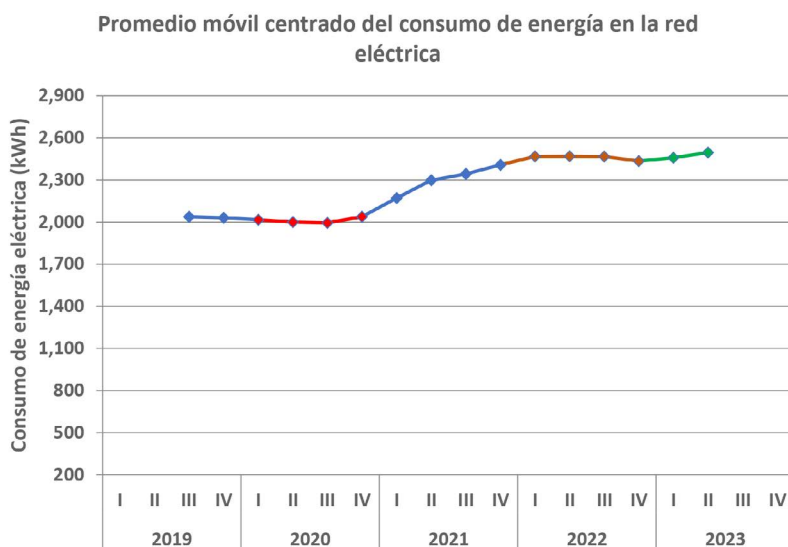
El procedimiento se repite hasta llegar al último total móvil, por ejemplo:  $972 + 1,867 + 4,465 + 2,610 = 9,914$ .

En la columna (5) el promedio móvil centrado se calcula dividiendo el total móvil entre los cuatro trimestres del periodo de análisis, por ejemplo:  $8,180 \div 4 = 2,045$  hasta el último valor  $9,914 \div 4 = 2,478$ .

La sexta columna se determina el promedio móvil centrado que resulta dividir los promedios móviles en pares, por ejemplo:  $(2,045 + 2,031) \div 2 = 2,038$  y así sucesivamente hasta el último valor  $(2,512 + 2,478) \div 2 = 2,495$ .

Por último, la columna (7) que corresponde al valor estacional específico, se calcula dividiendo el consumo de energía (3) entre el promedio móvil centrado (6), por ejemplo:  $(3,401 \div 2,038) \times 100 = 166.88 \%$  y así sucesivamente hasta el último valor  $(1,867 \div 2,495) \times 100 = 74.82 \%$ .

Gráfica 4. Promedio móvil centrado del consumo de energía en la red eléctrica.



Fuente: Elaboración propia.

Con la finalidad de describir el grado de la variación estacional en el comportamiento de la demanda de energía, los datos de la columna (6) son representados en la gráfica 4. Dichos datos muestran que el consumo promedio de energía oscila entre 2,000 y 2,600 kWh, además, se aprecia un comportamiento sin cambios entre los periodos 2019-II al 2020-III. Sin embargo, a partir de 2020-IV el consumo de energía ha ido en aumento gradualmente hasta el segundo trimestre de 2023.

*IV. Calcular la media modificada y lo índices estacionales trimestrales.* La media modificada representa un índice de la componente estacional, para calcularla, se debe descartar los valores más alto y bajo de cada trimestre, para que después los datos restantes sean promediados. Los valores estacionales recuperados de cada trimestre correspondientes a los datos de la columna (7) de la tabla 4, contienen aún las componentes cíclicas e irregulares de la serie tiempo. Por lo tanto, para contrarrestar esto, se eliminan los valores más alto y bajo de cada trimestre, tal como se muestra en la tabla 5.

Tabla 5. Media modificada.

| Año           | Trimestre |         |          |          |
|---------------|-----------|---------|----------|----------|
|               | I         | II      | III      | IV       |
| 2019          | ---       | ---     | 166.88 % | 122.68 % |
| 2020          | 46.02 %   | 64.70 % | 165.86 % | 120.14 % |
| 2021          | 42.65 %   | 71.71 % | 171.47 % | 114.44 % |
| 2022          | 39.55 %   | 85.70 % | 163.46 % | 112.71 % |
| 2023          | 46.20 %   | 74.82 % | ---      | ---      |
| Prom/<br>Trim | 44.33 %   | 73.27 % | 166.37 % | 117.29 % |

Fuente: Elaboración propia.

Al sumar los cuatro índices trimestrales  $44.33 + 73.27 + 166.37 + 117.29 = 401.26$ , sin embargo, dado que la base de un índice es 100, entonces, tomando en cuenta los cuatro índices trimestrales deberá dar un total de 400. Para corregir este error se debe multiplicar cada uno de los índices por una constante de ajuste, la cual se determina dividiendo la suma ideal de los índices entre la suma real, por ejemplo:  $400 \div 401.26 = 0.99688053$ .

$$\text{Factor de ajuste} = \frac{400}{401.26} = 0.99688053$$

Después, se aplica el factor de ajuste a cada periodo multiplicándolo por la media modificada para determinar el índice estacional (Tabla 6), por ejemplo:

$$\text{Índice estacional} = \left( \frac{44.33}{100} \right) \times 0.99688053 = 0.44193343$$

Tabla 6. Índice estacional trimestral.

| Trimestre | índices<br>ajustados | Factor de<br>de ajuste | índice<br>Estacional |
|-----------|----------------------|------------------------|----------------------|
| I         | 44.33%               | 0.99688053             | 0.44193343           |
| II        | 73.27%               | 0.99688053             | 0.73036836           |
| III       | 166.37%              | 0.99688053             | 1.6584702            |
| IV        | 117.29%              | 0.99688053             | 1.16922801           |

Fuente: Elaboración propia.

Los datos calculados muestran lo siguiente:

- a). Un índice de 44.19 % para el primer trimestre, significa que el consumo de energía se encontrará 55.81 % por debajo del promedio típico ( $100 - 44.19 = 55.81$ ).
- b). Para el segundo trimestre se tiene un índice de 73.03 %, esto representa un déficit del 26.97 % en el consumo de energía con respecto al promedio ( $100 - 73.03 = 26.97$ ).
- c). En el tercer trimestre, con un índice de 165.84 %, el consumo de energía estará un 65.84 % por encima del promedio ( $165.84 - 100 = 65.84$ ).
- d). En el cuarto trimestre, con un índice de 116.92 % representa que el consumo de energía estará un 16.92 % por arriba del promedio ( $116.92 - 100 = 16.92$ ).

Por lo tanto, de lo anterior se puede establecer que los mayores incrementos en el consumo de energía se encuentran en el tercer y cuarto trimestre.

V. Aplicar el ajuste estacional a los datos trimestrales. En este paso se extraen de los datos (consumo de energía) la variación estacional (datos desestacionalizados). El procedimiento consiste en dividir los datos originales de cada trimestre por el factor de ajuste correspondiente. Por ejemplo, el consumo de energía del primer trimestre de 2019 se divide por su índice estacional correspondiente ( $984 \div 0.4419 = 2,227$ ) y así sucesivamente hasta finalizar con el primer trimestre de 2023 ( $972 \div 0.4419 = 2,199$ ). En los casos del segundo, tercer y cuarto trimestre el procedimiento se repite utilizando el índice estacional 0.7303, 1.6584 y 1.1692 respectivamente, tal como se muestra en la Tabla 7.

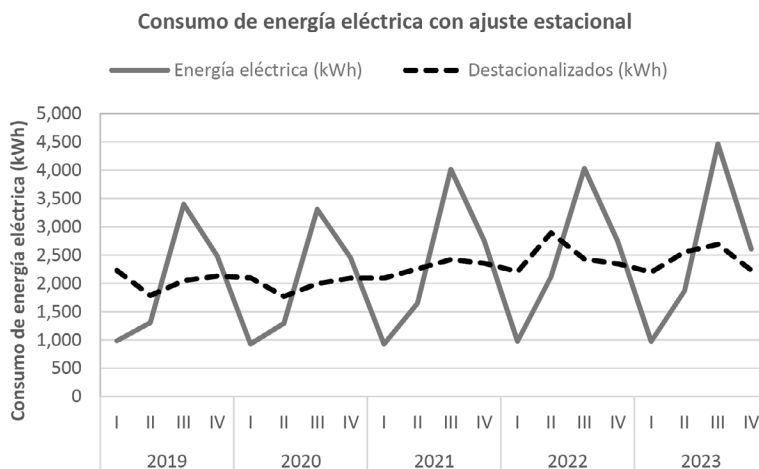
Tabla 7. Consumo de energía con ajuste estacional (kWh).

| Trimestre  | 2019  | 2020  | 2021  | 2022  | 2023  |
|------------|-------|-------|-------|-------|-------|
| <b>I</b>   | 2,227 | 2,100 | 2,095 | 2,208 | 2,199 |
| <b>II</b>  | 1,787 | 1,772 | 2,256 | 2,896 | 2,556 |
| <b>III</b> | 2,051 | 1,995 | 2,422 | 2,431 | 2,692 |
| <b>IV</b>  | 2,130 | 2,095 | 2,356 | 2,347 | 2,232 |

Fuente: Elaboración propia.

Con el cálculo del consumo de energía ajustado representado en la gráfica 5, permite observar que las diferencias en los consumos de energías de un trimestre a otro con los datos originales no son tan visibles, destacando que la demanda de energía aumenta de un año a otro.

Gráfica 5. Consumo de energía eléctrica con ajuste estacional.



Fuente: Elaboración propia.

VI. Obtener la línea de tendencia. Para determinar la línea de tendencia se aplicó el método de mínimos cuadrados. De esta manera, utilizando las ecuaciones (2) y (3) se obtienen los valores constantes  $a$  y  $b$  correspondientes a la ecuación de regresión lineal, quedando de la siguiente manera:

$$\hat{Y} = 2,242 + 15.203x \quad (\text{ec. 9})$$

Por lo tanto, la ecuación anterior representa la línea de tendencia de los datos en la demanda de energía, los cuales permiten establecer la relación existente entre el comportamiento de cada trimestre (Tabla 8) y la estimación de datos futuros.

Tabla 8. Codificación de la variable tiempo para datos pares en la demanda de energía.

| Trimestre |     | Energía (kWh)          | Tiempo              |                 |                |       |
|-----------|-----|------------------------|---------------------|-----------------|----------------|-------|
|           |     | Destacionali-<br>zados | codifica-<br>do (x) | xY              | x2             | =a+bx |
| Año       | X   | Y                      |                     |                 |                |       |
| 2019      | I   | 2,227                  | -19                 | -42,319         | 361            | 1,953 |
|           | II  | 1,787                  | -17                 | -30,375         | 289            | 1,984 |
|           | III | 2,051                  | -15                 | -30,760         | 225            | 2,014 |
|           | IV  | 2,130                  | -13                 | -27,685         | 169            | 2,045 |
| 2020      | I   | 2,100                  | -11                 | -23,099         | 121            | 2,075 |
|           | II  | 1,772                  | -9                  | -15,945         | 81             | 2,105 |
|           | III | 1,995                  | -7                  | -13,962         | 49             | 2,136 |
|           | IV  | 2,095                  | -5                  | -10,473         | 25             | 2,166 |
| 2021      | I   | 2,095                  | -3                  | -6,286          | 9              | 2,197 |
|           | II  | 2,256                  | -1                  | -2,256          | 1              | 2,227 |
|           | III | 2,422                  | 1                   | 2,422           | 1              | 2,258 |
|           | IV  | 2,356                  | 3                   | 7,069           | 9              | 2,288 |
| 2022      | I   | 2,208                  | 5                   | 11,042          | 25             | 2,318 |
|           | II  | 2,896                  | 7                   | 20,271          | 49             | 2,349 |
|           | III | 2,431                  | 9                   | 21,875          | 81             | 2,379 |
|           | IV  | 2,347                  | 11                  | 25,815          | 121            | 2,410 |
| 2023      | I   | 2,199                  | 13                  | 28,593          | 169            | 2,440 |
|           | II  | 2,556                  | 15                  | 38,344          | 225            | 2,470 |
|           | III | 2,692                  | 17                  | 45,768          | 289            | 2,501 |
|           | IV  | 2,232                  | 19                  | 42,404          | 361            | 2,531 |
|           |     | $\Sigma=44,847$        | $\Sigma=0$          | $\Sigma=78,603$ | $\Sigma=2,660$ |       |

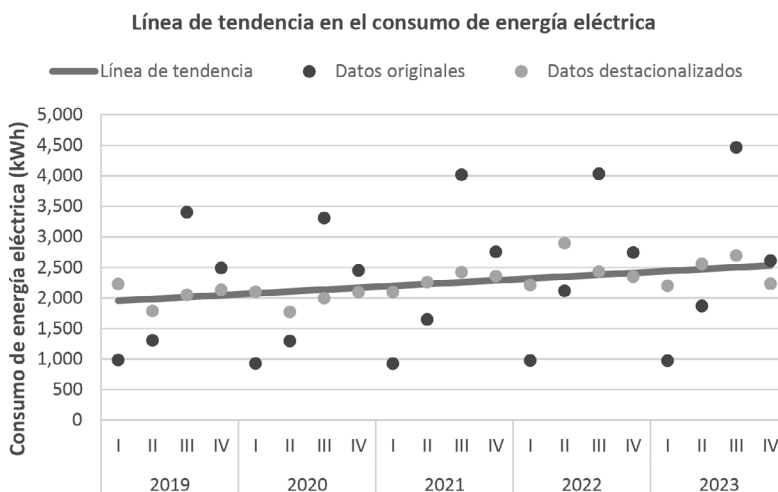
Fuente: Elaboración propia.

Para observar de una mejor forma el comportamiento de la línea de tendencia, la gráfica 6 muestra la tendencia del consumo de energía. Dado que la pendiente de la ecuación (9) es 15.20, se puede entender que alrededor de los últimos 15 trimestres el consumo de energía aumentó a razón de 15.20 unidades por trimestre.

Así mismo, empleando la ecuación (4) se obtuvo que el error estándar en la estimación ( $S_e$ ) fue 212, lo que indica una dispersión de solo el 9.5 % con relación al promedio de los datos desestacionalizados (2,242).

Por último, los resultados estadísticos obtenidos a través de las ecuaciones (5) y (6) para los coeficientes de determinación ( $r^2$ ) y correlación ( $r$ ), muestran los valores 0.43 y 0.66 respectivamente. Esto representa que el 43 % de la variación de los datos es explicada por la recta de regresión lineal, en cambio, el 66 % establece la relación que tienen las variables de estudio, para este caso la variable tiempo y el consumo de energía.

Gráfica 6. Línea de tendencia en el consumo de energía eléctrica.



Fuente: Elaboración propia.

## Conclusiones

El análisis de una serie de tiempo aplicado en la demanda de energía eléctrica permite observar con detalle el comportamiento histórico de las variables de estudio, así como también la estimación de valores futuros. Los resultados obtenidos en el análisis estadístico desarrollado muestran que la serie de tiempo tiene una relación de 66 % con el consumo de energía, lo que representa un aumento proporcional de acuerdo con la línea de tendencia. Dicho aumento es más evidente durante el tercer trimestre de cada año del periodo de análisis.

sis, pues los resultados muestran un incremento del 65.8 % en la demanda de energía, esto servirá como indicador para establecer estrategias que ayuden a disminuir el consumo de energía, adoptando medidas que mejoren las actitudes favorables para el uso correcto y adecuado de la energía eléctrica, contribuyendo a un consumo responsable y aportando directamente a la reducción de emisiones indirectas.

Adicionalmente, los resultados muestran de manera conjunta que el primer y segundo trimestre representan los periodos con menor consumo de energía, con un déficit del 56 % y 26 % por debajo del promedio típico en la demanda de energía respectivamente. Por otro lado, el tercer y cuarto semestre representan los periodos con mayor consumo de energía con un superávit del 66 % y 17 % respectivamente.

Por lo tanto, con la mejora de las herramientas de predicción en la demanda de energía generalmente asociadas a las emisiones de CO<sub>2</sub>, este parámetro podría verse reducido de forma considerable, disminuyendo así el impacto ambiental del sector energético en su conjunto.

Actualmente los efectos originados por el cambio climático son una realidad, por lo que se requiere seguir trabajando en el análisis de modelos estadísticos que aporten información en el comportamiento de la demanda de energía. Sin embargo, también es necesario actualizar los marcos regulatorios que permitan el desarrollo y expansión de nuevas alternativas de energías renovables con el fin de garantizar una energía asequible para todos, permitiendo mejorar las prácticas sostenibles para aumentar la eficiencia energética y combatir el cambio climático tal como lo establecen los Objetivos de Desarrollo Sostenible de la agenda 2030.

## Referencias

- Amador, L., Parra, R., y Castro, C. (2021). Cálculo de Emisiones de Gases de Efecto Invernadero (GEI) en la Red eléctrica del Tecnológico Nacional de México/IT de Los Mochis Campus Villa de Ahome (ITLMCVA). *RA XIMHAI*, 17(3), 47-70. doi: 10.35197/rx.17.03.2021.02.la.
- Amador, L., Parra, R., y Rodríguez, R. (2022). Comportamiento de los patrones de cambio en la demanda de energía eléctrica durante el Covid-19 y análisis de las actitudes favorables para un consumo responsable en el Instituto Tecnológico de Los Mochis, Campus Villa de Ahome. En C. S. En J. Rodríguez, *Aportes a los Objetivos del Desarrollo Sostenible de la Agenda 2030* (pág. 248). Los Mochis: Astra Ediciones S. A. de C. V.
- CEPAL. (2021). *Estadísticas ambientales y de cambio climático para América Latina y el Caribe*. Division de estadísticas ambientales.
- Chapra, S., y Canale, R. (2015). *Métodos numéricos para ingenieros* (Séptima ed.). México: McGraw-Hill Education.
- Gobierno del Estado de Sinaloa. (2018). *Plan Sectorial de Desarrollo Urbano, Medio Ambiente, Movilidad y Vivienda 2017-2021*. Culiacán: Gobierno del Estado de Sinaloa. Obtenido de <https://media.transparencia.sinaloa.gob.mx/uploads/files/1/PLAN%20SECTORIAL%20DE%20DESARROLLO%20URBANO.pdf>
- Gobierno del Estado de Sinaloa. (2022). *Plan Estatal de Desarrollo (PED) 2022-2027*. Culiacán: Gobierno del Estado de Sinaloa. Obtenido de <https://ped.sinaloa.gob.mx/wp-content/uploads/2022/04/PED27-compressed.pdf>
- Levin, R., y Rubin, D. (2004). *Estadística para Administración y Economía*. México: Pearson Educación.
- Moreno, T. (2019). *El pronóstico de ventas en los negocios: modelos y aplicaciones* (Primera ed.). Santiago de Chile: RiL Editores.
- ONU. (2021). *Objetivos de Desarrollo Sostenible 2030*. Recuperado el 25 de enero de 2021, de <https://www.un.org/sustainabledevelopment/es/>

# Capítulo 5

---

## **Regresión múltiple desde el punto de vista de la gestión-administración con un enfoque sustentable caso ganadería Rubio**

*Juan Héctor Alzate Espinoza  
Adalid Graciano Obeso  
Grace Erandy Báez Hernández*

## Introducción

### *Antecedentes*

La Ganadería Rubio es una empresa con más de 25 años de experiencia en el sector ganadero, reside en una de las zonas con mayor comercialización de productos cárnicos en el estado de Sinaloa, asumiendo una filosofía orientada a la mejora continua, medioambiente y la calidad de su ganado; constantemente está en búsqueda de estrategias y alternativas que le permitan seguir manteniéndose en la mente de sus clientes cómo una opción prioritaria en el mercado frente a negocios mejor posicionados como grupo SuKarne, se busca obtener un mejor rendimiento de sus productos, extrayendo la mayor cantidad de carne en su ganado, manteniendo el nivel de calidad estándar, permitiendo expandir el alcance de la empresa a un mercado más amplio y diversificar la cantidad de productos que maneja, durante este proceso se utilizará la metodología de regresión lineal como herramienta para alcanzar la sustentabilidad y rentabilidad de la Ganadería Rubio.

### *Justificación*

La cantidad de producto alimenticio desperdiciado en empresas cárnicas afecta negativamente al entorno que le rodea, representando un gasto innecesario de recursos, se debe promover el desarrollo comunitario generando una mayor calidad de vida y bienestar social humano, mediante la gesta en la calidad de los productos de la industria gestionando los residuos orgánicos para el aprovechamiento en empresas ganaderas mejorando las cadenas de suministro.

### *Objetivos*

- Lograr un mayor nivel de producción cárnico que mejore la rentabilidad de la empresa
- Aplicación de un modelo de regresión lineal para realizar proyecciones de venta
- Presentar una alternativa que contribuya a mejorar la cadena de suministro para el ganado
- Incrementar la cartera de clientes
- Diversificar la cadena de suministros

### *Metodología*

La Sustentabilidad es un tema de desarrollo y por tanto, de proceso de desenvolvimiento de capital, de variables como crecimiento económico, de PIB per cápita, de movilidad social y de mejoramiento de la población en términos de empleo, ingresos reales y bienestar social. (INEGI, 2009). La empresa que es sustentable y competitiva, diseña y trabaja con procesos productivos que respetan la naturaleza, diseña e implementar programas de educación ambiental para sus trabajadores y el público en general, exige a los stakeholder, políticas y acciones similares a las que implanta, para favorecer la tierra beneficiando a todos los stakeholder finalmente (Valderrama; 2007).

Ya no basta con tener una fundación, dar donativos para una cierta causa, desarrollar campañas de mercadotecnia social o reforestar el camellón frente a las oficinas... Hoy las empresas deben demostrar esta responsabilidad de forma congruente y permanente en todos los niveles y en todas las áreas de la organización... alineada con sus objetivos y tomada en cuenta en todas sus decisiones, para que esta tenga un verdadero sentido para la empresa y le permita identificar claramente su importancia y los beneficios que se pretenden generar, no solo a corto plazo, sino también a largo plazo. (Gutiérrez, P; 2009)

El concepto de sostenibilidad ambiental urbana logró abrir una nueva visión o perspectiva integral, transversal y a largo plazo. Inclusive, es preciso medir si el capital de los recursos naturales puede soportar indefinidamente el desarrollo urbano actual. Para ello, es imperante relacionar la demanda de los recursos necesarios para satisfacer las necesidades de los habitantes de las ciudades a través de indicadores e índices que midan estas relaciones e interacciones complejas (Bettini, 1998).

La ganadería plantea una grave amenaza a los ecosistemas y la biodiversidad locales, la tierra se utiliza para pastoreo y producción de ganado. Hasta una cuarta parte de la superficie terrestre se utiliza para pastoreo, mientras que aproximadamente un tercio de la tierra cultivable se utiliza para la producción de cosechas para animales con una tendencia de crecimiento. Como resultado, las especies animales luchan por la falta de hábitat y algunas incluso están al borde de la extinción. Cuando se talan árboles para dar paso a tierras de cultivo, las especies nativas supervivientes deben trasladarse para encontrar nuevos hábitats. Este tipo de deforestación, conocido como “cambio de uso

de la tierra”, es una de las principales causas del cambio climático porque los bosques son un importante sumidero de carbono que elimina los gases de efecto invernadero de la atmósfera. En el sector cárnico es de conocimiento empírico que no se puede aprovechar el cien por ciento de la carne en la mayoría de los casos, existen diversas variables que terminan afectando la calidad y el tejido de la misma así como las cualidades nutricionales posibles para el consumo humano, en general solo poco más de un sesenta por ciento del material orgánico es apto para ser procesado y vendido, sumado al impacto medioambiental que tiene el sector ganadero a escala global se requieren medidas para optimizar la producción de ganado en un entorno donde las demandas por alimentos incrementan desmesuradamente, algunas alternativas que permitirían reducir las consecuencias negativas al ecosistema son:

*La gestión de residuos orgánicos:* los planes de economía circular deben gestionar los residuos orgánicos a través de soluciones medioambientales para la industria cárnica, el uso de digestores anaeróbicos, la producción de biogás para calderas, el uso de plantas de cogeneración, el uso de biocombustibles en vehículos.

*Eficiencia energética:* Existen opciones como la cogeneración, así como el uso de equipos energéticamente eficientes y combustibles renovables.

*Consumo de agua:* Esta es una de las soluciones medioambientales en la industria cárnica, donde es necesario ahorrar agua, reutilizarla cuando sea posible, reciclar agua e incluso reducir el consumo energético en todos los procesos involucrados.

*Bio-Soluciones:* Requiere el uso de bioplásticos y biosurfactantes para reducir el uso de derivados del petróleo. Esto se aplica desde los cultivos hasta el procesamiento de subproductos de la industria cárnica. Implementar el mejoramiento genético para aumentar la producción de leche y carne, es lo que se busca con la inseminación artificial en Santo Domingo de los Tsáchilas, a través del Proyecto de Ganadería Sostenible, que ejecuta el Ministerio de Agricultura, Ganadería, Acuacultura y Pesca (Leopoldo, 2005).

Actualmente existe una tendencia al desarrollo de nuevas herramientas para realizar investigaciones, el almacenamiento y procesamiento de datos procedente de varios sensores. Generalmente estos dispositivos están localizados en lugares remotos. Los métodos típicos de recopilación de datos implican el uso de dispositivos para cada variable de interés, lo que dificulta y encarece la integración y el procesamiento conjunto.

El análisis de regresión trata del estudio de la dependencia de una variable dependiente respecto de una o más variables independientes con el objetivo de

estimar o predecir el valor promedio poblacional de la variable dependiente en términos de REICE | 136 los valores conocidos de la variable independiente. (Gujarati y Porter, 2010). Se considera la técnica que se ocupa de analizar la dependencia entre una variable dependiente y una o más variables explicativas. Su objetivo consiste en estimar y/o predecir el valor medio poblacional de la variable dependiente a partir de los valores conocidos y fijos de las variables explicativas, obtenidos mediante un proceso de muestreo repetido. (Díaz Fernández y Llorente Marrón, 2013). Se trata de un estudio no experimental de tipo prospectivo, donde los sujetos son seleccionados por poseer determinados valores que pueden ejercer cierta influencia y de grupo único (se selecciona una muestra de sujetos que presenten todos los niveles de las variables independientes) con finalidad predictiva sobre la variable predicha (VP) (Montero y León, 2002). Los métodos multivariantes permiten reducir o sintetizar información sin perder la brindada por las mediciones originales, muy utilizada en el campo de las ciencias sociales (Rodríguez, 2022). La regresión lineal es una herramienta versátil capaz de generar una vasta cantidad de información útil para los líderes en empresas que necesitan orientación al momento de tomar decisiones estratégicas. Existen investigaciones sobre la construcción de índices e indicadores ambientales y de desarrollo sostenible utilizando técnicas de análisis multivariante, como el análisis de componentes principales (ACP) y el análisis de conjuntos difusos (DP2) (Montosa, 2014). Las técnicas para la regresión lineal implican una metodología de clasificación automática jerárquica, donde se reduce el número de individuos en grupos homogéneos y significativos, denominados conglomerados. En este punto, es preciso, por un lado, definir una medida de la distancia entre individuos que nos indique su grado de similitud o de semejanza y, por otro lado, determinar una estrategia de agrupación de los individuos para la constitución de las sucesivas clases (Gómez, 2021). Incidiendo en la utilidad medioambiental que provee el sistema de regresión lineal se cita un ejemplo en el cual se realizó un estudio estadístico sobre la correlación entre contaminantes atmosféricos y variables meteorológicas en la zona norte de Chiapas, México. Los autores aplicaron análisis de RLM para predicción de PM10. Las variables significativas en el modelo fueron temperatura, humedad relativa y dirección del viento (Ramos-Herrera et al, 2010). Es una técnica que se basa en la clasificación de los individuos o elementos de estudio. Esta técnica, se observó que ha sido aplicada cuando el objetivo de estudio es la segmentación de los clientes con respecto a la satisfacción, luego se identifica los factores mediante la aplicación de un análisis discriminante (Legarreta, 2019).

## **Análisis de proceso en la empresa ganadera Rubio**

El modelo de negocios en la ganadera Rubio empieza desde la producción primaria cuando el ganado proveniente de fincas aledañas se realiza la crianza, alimentación y eventual sacrificio o exportación de los especímenes, cuenta con cuartos refrigerados cuya capacidad permite almacenar hasta 20 toneladas de carne variada, huesos, sebo, harinas de carne y otros subproductos; la carne es evaluada, procesada y posteriormente empacada para distribuirse a su cartera de clientes frecuentes, es un modelo estándar que prácticamente cualquier empresa cárnica efectúa; pese a ello se detectaron algunos elementos clave que podrían ayudar a mejorar su desempeño, reducir costos de producción y elevar la calidad de sus productos con un enfoque sustentable.

La ganadera enfrenta un problema respecto a sus fuentes de energía, debido a su localización esta empresa cuenta con un ambiente propicio para el cultivo de alimentos y la crianza adecuada del ganado, pero generalmente enfrentan fallos eléctricos que dificultan realizar actividades cotidianas desde el área de recursos humanos, contabilidad e incluso interfieren en la conservación de carne cuando la refrigeración de los cuartos deja de funcionar, desencadenando en productos mermados y pérdidas significativas de ingresos, por este motivo se propuso la negociación para instalar una mayor cantidad de celdas solares que sean suficientes para cubrir esta necesidad, siendo una alternativa de emergencia que a largo plazo se vuelva una inversión rentable. Se evaluará el nivel de calidad de la carne; la importancia de esta actividad radica en que el aporte nutritivo de las cecinas lo constituye principalmente sus proteínas de origen animal, debido a su elevado valor biológico. En este ejercicio se trabajó conforme las necesidades de la compañía ganadera Rubio en base a una problemática central, la cantidad de carne desperdiciada y los agentes que afectan a su línea de producción la técnica de la regresión lineal fue aplicada para determinar las posibles variables que afectan a su calidad. En relación con los productos cárnicos, es muy importante la selección de la carne a utilizar, de acuerdo con el tipo de producto que se desea fabricar. Para realizar esta selección se deben conocer las siguientes propiedades de la carne que inciden en su calidad como materia prima para fabricar cecinas:

*Carga microbiana:* El grado de contaminación microbiana de la carne es vital para la obtención de un producto final que no constituya un riesgo para la salud del consumidor, especialmente si el producto cárnico se expende crudo. Se recomienda seleccionar a los proveedores realizando los controles microbiológicos más necesarios a la carne empleada como materia prima (Leítao, 1999).

*Capacidad de retención de agua:* Corresponde al porcentaje de agua que la carne es capaz de retener cuando se le somete a la acción de fuerzas externas (operaciones de cortado, presión, tratamiento térmico). Las proteínas de tipo miofibrilar son responsables del setenta por ciento de fijación de agua, las sarcoplasmáticas veinte por ciento y el tejido conectivo diez por ciento.

*Capacidad emulsionante:* Es la habilidad de la carne de sostener la grasa y producir emulsiones estables. Las carnes más apropiadas para formar emulsión son aquellas que poseen elevado contenido de proteínas contráctiles. Estas proteínas recubren o envuelven los glóbulos de grasa y al someter la emulsión a la acción del calor coagulan formando una especie de matriz rígida que atrapa cada partícula grasa.

*Color de la carne:* La proporción de mioglobina está aumentada en aquellos músculos de mayor actividad y por lo tanto mayor demanda de oxígeno; varía además según la especie y edad del animal: es unas 3 veces mayor en la de vacuno que en la de cerdo y animales más viejos presentan un pigmento más oscuro (Fernandez, 2001).

*Calidad sanitaria:* Los controles para evaluar la calidad sanitaria contemplan todos aquellos aspectos relacionados con la protección de la salud del consumidor y por lo tanto son regulados por las legislaciones de carácter obligatorio existentes en cada país. Todas las reglamentaciones consideran la calidad sanitaria como un requisito básico que debe ser cumplido satisfactoriamente, cualquiera sea el producto cárnico.

*Calidad sensorial:* Es indudable que las características organolépticas de una cecina influyen en gran medida en el consumidor. La calidad sensorial es factor decisivo para decidir la adquisición de un producto cárnico y por lo tanto los atributos sensoriales deben ser controlados a nivel industrial, si se desea mantener y/o aumentar el mercado.

*Calidad nutritiva:* Interesa conocer la composición centesimal del derivado cárnico: humedad, proteínas, grasas y cenizas. De las determinaciones recién señaladas el contenido de proteínas es determinante para evaluar la calidad nutritiva de la cecina, debido a la importancia de su presencia en una correcta dieta alimenticia.

Para evaluar la calidad del ganado a cada una de estas características se le asignó un valor en una escala de 1 a 10 fueron revisadas por el equipo de control de calidad de la ganadería centrándose en especímenes de res y cerdo, cada variable es representada por la siguiente abreviación:

CM: Carga Microbiana  
RA: Retención de agua  
CE: Capacidad emulsionante  
CC: Color de la carne  
CSA: Calidad sanitaria  
CSE: Calidad sensorial  
CN: Calidad Nutritiva

Selección y muestra: Se optó por una muestra de 8 especímenes tanto para la población de res como de los cerdos para obtener la evaluación promedio inicial del nivel de calidad del ganado. Este ejercicio se realizará con una evaluación Inicial efectuada durante el periodo de agosto del 2023 y una evaluación Final en el periodo de octubre del 2023 donde se comparará ambos niveles de calidad en torno al nuevo proceso de cuidado implementado. Los resultados iniciales obtenidos se plasman en las Tablas 1 y 2.

Tabla 1. Evaluación Inicial de la calidad en la Res.

| Evaluación Inicial de la calidad (Res) |       |       |       |       |       |       |       |       |          |
|----------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| Variable                               | Res 1 | Res 2 | Res 3 | Res 4 | Res 5 | Res 6 | Res 7 | Res 8 | Promedio |
| CM                                     | 2     | 1     | 2     | 3     | 2     | 2     | 1     | 1     | 1.75     |
| RA                                     | 6     | 7     | 7     | 6     | 8     | 9     | 8     | 8     | 7.375    |
| CE                                     | 9     | 7     | 8     | 8     | 7     | 6     | 7     | 8     | 7.5      |
| CC                                     | 8     | 6     | 9     | 8     | 9     | 8     | 8     | 9     | 8.125    |
| CSA                                    | 5     | 7     | 9     | 8     | 6     | 7     | 7     | 8     | 7.125    |
| CSE                                    | 8     | 9     | 8     | 7     | 7     | 6     | 8     | 8     | 7.625    |
| CN                                     | 6     | 7     | 8     | 7     | 7     | 8     | 6     | 9     | 7.25     |

Tabla 2. Evaluación Inicial de la Calidad en los Cerdos.

| Evaluación Inicial de la calidad (cerdos) |       |       |       |       |       |       |       |       |          |
|-------------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| Variable                                  | Res 1 | Res 2 | Res 3 | Res 4 | Res 5 | Res 6 | Res 7 | Res 8 | Promedio |
| CM                                        | 3     | 2     | 4     | 3     | 2     | 3     | 3     | 2     | 2.75     |
| RA                                        | 7     | 6     | 7     | 7     | 6     | 6     | 7     | 6     | 6.5      |
| CE                                        | 8     | 8     | 6     | 5     | 7     | 6     | 8     | 7     | 6.875    |
| CC                                        | 9     | 6     | 8     | 6     | 8     | 5     | 7     | 6     | 6.875    |
| CSA                                       | 8     | 5     | 8     | 7     | 7     | 6     | 6     | 5     | 6.5      |
| CSE                                       | 5     | 6     | 9     | 7     | 8     | 6     | 5     | 6     | 6.5      |
| CN                                        | 6     | 7     | 6     | 8     | 6     | 5     | 6     | 8     | 6.5      |

Un factor externo al nivel de calidad de la carne es la cantidad de material que puede extraerse de cada espécimen en particular, cada industria desarrolla sus propias técnicas para incrementar el índice de masa corporal en sus ganados y de este modo obtener mayor rendimiento a un coste menor, pero generalmente se recurren al uso de insumos que alteran el grado de pigmentación natural de la carne, sus cualidades nutritivas y sanitarias; afectando de forma negativa tanto al espécimen sacrificado y al consumidor final; La evolución de los sistemas de evaluación de alimentos hacia herramientas que permiten la predicción de diversas formas de respuesta animal tales como eficiencia, conversión y calidad del producto requieren de una integración cuantitativa del conocimiento del sitio y la extensión de la digestión de los alimentos, la cual impacta fuertemente en la cantidad y tipo de nutrientes que se liberan hacia los tejidos periféricos (Nozière et al., 2010). En la ganadería Rubio se utilizan diversas variedades de trigo, heno, soja y maíz, pese a ello en ciertas épocas del año estos alimentos escasean y se vuelven insumos costosos para la empresa, provocando que en algunas ocasiones exista una baja en el peso y características de los animales que exportan; cómo una medida para contrarrestar la escasez de comida para el ganado se propuso establecer una cosecha de sorgo personal en un territorio cercano a la ganadería. Considerando que la inclusión de sorgo en las dietas para engorda de bovinos en México es alta y que de acuerdo con Peel et al. (2011), en 2011 se engordaron alrededor de 780,000 cabezas en México, esto representa una gran cantidad de grano de sorgo utilizado en las dietas. De acuerdo con Calderón et al. (2011) la calidad del sorgo es variable, y en función a su variedad es su valor nutricional. En la alimentación de ganado bovino es fundamental conocer la calidad de un ingrediente, ya que de esa manera se puede

estimar la productividad en el animal. Las siembras de sorgo serán utilizadas para alimentar a los animales de la Ganadera Rubio cómo una medida para verificar si existe una mejora respecto a los índices de calidad y peso tomando como base las propiedades benéficas del sorgo, a continuación, se presenta el pesaje de una muestra de ganado comprendida por reses y cerdos anterior al nuevo régimen de alimentación ilustradas en las tablas 3 y 4.

Tabla 3. Pesaje de reses criados en el periodo Enero a Febrero del 2023.

| Ejemplar de Res Criado entre Enero-Febrero 2023 |        |        |        |        |        |        |        |       |               |
|-------------------------------------------------|--------|--------|--------|--------|--------|--------|--------|-------|---------------|
| Edad en Semanas                                 | Res 1  | Res 2  | Res 3  | Res 4  | Res 5  | Res 6  | Res 7  | Res8  | Peso promedio |
| 15                                              | 93.19  | 92.36  | 90.31  | 90.85  | 101.70 | 94.93  | 99.84  | 86.94 | 93.71         |
| 16                                              | 95.98  | 95.13  | 93.02  | 93.58  | 104.76 | 97.78  | 102.37 | 89.55 | 96.52         |
| 17                                              | 99.34  | 98.46  | 96.28  | 96.85  | 108.42 | 101.20 | 105.95 | 92.68 | 99.90         |
| 18                                              | 103.51 | 102.59 | 100.32 | 100.92 | 112.98 | 105.45 | 110.40 | 96.58 | 104.09        |

Tabla 4. Pesaje de Cerdos criados en el periodo Enero a Febrero del 2023.

| Ejemplar de Cerdo Criado entre Enero-Febrero 2023 |         |         |         |         |         |         |         |         |               |
|---------------------------------------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------------|
| Edad en Semanas                                   | Cerdo 1 | Cerdo 2 | Cerdo 3 | Cerdo 4 | Cerdo 5 | Cerdo 6 | Cerdo 7 | Cerdo 8 | Peso promedio |
| 15                                                | 58.3    | 56.7    | 59.2    | 56.47   | 58.13   | 57.2    | 56.5    | 59.8    | 57.78         |
| 16                                                | 72.8    | 70.8    | 74      | 70.5    | 72.69   | 71.5    | 70.62   | 74.75   | 72.2          |
| 17                                                | 88.1    | 85.75   | 89.54   | 85.30   | 87.5    | 86.55   | 85.45   | 90.44   | 87.32         |
| 18                                                | 105.8   | 102.9   | 107.44  | 102.36  | 105.45  | 103.18  | 102.54  | 108.53  | 104.77        |

## Aplicación de la regresión múltiple

Es necesario que se cumplan una serie de supuestos básicos para que sea viable la aplicación de una regresión lineal, en consecuencia, se evaluaron los supuestos básicos del modelo de regresión (linealidad, independencia, normalidad,

homocedasticidad y no-colinealidad) para confirmar la pertinencia de llevar a cabo el modelo estadístico propuesto Kerlinger y Lee (2002).

Para poder estimar el ingreso promedio futuro a obtener por cada animal exportado en la ganadería se utilizó un modelo de dos variables: la variable dependiente y la variable independiente. El precio promedio cotizado por cada ejemplar se estableció como la variable dependiente la cual estaría en función del tiempo, expresado en meses. Es decir, la variable dependiente del modelo son los precios y la variable independiente es el tiempo.

### **Precio promedio de cada ejemplar (Y) = F(tiempo (x) dado en meses**

#### *Ec.1 Planteamiento del modelo de Regresión Lineal Simple*

El modelo utilizado quedó expresado de la siguiente manera, siendo esta la Función de Regresión Poblacional (FRP) del modelo:

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

Figura 1. Ec.2 Función de Regresión Poblacional. Fuente González (2017).

A partir de la FRP, se estima la Función de Regresión Muestral (FRM), la cual, sirvió para la estimación del modelo de regresión y que se definió a como se muestra a continuación:

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$$

Figura 2. Ec.3 Función de Regresión Poblacional. Fuente González (2017).

Donde:

Yi: El precio mensual estimado

B1: Constante de la función

B2: El coeficiente estimado de la pendiente de la función

Xi: El tiempo expresado en meses

Ui: Término de perturbación

Las fórmulas de regresión se aplicarán para lograr calcular el valor de los ingresos obtenibles a futuro con relación a la nueva metodología de trabajo interno con enfoque sustentable comprobando la rentabilidad de la ganadería y cómo está se desempeñará en meses y/o años próximos con estos cambios en el sistema.

## Resultados y discusión

Se registraron los pesajes de una muestra de ganado de res, posteriormente se realizó el mismo procedimiento con ejemplares de cerdo donde se obtuvo una media de peso aproximado y se comparó con resultados anteriores de ejemplares que no se les daba este tratamiento alimenticio para verificar si existía un incremento significativo respecto al peso y la cantidad de carne aprovechada; igualmente se repitió el proceso de evaluación en la calidad final para comparar las calificaciones obtenidas y verificar si el uso del sorgo como fuente de alimento ayudo a mejorar la calidad de la carne. Los resultados se exponen en las Tablas 5, 6, 7 y 8 respectivamente.

| Evaluación final de la calidad (Res) |       |       |       |       |       |       |       |       |          |
|--------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| Variable                             | Res 1 | Res 2 | Res 3 | Res 4 | Res 5 | Res 6 | Res 7 | Res 8 | Promedio |
| CM                                   | 1     | 1     | 1     | 2     | 1     | 2     | 1     | 1     | 1.25     |
| RA                                   | 8     | 9     | 10    | 8     | 9     | 10    | 9     | 9     | 9        |
| CE                                   | 9     | 8     | 9     | 8     | 10    | 9     | 9     | 8     | 8.75     |
| CC                                   | 8     | 9     | 9     | 8     | 9     | 8     | 8     | 9     | 8.5      |
| CSA                                  | 9     | 8     | 9     | 8     | 9     | 8     | 10    | 8     | 8.625    |
| CSE                                  | 10    | 10    | 8     | 9     | 8     | 9     | 8     | 8     | 8.75     |
| CN                                   | 9     | 8     | 10    | 9     | 8     | 8     | 9     | 9     | 8.75     |

Tabla 5. Evaluación final de la calidad en la Res.

| Evaluación final de la calidad (cerdos) |       |       |       |       |       |       |       |       |          |
|-----------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| Variable                                | Res 1 | Res 2 | Res 3 | Res 4 | Res 5 | Res 6 | Res 7 | Res 8 | Promedio |
| CM                                      | 1     | 2     | 2     | 1     | 2     | 1     | 2     | 1     | 1.5      |
| RA                                      | 9     | 9     | 9     | 9     | 8     | 9     | 9     | 10    | 9        |
| CE                                      | 8     | 8     | 8     | 8     | 9     | 8     | 10    | 9     | 8.5      |
| CC                                      | 9     | 8     | 9     | 9     | 10    | 9     | 8     | 9     | 8.875    |
| CSA                                     | 9     | 9     | 9     | 9     | 10    | 10    | 9     | 10    | 9.375    |
| CSE                                     | 10    | 9     | 9     | 10    | 8     | 9     | 9     | 9     | 9.125    |
| CN                                      | 9     | 10    | 10    | 9     | 9     | 9     | 10    | 8     | 9.25     |

Tabla 6. Evaluación final de la calidad en los cerdos.

| Ejemplar de Res Criado entre Agosto-Septiembre 2023 |        |        |       |        |        |        |        |       |               |
|-----------------------------------------------------|--------|--------|-------|--------|--------|--------|--------|-------|---------------|
| Edad en Semanas                                     | Res 1  | Res 2  | Res 3 | Res 4  | Res 5  | Res 6  | Res 7  | Res 8 | Peso promedio |
| 15                                                  | 95.18  | 101.61 | 89.34 | 90.44  | 101.46 | 96.49  | 100.98 | 86.28 | 95.22         |
| 16                                                  | 98.21  | 104.86 | 92.20 | 93.34  | 104.74 | 99.57  | 104.22 | 89.04 | 98.27         |
| 17                                                  | 101.84 | 108.74 | 95.61 | 96.79  | 108.58 | 103.26 | 108.07 | 92.33 | 101.90        |
| 18                                                  | 106.32 | 113.53 | 99.82 | 101.05 | 113.36 | 107.80 | 112.83 | 96.40 | 106.39        |

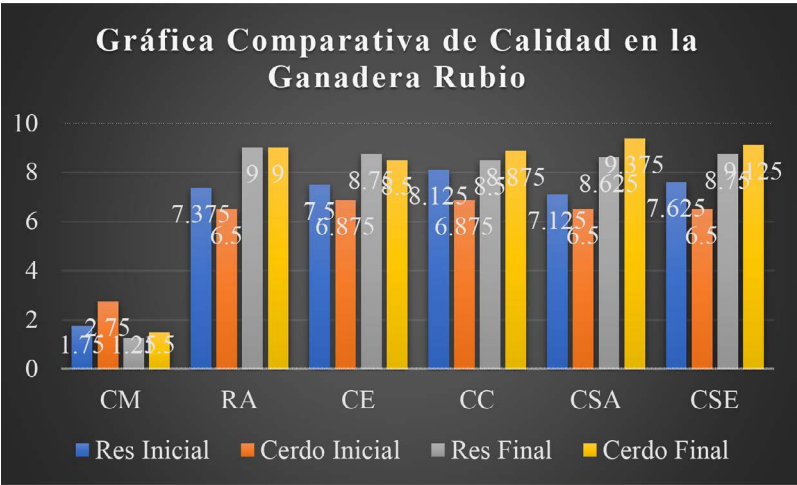
Tabla 7. Pesaje de res criados entre agosto y septiembre del 2023.

| Ejemplar de cerdo criado entre agosto y septiembre 2023 |         |         |         |         |         |         |         |         |               |
|---------------------------------------------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------------|
| Edad en semanas                                         | Cerdo 1 | Cerdo 2 | Cerdo 3 | Cerdo 4 | Cerdo 5 | Cerdo 6 | Cerdo 7 | Cerdo 8 | Peso promedio |
| 15                                                      | 58.3    | 55.4    | 58.32   | 56.21   | 57.43   | 58.67   | 55.42   | 58.38   | 57.26         |
| 16                                                      | 74.01   | 70.38   | 74.06   | 71.38   | 72.93   | 74.51   | 70.38   | 74.14   | 72.35         |
| 17                                                      | 91.81   | 87.24   | 91.33   | 88.98   | 90.44   | 92.39   | 87.27   | 91.93   | 90.15         |
| 18                                                      | 113.8   | 108.8   | 113.6   | 109.76  | 112.65  | 114.6   | 108.58  | 114.8   | 111.42        |

Tabla 8. Pesaje de cerdos criados entre agosto y septiembre del 2023.

Los resultados obtenidos respecto a la calidad fueron plasmados en la figura 3:

Figura 3. Gráfica comparativa de calidad.

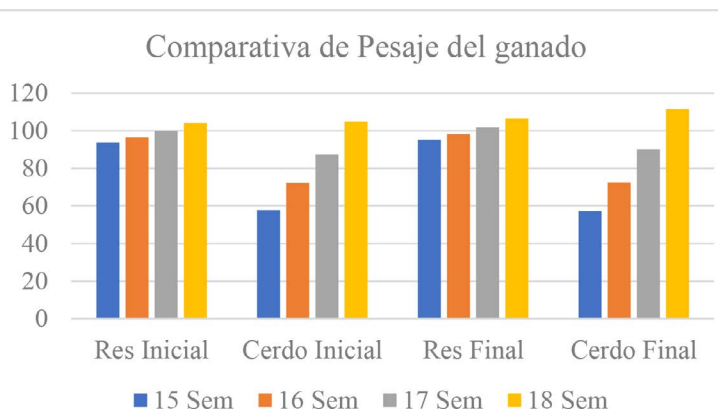


Fuente: elaboración propia.

Existe una diferencia notoria respecto a los indicadores de calidad, el uso del sorgo como fuente principal de alimentación incremento cada uno de los aspectos evaluados en el periodo de febrero, el único indicador con calificación disminuida resulto la carga microbiana pero contrariamente es el mejor indicativo para afirmar que los animales cuentan con una mejor condición de vida y características óptimas para su consumo, la coloración de la carne, sus cualidades

nutricionales y sensoriales resultaron demasiado elevadas prácticamente siendo casi impecables; debido a esto se ha provisto un impacto significativo en el índice de carne que es aprovechable, generalmente solo el sesenta por ciento de la carne en los animales de las ganaderas es utilizado para la distribución, principalmente debido a ciertas deficiencias y áreas que no obtienen el certificado de calidad, pero en este caso ahora se está efectuando un mayor nivel del aprovechamiento por ejemplar, implicando índices de rentabilidad superiores con la misma cantidad de reses. En la figura 4 se ilustra la comparativa de los pesajes en el ganado.

Figura 4. Gráfica comparativa de pesaje.



Fuente: Elaboración propia.

Los indicadores de pesaje no representan una diferencia tan notoria cómo ocurrió en los medidores de calidad pero existen dos aspectos remarcables que convierten estos avances ligeros de peso en un valor más significativo, primero cabe destacar que debido a los mejores resultados obtenidos en los indicadores de calidad el aprovechamiento de la carne es más elevado volviendo relevante cualquier gramaje superior a periodos de tiempo pasado, en segundo plano se debe aclarar que los especímenes tomados para la muestra son algo más jóvenes a la media empleada para fabricación de productos, en este sentido cabe mencionar que se pueden obtener mejores resultados con la parte del ganado que tenga mayor longevidad, se utilizó una principalmente una muestra joven para fines de optimización de tiempos al ser estos más propensos a demostrar un cambio antes que un miembro en etapa adulta.

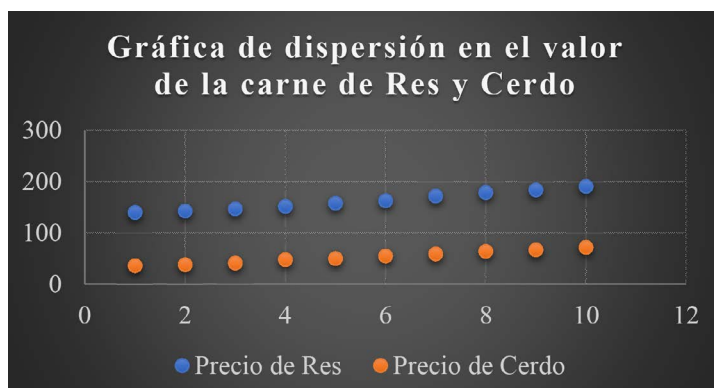
En lo que se refiere al precio promedio del kilo de cerdo oscila entre un mínimo de 36 y un máximo de 72 pesos mexicanos mientras que el kilo de res oscila entre un mínimo de 140 y un máximo de 196 pesos en promedio teniendo un rango de 36 y 56 pesos respectivamente, así mismo se observó un valor promedio de 162.9 para el kilo de res y 53 para el kilo de cerdo. Estos valores se aprecian más claramente en la tabla 9.

Tabla 9. Valores estadísticos de los precios de la carne de res y cerdo.

| Estadísticos | Valor Res | Valor Cerdo |
|--------------|-----------|-------------|
| Mínimo       | 140       | 36          |
| Máximo       | 196       | 72          |
| Rango        | 12        | 8           |
| Promedio     | 162.9     | 53          |

Igualmente, cuando se relaciona la variable de precio con el tiempo puede apreciarse una dispersión en los datos donde se puede enfocar una relación lineal en las variables apreciadas en la figura 5.

Figura 5. Gráfica de dispersión del valor de la carne de res y cerdo.



Fuente: Elaboración propia.

Existe una relación lineal positiva notoria en la figura anterior, se espera que el coeficiente beta de la pendiente de la ecuación adquiera un valor superior a cero, es decir, que sea positiva, al realizar el cálculo del valor de la pendiente se obtuvo la cifra de 2.56, el valor de la pendiente indica la variabilidad que tiene el precio en relación con el tiempo manteniéndose constante.

Una vez se ha comprobado el movimiento lineal de la variable es necesario determinar cómo afectará esta vertiente a los fines lucrativos de la empresa Ganadera Rubio, haciendo uso del modelo de ecuación lineal propuesto en este ejercicio y tomando como referencia los datos recopilados por el personal de contabilidad y área financiera en la empresa se registró el coste promedio de mantenimiento por cada especie y los costes de venta a futuro derivados de la dispersión planteamiento de regresión lineal simple, estos resultados se abordan en la figura 6.

Donde:

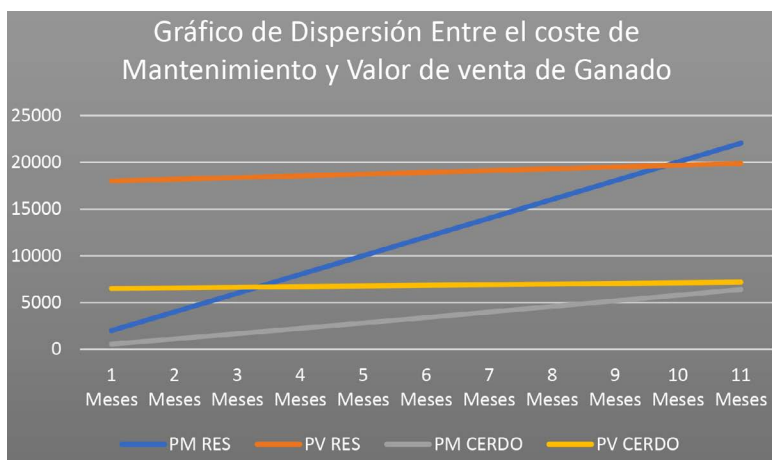
**PM Res:** Precio de Mantenimiento Mensual por Res

**PV Res:** Precio de venta de Res por Kilo

**PM Cerdo:** Precio de Mantenimiento Mensual por Cerdo

**PV Cerdo:** Precio de Mantenimiento Mensual por Cerdo

Figura 6. Gráfica de dispersión coste-mantenimiento.



Fuente: Elaboración propia.

Con base en la información recopilada a lo largo de esta investigación utilizando tanto las fórmulas de regresión lineal, la información proporcionada por la Ganadera y los nuevos datos generados a partir de las mejoras en los indicadores de calidad contemplando un rendimiento superior de la capacidad cárnica se obtiene la última gráfica de dispersión que en síntesis refleja el punto de equilibrio o puede considerarse también el margen de utilidad por cada ejemplar de res y cerdo en la empresa, los costes de mantenimiento mensual en relación

con el valor promedio obtenible por espécimen considerando la inflación y la reducción de costos por mantenimiento nos indican que en el área de la res estás presentan una permanencia promedio de nueve meses en la ganadera antes de que su mantenimiento deje de ser redituable, mientras en el área de los cerdos estos tienen una estancia algo más prolongada llegando hasta prácticamente el año de mantenimiento considerando su aumento por masa, asimismo esto no presenta ninguna dificultad o preocupación económica debido a que las medidas implementadas durante el proceso agilizarán la comercialización del ganado traducido en mayores utilidades para la empresa y reduciendo significativamente el impacto ecológico que se pueda presentar en años futuros.

## **Conclusiones**

Alcanzar la sustentabilidad es el hito al que toda organización aspira, el impacto medioambiental que tienen las empresas es mayor año con año y esto conduce a seguir optando más ecológicas al momento de estructurar el plan de negocio en una organización. El Modelo de regresión lineal es una metodología útil aplicable a cualquier organización que esté orientada a tener un mayor índice de calidad, la investigación aplicada y la opinión general de los clientes manifiestan que la cadena de producción en ganadería rubio ha mejorado, en función de que ahora existe un mayor control sobre la cantidad de productos elaborados diariamente, el estándar de calidad logra mantenerse respecto a los regímenes actuales, se encontró un nicho de mercado dónde la ganadería puede desempeñarse mejor y cuyo nivel de competencia es relativamente bajo en el área de los embutidos, un mayor aprovechamiento de las condiciones del entorno permitió favorecer a una empresa que estaba cómodamente posicionada en el mercado pero ahora también tiene mayor cadena de productos y distribución.

Es importante revisar todos los nutrientes (vitaminas, minerales, energía, proteínas, FND) en las diferentes etapas del desarrollo del ganado en el corral. Se debe aprender a manejar animales alimentados con cereales y tomar decisiones basadas en aspectos biológicos y económicos, teniendo en cuenta aspectos de bienestar animal.

## Referencias

- BETTINI, V. (1998). *Elementos de ecología urbana*. Valladolid: Trotta.
- Calderón, S. F., Mendoza, M. G. D., Hernández, G. A., Muñoz, O. A., Ramírez, V. G., Plata, P. F. X. (2011). *Concentración de nutrientes y rendimiento de diez variedades de sorgo (Sorghum bicolor L. Moench) usados para alimentación de rumiantes*. Universidad y Ciencia, 27: 169-177
- Díaz Fernández, M., & Llorente Marrón, M. D. (2013). *Econometría* (Cuarta ed.). Madrid, España: Ediciones Pirámide. Recuperado el 9 de noviembre de 2023, de <http://site.ebrary.com/lib/bibliotecaunansp/reader.action?docID=11072811&page=36&ppg=36>
- Fernández. (2001). “Centro de investigación y Tecnología de Carnes, CITECA del sistema INTI: Limpieza y desinfección en la industria cárnica”. Buenos Aires. Argentina.
- Gómez. (2021). “Técnicas estadísticas multivariantes para valorar la satisfacción de clientes” Recuperado el 28 de noviembre del 2023 en: [https://tauniversity.org/sites/default/files/articulo\\_tecnicas\\_estadisticas\\_multivariantes\\_para\\_valorar\\_la\\_satisfaccion\\_de\\_clientes.pdf](https://tauniversity.org/sites/default/files/articulo_tecnicas_estadisticas_multivariantes_para_valorar_la_satisfaccion_de_clientes.pdf)
- Gonzalez. (2017). “Aplicación del análisis de regresión lineal simple para la estimación de los
- Gujarati, D. N., & Porter, D. C. (2010). *Econometría* (Quinta ed.). México, D.F., México: Mc Graw Hill.
- Gutiérrez P. (2009). Desarrollo sustentable visto por un ecólogo: Mitos, controversias y futuro (primera parte), *Derecho Ambiental y Ecología*, No. 29, p.p. 67-63
- INEGI. (2009). Instituto Nacional de Estadística Geografía e Informática. *Indicadores de desarrollo sustentable en México*. México
- KERLINGER, F. y LEE H. (2002). *Investigación del comportamiento. Método de investigación en Ciencias Sociales*. México: Mc Graw Hill
- Legarreta. (2019). “Aplicación de un análisis clúster para el estudio de la segregación social en el municipio de Bilbao,” *Boletín la Asoc. Geógrafos Españoles*, no. 81, pp. 1–35, doi: 10.21138/bage.2763.
- Leitão. (1999). “Microbiología da carne e produtos carneos”. *Secretaria de Agricultura e Abastecimento*. São Paulo, Brasil.
- Leopoldo. 2005. “La escasez, el costo y el precio del agua en México: Its Cost and Price”. *Economía UNAM* 2(6):24-42.
- MONTERO, I. y LEÓN, O. G. (2002). “Clasificación y descripción de las metodologías de investigación en psicología”. *International Journal of Clinical and Health Psychology*, 5: 115-127.

- MONTOSA, J. (2014). “Aplicación del análisis multivariante a espacios en transformación: Las periferias de las mayores aglomeraciones urbanas andaluzas”. *Boletín de la Asociación de Geógrafos Españoles*, 65: 87-112.
- Nozière, P., Ortigues-Marty, I., Loncke, C., Sauvant, D. (2010). Carbohydrate quantitative digestion and absorption in ruminants: from feed starch and fibre to nutrients available for tissues. *Animal*, 4: 1057–1074.
- Peel, D. S., Mathews, K. H., Johnson, R. J. (2011). *Trade, the expanding Mexican beef industry, and feedlot and stocker cattle production in Mexico*. ldp-m-206-01, U. S. Department of Agriculture, Economic Research Service. Disponible en línea: Recuperado el 8 de noviembre del 2023 en: <http://www.ers.usda.gov/media/118317/ldpm20601.pdf>.
- Precios de las acciones de Facebook, Inc.
- Ramos-Herrera S., Bautista-Margulis R. y ValdezManzanilla A. (2010). *Estudio estadístico de la correlación entre contaminantes atmosféricos y variables meteorológicas en la zona norte de Chiapas, México*. Universidad y Ciencia 26 (1), 65-80.
- Rodríguez. (2022). “A students academic index using multivariate techniques” HAL Id: hal-03551705,” HAL open Sci., 2022.
- Valderrama. (2007). Reflexiones en torno a la responsabilidad social empresas empresarial [Resumen]. *Teoría y Praxis*, ISSN 1870-1582, N°. 3, 2007, 128, 125-134.

# Capítulo 6

---

## **Metodología para simulaciones de liberación de sustancias peligrosas empleando el programa Aloha**

*Marco Arturo Arciniega Galaviz*

## Introducción

Un riesgo ambiental se define como la probabilidad de que ocurran accidentes en las que intervengan sustancias químicas peligrosas y que puedan trascender los límites de sus instalaciones y afectar adversamente a la población, los bienes, el ambiente y los ecosistemas (SEMARNAT, 2008). Es por eso que las empresas deben implementar mecanismos, procedimientos, instalar equipos, capacitar al personal en el manejo de materiales peligrosos para disminuir la probabilidad de que ocurra algún evento catastrófico.

Algunas de las sustancias químicas peligrosas que comúnmente son usadas en las industrias se mencionan la tabla 1. Las empresas que las utilizan deben de conocer los efectos a la salud y medioambiente que provocaría una liberación, saber cuál es el alcance de los daños y por lo tanto implementar dispositivos de seguridad, procedimientos y capacitar al personal para evitar algún evento catastrófico.

Tabla 1. Sustancias químicas peligrosas, usos y efectos a la salud.

| Sustancia | Usos                                                                                                                                                                                                                                                                  | Efectos a la salud                                                                                                                                                                                                                                                                                                                                                                                                                        |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Amoniaco  | En las últimas décadas se sabe que casi el 90 % del amoniaco producido en el mundo se usa para la elaboración de fertilizantes alrededor de todo el planeta. Es usado como gas refrigerante en aparatos de aire acondicionado y en fábricas de hielo (Fuideco, 2019). | Puede causar irritación intensa y quemaduras en la piel y los ojos con la posibilidad de daño ocular. Si se inhala puede irritar la nariz y la garganta, irritar el pulmón, causando tos o falta de aire. La exposición más alta puede causar asfixia por acumulación de líquido en el pulmón (edema pulmonar), que es una emergencia médica. El contacto con el amoníaco líquido puede causar lesiones por congelación (NJHealth, 2016). |

| Sustancia   | Usos                                                                                                                                                                                                                                                          | Efectos a la salud                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Cloro Gas   | Se utiliza principalmente para esterilizar el agua potable y el agua de las piscinas. En la industria química con aplicaciones para el blanqueado (Bleaching) en la industria del papel, desinfección, tratamiento de aguas y plásticos (Lana Sarrate, 2020). | Puede producir irritación de la nariz, la garganta y los ojos. La exposición a niveles más altos puede producir tos y alteraciones del ritmo respiratorio y daño de los pulmones.                                                                                                                                                                                                                  |
| Gas LP      | Se usa principalmente como combustible en todo tipo de cocinas domésticas y profesionales, de igual manera se utiliza como combustible para eliminar humedad de granos, frutas y hortalizas.                                                                  | Por inhalación: los efectos por respirar GAS LP son, dolor de cabeza, náuseas, vómitos, tos, mareos, somnolencia, desorientación y en, casos extremos, inconsciencia, incluso la muerte. Si el gas entra en contacto con los ojos provoca congelamiento, hinchazón y daño ocular. Por ser un combustible puede producir incendios o explosiones con fuego (Gobierno de la Ciudad de México, s.f.). |
| Gas Natural | Se utiliza como combustible, tanto en el ámbito doméstico como en el industrial. Se usa como materia prima para la obtención de propano, amoníaco, metanol, etileno o butadieno. También como combustible para generar electricidad (Suroeste Grupo, 2023).   | Una gran cantidad de gas natural o metano desplaza el aire, la falta de oxígeno podría provocar asfixia. Por ser un gas altamente inflamable puede producir incendios o explosiones.                                                                                                                                                                                                               |
| Gasolina    | Es una fuente importante de combustible para propulsar vehículos, botes, motocicletas, tractores y algunas aeronaves (Pelayo, 2001).                                                                                                                          | Se evapora fácilmente, es muy inflamable y puede formar explosivas mezclas en el aire. El fuego y la explosión son los principales peligros físicos de la gasolina. La gasolina es altamente inflamable.                                                                                                                                                                                           |

Existen simuladores que dan como resultado las áreas de afectación en caso de la liberación de una sustancia química peligrosa, en base a las características fisicoquímicas, condiciones meteorológicas que prevalecen en el lugar de la emisión, las características de la fuente, por ejemplo, tipo de tanque de almacenamiento, volumen y masa, el estado físico de la sustancia química, la presión y temperatura a la que se encuentra dentro de este y la forma de su liberación.

Los resultados de la simulación también son parte del estudio de riesgo ambiental que tiene que presentar la empresa a las autoridades competentes para que sea autorizada su construcción y operación.

En el periodo de 2007-2017, en México se ingresaron para su evaluación y dictaminación 4,262 estudios de riesgo ambiental, de los cuales el 28 % fueron de industrias del petróleo, 15 % del giro alimenticio, 8.5 % industrias químicas, 5.2 industrias del gas LP y 3.8 % industrias mineras (Semarnat, 2018).

Como parte de un estudio de riesgo ambiental, se debe mencionar los riesgos potenciales de accidentes ambientales, para esto se debe de realizar una modelación de él o los eventos probables máximos de riesgo y hacer la descripción de riesgos que tengan afectación potencial al entorno de la planta. La metodología presentada en este capítulo de libro describe los pasos para realizar la modelación (simulación) de un evento para determinar los radios de afectación en caso de una emisión de sustancias químicas y cuáles serían sus efectos negativos al entorno de la empresa.

Se han presentado eventos que han cobrado vida en Sinaloa y algunos lugares de México al manejar amoniaco anhidro, en el estado de Sinaloa Municipio de Guasave se registró una fuga de 5 toneladas de amoniaco, en el transcurso de la mañana se presentó un reporte a las autoridades, ya que lograba verse desde la carretera Internacional México 15, se tuvieron que evacuar a más de 3,604 persona ya que el gas formo una nube muy extensa sobre las comunidades cercanas, para evitar que ocurriera una tragedia sobre la población (Torres, 2018).

En Culiacán, Sinaloa, se registró un accidente a las 4:30 de la mañana en una espuela de ferrocarril —desviación de la vía de una fábrica—, donde se abastece la empresa de Fertilizantes Agrícolas, S.A. ubicada en el ejido Bario-meto, 20 Km al poniente de la capital del estado. Según investigaciones el carro de cisterna que descargaba el gas en un furgón de ferrocarril presentaba una fisura provocada por un golpe en la costura de los 20 centímetros, por donde libero grandes volúmenes de amoniacos. El hecho tuvo consecuencias graves ya que se presentaron dos muertos, 58 intoxicados y 800 familias desalojadas de nuevas colonias y tres núcleos ejidales fue el saldo que dejo la fuga de gas de amoniaco (Martínez, 2006).

Una mujer de 26 años, trabajadora de industrias frutícola, sin antecedentes mórbicos, accidentalmente se encontró expuesta a altas concentraciones de  $\text{NH}_3$  (sobre 1.720 ppm) por 15 min al romperse un tubo de refrigeración. Es rescatada e intubada en el lugar del accidente y trasladada al hospital, presentando insuficiencia respiratoria aguda grave con obstrucción de la vía aérea alta y baja, insuficiencia renal aguda. A los dos años que se registró el accidente inhalatorio, la paciente presenta disnea a mínimos esfuerzos, insuficiencia respiratoria crónica parcial con necesidades de oxigenoterapia permanente y colonización de la vía aérea por pseudomonas aeruginosas (Gutiérrez, 2014).

En base a lo anterior, es importante realizar un estudio de riesgo principalmente cuando se manejen sustancias químicas peligrosas con el objetivo de determinar los efectos que se tendría a las personas, instalaciones o edificios en caso de la liberación de estas.

La Ley del Equilibrio Ecológico y Protección al Ambiente del estado de Sinaloa (LEEPA), en el artículo 34, menciona que las empresas que presenten una manifestación de impacto ambiental, deberán ir acompañada de un estudio de riesgo de la obra o actividad y de sus modificaciones, consistente en las medidas técnicas preventivas y correctivas para mitigar los efectos adversos al equilibrio ecológico durante su ejecución y operación normal conforme a la Ley en caso de accidente. Y en el artículo 120, “los responsables de los establecimientos en los que se realicen actividades riesgosas, están obligados a elaborar, actualizar y someter a la autorización de la Secretaría, sus programas para la prevención de accidentes en el desarrollo de dichas actividades para lo cual la Secretaría solicitará la intervención de otras dependencias estatales, que, de acuerdo con sus atribuciones, deban participar en la autorización antes referida”.

La evaluación del riesgo es el uso de los datos y observaciones científicas para definir los efectos para la salud o los ecosistemas causados por la exposición a materiales o situaciones peligrosas (NAS, 1983).

Como parte del estudio de riesgo ambiental, se debe de realizar simulaciones en caso de la liberación de sustancias peligrosas y conocer cuáles son sus efectos a las personas y medio ambiente, una de las herramientas utilizadas es el programa Aloha de la Environmental Protection Agency (EPA), el cual permite ingresar detalles sobre una liberación química real o potencial y luego generará estimaciones de zonas de amenaza para varios tipos de peligros. Aloha puede modelar nubes de gases tóxicos, nubes de gases inflamables, BLEVE (explosiones de vapor en expansión de líquidos en ebullición), incendios de chorros, incendios de piscinas y explosiones de nubes de vapor (EPA, 2022).

Existen investigaciones para determinar los impactos que pudieran generar eventos en empresas donde se manejan sustancias químicas peligrosas (tabla 2).

Tabla 2. Investigaciones donde fue empleado el programa Aloha como un instrumento de simulación.

| <b>Autores</b>                                                                                          | <b>Nombre de la Investigación</b>                                                                                                                           | <b>Resultados</b>                                                                                                                                                                                                                                                                                                                                                                                                                      |
|---------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Soto Pacheco Samantha Abigail, Alcántara Garduño Martha Elena (2022)                                    | Determinación de consecuencias ocasionadas por la explosión de nitrato de amonio en Celemania, Coahuila.                                                    | Se realizó la simulación de la dispersión de gases liberados de 25 toneladas de nitrato de amonio ( $\text{NH}_4\text{NO}_3$ ), la zona de amenaza resultó de 5 kilómetros. Se propusieron los requisitos y acciones inmediatas que debían llevar a cabo los cuerpos de primera respuesta para la atención del incidente.                                                                                                              |
| Luis Carlos Valdés L., Gustavo Adolfo Ortega B., Jesús María Consuegra G. (2019)                        | Mapa de riesgos tecnológicos asociados al manejo de mercancías peligrosas en los puertos de Barranquilla                                                    | Los autores realizaron una simulación de incendio de líquido inflamable en almacenamiento del puerto. La cantidad total de líquido quemado es de 134.861 kg durante 1 hora de conflagración. El incendio generó radiación térmica en un área radial de 127,0 m, desde el foco del siniestro, en donde la radiación alcanza 2,0 kW/m <sup>2</sup> , esto implica quemaduras de 1er grado en personas expuesta a esta radiación térmica. |
| Catherine Andrea Parra Abril, Jonathan Fernando Ortiz Rey, Coautora, Sonia Lucia Meneses Velosa. (2015) | Uso de herramientas de ingeniería para la gestión de riesgo tecnológico en una central de abastecimiento agrícola con aglomeraciones permanentes en Bogotá. | Simulación Explosión Pipetas de Gas Oxietileno, de acuerdo a los resultados, se tienen una radiación térmica de 10 kW/m <sup>2</sup> a los 30 metros, esto significa que es potencialmente letal a las personas, a los 50 metros se tiene 5 kW/m <sup>2</sup> , significa quemaduras de segundo grado, y a los 75 metros se tiene 2kW/m <sup>2</sup> que implica dolor a las personas que estén dentro de este radio.                  |
| Micaela Belén del Solé (2019).                                                                          | Simulación de incidentes en el transporte de materiales peligrosos en Gran la Plata                                                                         | Se realizó una simulación de una emisión de metanol en una cantidad de 37,000 litros. Los resultados muestran que la zona de amenaza es de 600 metro desde el origen de la emisión. De acuerdo a los resultados las áreas de afectación abarcan escuelas, clubes y espacios recreativos, para reducir el impacto de los desastres se requiere de un análisis y de la mitigación del riesgo.                                            |

| Autores                           | Nombre de la Investigación                                                                                                             | Resultados                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Andrea León Sierra (2022).        | Modelización de accidentes en una refinería de Tarragona usando ALOHA                                                                  | Se realizó una simulación de BLEVE (explosión de vapores que se expanden al hervir el líquido) con propano en un volumen de 10,000 m3, los resultados muestran que el área de afectación con 5kW/m2 es de 1.7 millas, 0.6 millas se tiene una radiación térmica de 40 kW/m2 y 0.8 millas se tiene 20 kW/m2. Tras elaborar la simulación se aportan una serie de medidas para la prevención y protección contra estos accidentes, de forma que pueda evitarse que se produzca y en caso de que esto no sea posible, para reducir las consecuencias que pueden producirse sobre otros equipos o el personal presente en la zona. |
| Duche Chaquay Carlos Paul (2021). | Evaluación cuantitativa de explosión por riesgo químico en sistemas presurizados de la unidad separadora de aire de Linde Ecuador S. A | Se realizó una simulación de explosión de un tanque de oxígeno con capacidad de 250,000 litros. Dio como resultado una expansión del químico (oxígeno), de un diámetro de 7.24 km cerca de la zona de explosión, las personas que se encuentren en este rango tendrán síntomas como dolores de cabeza, irritación en los ojos, falta de aliento, posible riesgo de convulsión, los que se encuentren dentro de la zona naranja podrá sufrir de un leve dolor de cabeza.                                                                                                                                                        |

El objetivo de este capítulo de libro es mostrar paso a paso el uso del programa Aloha y pueda contribuir a que personas aprendan a realizar simulaciones que les permitirá ser parte de un estudio de riesgo el cual buscara disminuir la probabilidad de que ocurra un accidente que pudiera causar muerte de personas, dañar la salud desde leve a gravemente, así como proteger instalaciones y el medio ambiente.

## Metodología

Aprender a realizar modelaciones utilizando el programa Aloha de la EPA y que den como resultado áreas de afectación cuando se liberan sustancias químicas al ambiente que puedan dañar al entorno que rodean a la fuente de la emisión, permite que las empresas apliquen estrategias para evitar estos eventos que causarían efectos negativos a la salud de las personas y a los ecosistemas.

### *Revisión bibliográfica*

Con base en el objetivo de mostrar la metodología para realizar una simulación de la liberación de una sustancia química, primeramente se realizó una revisión bibliográfica realizando una búsqueda con las palabras clave: programa Aloha, hoja de seguridad del Amoniaco, condiciones atmosféricas del norte de Sinaloa, legislación mexicana en materia de estudio de riesgo, en sistemas de indización como Scielo, Redalyc, Redib y Doaj, páginas gubernamentales y en directorios y catálogos como Latindex y Google académico, principalmente en investigaciones realizados en México y países de Latinoamérica.

### *Descarga de simulador Aloha*

Posteriormente, el programa Aloha, se descargó de la siguiente liga <https://www.epa.gov/cameo/aloha-software>. Para mostrar la metodología de la simulación se fueron tomando capturas de pantalla para cada paso que se avanzaba y se describió cada una de las ventanas que aparecían, así como las distintas opciones que te ofrece el programa.

### *Información de datos atmosféricos*

Debido a que en Sinaloa la actividad predominante es la agricultura y el uso de amónico como fertilizante es en grandes cantidades, la simulación se realizó empleando datos atmosféricos del norte de Sinaloa. El amoniaco aporta nitrógeno como nutriente primario y de acuerdo con Escobedo (2021), en el distrito de riego 075, el cual es el más grande de México y se encuentra ubicado en áreas del municipio de Ahome, El Fuerte y Guasave, se aplican más de 50,000 toneladas de unidades de nitrógeno por ciclo en cultivo. Esto significa que el amoniaco es almacenado, transportado y usado en grandes cantidades en esta región del país.

Los datos que se investigaron previo a usar el programa se obtuvieron del Servicio Meteorológico Nacional (Conagua, 2023).

- Velocidad del viento: 3 m/s.
- Dirección del viento: Suroeste (SW).
- Nubosidad: Parcialmente nublado.
- Porcentaje de humedad: 75 %.
- Temperatura del aire: 30°C.

Para describir la fuente donde se libera el amoniaco, se ensayó con las siguientes características:

- Tanque cilíndrico horizontal.
- Dimensiones: 3 metros de diámetro, 10 metros de longitud y volumen de 70,686 litros.
- Estado de la sustancia dentro del tanque: Líquido.
- Temperatura del líquido dentro del tanque: -30°C.
- Masa del líquido dentro del tanque: 45,000 kilogramos.
- Tipo de falla en el tanque: Liberación de líquido no inflamable que escapa a atmósfera.
- Forma que mejor representa la liberación de la sustancia hacia fuera del tanque: circular.
- Diámetro de la abertura: 2 metros.
- Posición de la abertura en el tanque por donde se libera la sustancia: 1.5 metros

#### *Información de efectos a la salud del amoniaco*

Los parámetros de control de exposición para el amoniaco anhidro, se obtuvieron de las hojas de seguridad, ACGIH TLV: 25 ppm (17 mg/m<sup>3</sup>) TWA; 35 ppm (24 mg/m<sup>3</sup>) STEL OSHA PEL: 50 ppm (35 mg/m<sup>3</sup>) TWA.

#### *Mapa para plasmar los resultados*

Las zonas de afectación obtenidas fueron plasmadas sobre un mapa empleando el programa MARPLOT, lo cual permitió poder visualizar de una manera fácil las zonas de afectación y las diferentes concentraciones a distintas distancias.

#### *Interpretación de resultados:*

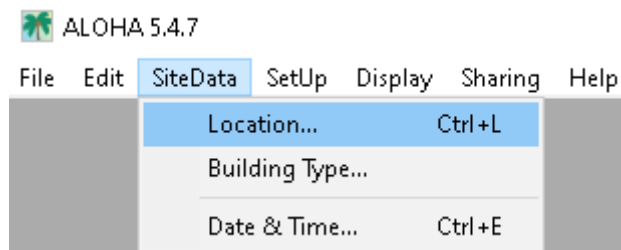
Una vez que se muestran los resultados sobre una imagen satelital, es posible ver de una manera precisa el área de afectación de la nube de amoniaco, se puede conocer la distancia en metros y sus respectivas concentraciones, y saber cuáles son los efectos a la salud de las personas que pudieran estar dentro de las zonas de amenaza dependiendo de la concentración o a la que están expuestos.

## **Resultados**

Los resultados se mostrarán señalando paso a paso la metodología para las simulaciones de liberación de amoniaco en el aire, la formación de un charco de fuego de etanol y un BLEVE con gas propano. Se mostrarán las ventanas que aparecen en el simulador conforme se va realizando la simulación.

Para realizar la simulación de liberación de amoniaco en el aire, primeramente, se selecciona la ciudad donde será el evento de emisión de la sustancia peligrosas (Figura 1).

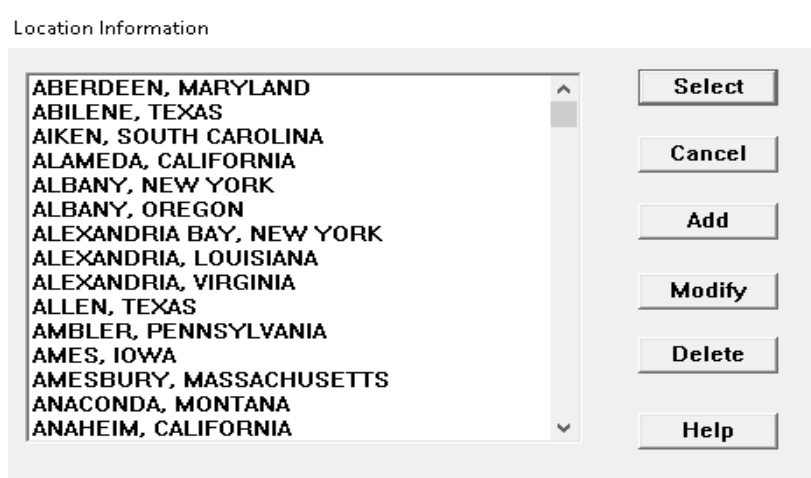
Figura 1. Datos del lugar de la simulación (localización, tipo de edificio y fecha y hora).



Fuente: Office of Emergency Management, EPA.

Por ser un simulador de los Estados Unidos de Norte América (EE. UU.), nada más te muestra ciudades de este país, te da la opción de agregar otros lugares que no sean de EE. UU., para esto se selecciona la opción “Add” (figura 2).

Figura 2. Nombres de ciudades donde se realiza la simulación.



Nota. La figura muestra opciones de agregar ciudades, cancelar, modificar y borrar. Fuente: Office of Emergency Management, EPA.

Se agrega el nombre del lugar donde se hará la simulación, para este ejemplo se escribió la ciudad de Los Mochis, la cual no es una ciudad de EE. UU., se encuentra aproximadamente a nueve metros sobre el nivel del mar (INEGI, 2021). Es importante seleccionar las unidades en las que se expresa la altura sobre el nivel del mar, ya sea en pies (ft) o metros (m), sus coordenadas geométricas son Latitud Norte: 25° 47' y Longitud Oeste: 108° 00' (Figura 3).

Figura 3. Agregar un lugar nuevo a la lista de localizaciones.

Location Input

**Enter full location name:**

Location is

---

**Is location in a U.S. state or territory ?**

☐ In U.S.    ☒ Not in U.S.

---

**Enter approximate elevation**

Elevation is  ☐ ft    ☒ m

---

**Enter approximate location**

|           | deg.                             | min.                            |                                                            |
|-----------|----------------------------------|---------------------------------|------------------------------------------------------------|
| Latitude  | <input type="text" value="25"/>  | <input type="text" value="47"/> | <input checked="" type="radio"/> N <input type="radio"/> S |
| Longitude | <input type="text" value="108"/> | <input type="text" value="00"/> | <input type="radio"/> E <input checked="" type="radio"/> W |

Nota. La figura muestra la ventana con datos del lugar nuevo que se agrega a la lista. Fuente: Office of Emergency Management, EPA.

Y finalmente se escribe el país donde se encuentra ubicada la localidad donde se hará la simulación y de igual manera se menciona la diferencia en horas de Greenwich Mean Time (GMT) y del tiempo de la localidad, siendo siete horas de diferencia (HORA.MX, 2023) (figura 4).

Figura 4. Información del nuevo lugar adicionado a la lista de localizaciones.

Foreign Location Input

Country name:

---

Offset from local STANDARD time to GMT:  hours  
(eastern hemisphere is a negative offset)

---

Is current model time standard or daylight savings time ?

☒ Standard Time    ☐ Daylight Savings Time

Fuente: Office of Emergency Management, EPA.

Una vez ingresados los datos de la localidad donde se realizará la simulación, aparecerá dentro del listado de ciudades y de esta manera se puede seleccionar (figura 5).

Figura 5. Selección de la localización donde se realiza la simulación.

Location Information

LODI, CALIFORNIA  
LONG BEACH, CALIFORNIA  
LORAIN, OHIO  
LOS ALAMITOS, CALIFORNIA  
LOS ALAMOS, NEW MEXICO  
LOS ANGELES, CALIFORNIA  
LOS GATOS, CALIFORNIA  
**LOS MOCHIS, SINALOA, MEXICO**  
LOUISVILLE, COLORADO  
LOUISVILLE, KENTUCKY  
LOVELAND, COLORADO  
MACON, GEORGIA  
MADISON, WISCONSIN  
MADISONVILLE, KENTUCKY  
MANCHESTER, IOWA

Fuente: Office of Emergency Management, EPA.

Después se describe el tipo de edificio donde se realizará la liberación del amoníaco. Para esta etapa de la simulación, te muestran diferentes tipos de edificios donde será el evento, por ejemplo, si es un lugar cerrado como una oficina, o si es un edificio de un solo piso o de dos pisos, en caso de conocer el número de cambio de aire por hora que se realiza en el lugar se puede agregar, y para completar el tipo de edificio debes seleccionar si está rodeado de árboles o arbustos o no existe ningún tipo de barrera (figura 6).

Figura 6. Opciones del tipo de edificación del lugar donde se realiza la simulación.

Infiltration Building Parameters

**Select building type or enter exchange parameter** Help

☐ Enclosed office building

☒ Single storied building

☐ Double storied building

☐ No. of air changes is  per hour

---

**Select building surroundings** Help

☒ Sheltered surroundings (trees, bushes, etc.)

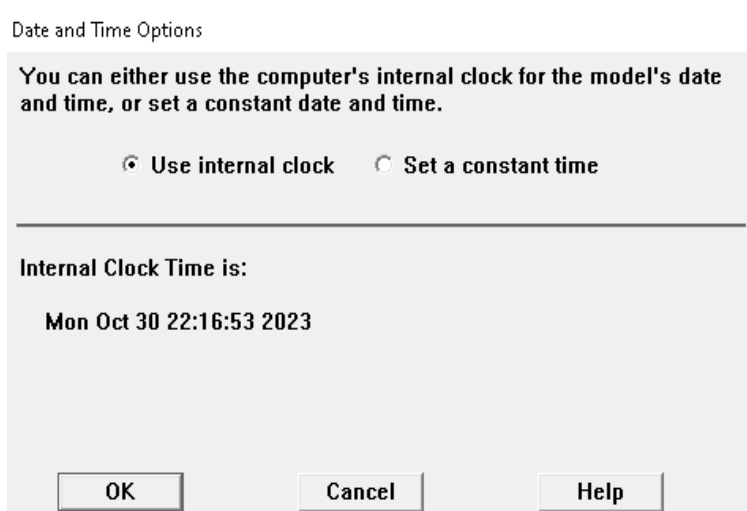
☐ Unsheltered surroundings

OK Cancel

Nota. En la figura se muestra las opciones para el tipo de edificio y los alrededores.  
Fuente: Office of Emergency Management, EPA.

Para la fecha y hora, se recomienda utilizar el reloj interno de la computadora (figura 7).

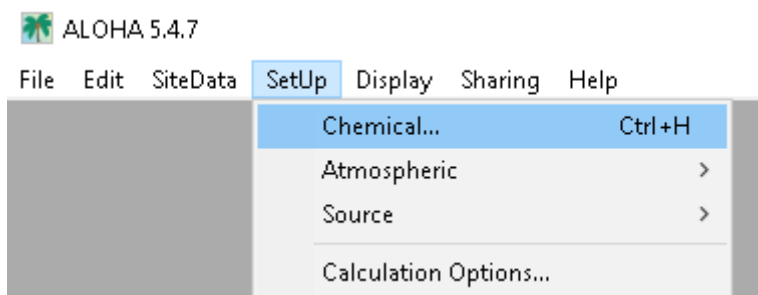
Figura 7. Opciones de seleccionar la fecha y hora de la simulación.



Fuente: Office of Emergency Management, EPA.

Posteriormente pasamos a la pestaña de “SetUp”, en la cual se muestran las opciones para seleccionar la sustancia química, las condiciones atmosféricas del lugar donde se realiza la simulación y los datos de la fuente donde se libera la sustancia química (Figura 8).

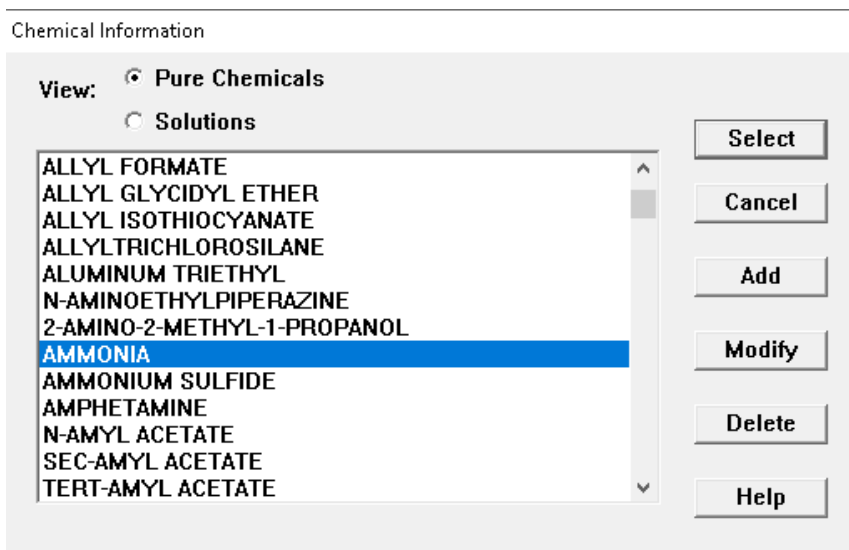
Figura 8. Configuración de la simulación (sustancia química, condiciones atmosféricas y fuente).



Fuente: Office of Emergency Management, EPA.

En la opción de “Chemical”, aparecerán una gran cantidad de sustancias químicas para ser seleccionada la de interés, para este trabajo se selecciona la sustancia “Ammonia”, inmediatamente el programa te muestra propiedades químicas y físicas del amoniaco. Aloha cuenta con una biblioteca de una amplia cantidad de sustancias químicas y sus propiedades físico-químicas las cuales utilizará para realizar los cálculos de la simulación (figura 9).

Figura 9. Selección de la sustancia química liberada.



Nota. En la figura se muestra el listado de sustancias químicas peligrosas que pueden ser seleccionadas para realizar la simulación. Fuente: Office of Emergency Management, EPA.

Para la selección de las condiciones atmosféricas del lugar donde será la simulación, te ofrece la opción que se ingrese de manera manual los datos atmosféricos o los datos que proporciona una estación meteorológica. Dentro de las condiciones atmosféricas que requiere el simulador se encuentra la velocidad del viento la cual puede expresarse en nudos (knots), millas por hora (mph), metros por segundo (metres/sec).

La dirección del viento es necesario ingresarla por sus iniciales en inglés, por ejemplo, suroeste (SW), sureste (SE), Noroeste (NW), noreste (NE), norte (N), sur (S), Oeste (W) Y ESTE (E). De igual manera se ingresa el dato de la altura a la cual se realizó las mediciones anteriores.

Es necesario indicar la nubosidad del lugar donde se realiza la simulación, las opciones son, completamente nublado, parcialmente nublado y despejado (figura 10).

Figura 10. Alimentación de las condiciones atmosféricas que prevalecen en el lugar de la liberación.

Atmospheric Options

Wind Speed is :  ☐ knots ☐ mph ☒ meters/sec

Wind is from :  Enter degrees true or text (e.g. ESE)

Measurement Height above ground is:

☒  ☐  OR ☐ enter value :  ☐ feet ☒ meters

---

Ground Roughness is :

☒ Open Country ☐ Urban or Forest OR ☐ Input Roughness [Z<sub>0</sub>] :

☐ Open Water

---

Select Cloud Cover :

☐  ☐  ☐ 

☐ complete ☐ partly ☐ clear

OR ☐ enter value :  [0 - 10]

Fuente: Office of Emergency Management, EPA.

La temperatura del aire puede proporcionarse en grados centígrados o grados Fahrenheit, una vez ingresados la velocidad y temperatura del aire, nubosidad, dirección del viento, el programa Aloha, se arroja la clase de estabilidad atmosférica que existe en el lugar de la simulación (figura 11).

Figura 11. Datos atmosféricos del lugar donde se realiza la simulación.

Atmospheric Options 2

Air Temperature is :  Degrees ☐ F ☒ C

Stability Class is :  ☐ A ☐ B ☐ C ☐ D ☒ E ☐ F

Inversion Height Options are :

☒ No Inversion ☐ Inversion Present, Height is :  ☒ feet ☐ meters

---

Select Humidity :

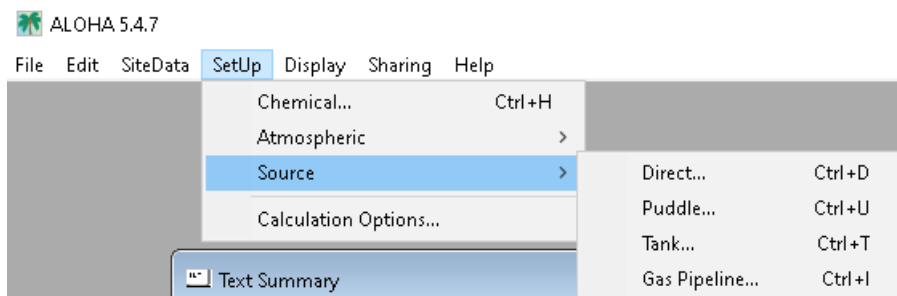
☒  ☐  ☐  OR ☒ enter value :  %  
[0 - 100]

Nota. Fuente: Office of Emergency Management, EPA.

Y finalmente el programa pregunta si existe inversión y que altura en pies o metros, y el porcentaje de humedad del lugar, donde las opciones son húmedo, medio o seco, de igual manera tienes la opción de ingresar el valor del porcentaje (100 %, 75 %, 50 %, 25 % o 5 %).

En relación con los datos de la fuente donde se libera la sustancia química y de la cual se hará la simulación, primeramente, se selecciona si es de manera directa, si es forma de charco, si se libera de un tanque o una tubería (figura 12).

Figura 12. Opciones del tipo de fuente de la emisión de la sustancia química.



Nota. En la figura se muestran las opciones de Charco, Tanque o Tubería. Fuente: Office of Emergency Management, EPA.

En caso de que la liberación sea de un tanque, el programa te pídela forma y las dimensiones de este, por ejemplo, si el tanque es esférico, cilíndrico horizontal, cilíndrico vertical. La longitud y diámetro del tanque es en metros o pies y el volumen en galones o pies cúbicos (figura 13).

Figura 13. Tipo y dimensiones del tanque.

Tank Size and Orientation

Select tank type and orientation:

Horizontal cylinder ☒ Vertical cylinder ☐ Sphere ☐

Enter two of three values:

diameter  length  volume

☐ feet ☒ meters ☐ liters ☒ cu meters

OK Cancel Help

Nota. En la figura se muestra las medidas del diámetro y largo del tanque. Fuente: Office of Emergency Management, EPA.

El estado de la sustancia química dentro del tanque puede ser líquido o solamente gas o desconocido. Existen sustancias químicas que en condiciones atmosféricas son gases, pero al almacenarse en un tanque a presión o temperaturas muy bajas, la sustancia pasa a estado líquido (se licúa). O si las condiciones del tanque son atmosféricas el gas dentro del tanque permanecerá en su mismo estado (figura 14).

Figura 14. Estado y Temperatura de la sustancia química dentro del tanque.

Chemical State and Temperature

**Enter state of the chemical:** Help

☒ Tank contains liquid

☐ Tank contains gas only

☐ Unknown

---

**Enter the temperature within the tank:** Help

☐ Chemical stored at ambient temperature

☒ Chemical stored at  degrees ☐ F ☒ C

OK Cancel

Fuente: Office of Emergency Management, EPA.

La cantidad de masa que se encuentra dentro del tanque se puede expresar en libras, toneladas o kilogramos. De acuerdo con la densidad de la sustancia química dentro del tanque, el programa realiza las conversiones para indicar el volumen de llenado del tanque en galones, pies cúbicos, litros o metros cúbicos y mostrar el nivel de llenado del tanque, se recomienda realizar la simulación en las condiciones de llenado máximas (Figura 15).

Figura 15. Volumen y masa de la sustancia química dentro del tanque.

Liquid Mass or Volume

Enter the mass in the tank OR volume of the liquid

The mass in the tank is:  ☐ pounds  
☐ tons(2,000 lbs)  
☒ kilograms

OR

Enter liquid level OR volume

☐ gallons  
☐ cubic feet  
☒ liters  
☐ cubic meters

% full by volume

Nota. En la figura se muestra aparte de la masa y el volumen el porcentaje de llenado del tanque. Fuente: Office of Emergency Management, EPA.

El programa se da la opción para que se seleccione el escenario que se presenta con la liberación de la sustancia química. Dependiendo de las características fisicoquímicas de la sustancia será el escenario, si se trata de una sustancia inflamable, al momento de liberarse como gas y encuentre un punto de ignición se formara una bola de fuego (BLEVE), si la liberación es en una tubería o un orificio pequeño, la sustancia se libera como un dardo de fuego, y en caso de que sea una sustancia química no flamable, se libera y se dispersa en el aire, dependiendo de las condiciones atmosféricas y topográficas será la distancia que recorrerá (figura 16).

Figura 16. Tipo de falla en el tanque.

Type of Tank Failure

**Scenario:**  
Tank containing a pressurized flammable liquid.

**Type of Tank Failure:**

☒ Leaking tank, chemical is not burning as it escapes into the atmosphere

☐ Leaking tank, chemical is burning as a jet fire

☐ BLEVE, tank explodes and chemical burns in a fireball

**Potential hazards from flammable chemical which is not burning as it leaks from tank:**

- Downwind toxic effects
- Vapor cloud flash fire
- Overpressure (blast force) from vapor cloud explosion

OK Cancel Help

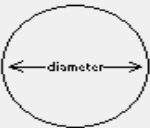
Fuente: Office of Emergency Management, EPA.

Es necesario indicar la forma de la abertura por donde se liberará la sustancia química, puede ser una abertura circular o rectangular, si la abertura es circular es necesario indicar el diámetro en pulgadas, pies, centímetros o metros, si la abertura es rectangular se indica el largo y ancho en las mismas unidades del diámetro. Se indica si la liberación es por una abertura circular o por una tubería (figura 17).

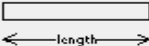
Figura 17. Área y tipo de la fuga en el tanque.

Area and Type of Leak

**Select the shape that best represents the shape of the opening through which the pollutant is exiting**



☒ **Circular opening**



☐ **Rectangular opening**

---

Opening diameter:

☒ **inches**  
☐ **feet**  
☐ **centimeters**  
☐ **meters**

---

**Is leak through a hole or short pipe/valve?**

☒ **Hole**
☐ **Short pipe/valve**

Nota. En la figura se muestra los datos que mejor muestran la forma de la abertura.  
Fuente: Office of Emergency Management, EPA.

Por último, se propone en que altura del tanque podría ser la abertura por donde se liberaría la sustancia química, en el fondo, en la parte superior o en alguna parte del medio del tanque (figura 18).

Figura 18. Altura de la abertura en el tanque.

Height of the Tank Opening



**The bottom of the leak is:**

☐ in
 ☐ ft
 ☐ cm
 ☒ m

above the bottom of the tank

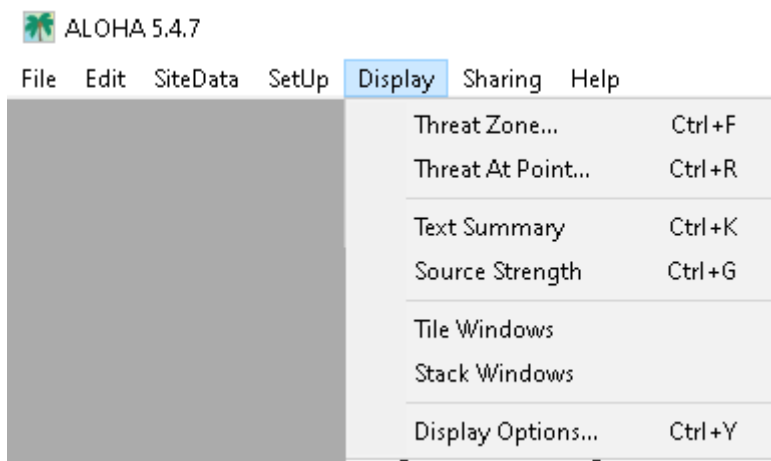
\_\_\_\_\_ OR \_\_\_\_\_

% of the way to the top of the tank

Fuente: Office of Emergency Management, EPA.

Una vez ingresado los datos de la sustancia química, condiciones atmosféricas e información de la fuente emisora, se podrán mostrar los resultados de la simulación. Se puede visualizar la zona de amenaza (threat zone), la amenaza en el punto (threat at point), y un resumen de la información empleada en la simulación (text summary), la velocidad con que es liberada la sustancia química en kilogramos/minuto (source strength) y en display options puedes escoger el sistema de unidades en las que se puede mostrar los resultados, ya sea unidades inglesas o unidades métricas (Figura 19).

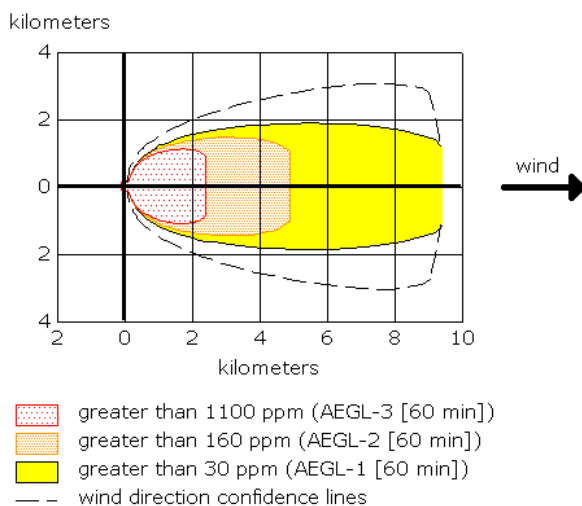
Figura 19. Opciones para mostrar los resultados de la simulación.



Nota. En la figura se muestra las opciones para dar los resultados en diferentes formas. Fuente: Office of Emergency Management, EPA.

En la zona de amenaza (threat zone) se muestra las áreas de la dispersión a diferentes concentraciones mediante el uso de colores, rojo muestra la concentración más alta, naranja concentración media y amarilla la menor concentración (figura 20).

Figura 20. Áreas de amenaza por la liberación de la sustancia química.



En la figura 21, se muestra el resumen de la información mostrada en los pasos anteriores.

Figura 21. Resumen de información sobre la sustancia química, condiciones atmosféricas y fuente.

CHEMICAL DATA:  
Chemical Name: AMMONIA  
CAS Number: 7664-41-7 Molecular Weight: 17.03 g/mol  
AEGL-1 (60 min): 30 ppm AEGL-2 (60 min): 160 ppm AEGL-3 (60 min): 1100 ppm  
IDLH: 300 ppm LEL: 150000 ppm UEL: 280000 ppm  
Ambient Boiling Point: -33.4° C  
Vapor Pressure at Ambient Temperature: greater than 1 atm  
Ambient Saturation Concentration: 1,000,000 ppm or 100.0%

ATMOSPHERIC DATA: (MANUAL INPUT OF DATA)  
Wind: 3 meters/second from SW at 3 meters  
Ground Roughness: open country Cloud Cover: 5 tenths  
Air Temperature: 30° C Stability Class: E  
No Inversion Height Relative Humidity: 75%

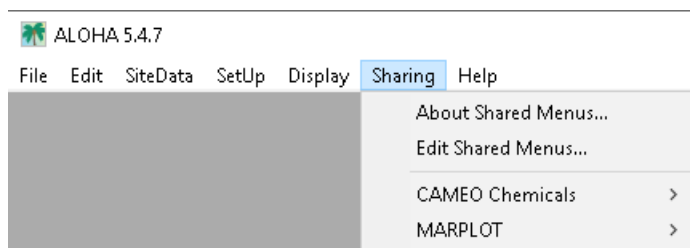
SOURCE STRENGTH:  
Leak from hole in horizontal cylindrical tank  
Flammable chemical escaping from tank (not burning)  
Tank Diameter: 3 meters Tank Length: 10 meters  
Tank Volume: 70.7 cubic meters  
Tank contains liquid Internal Temperature: -30° C  
Chemical Mass in Tank: 45000 kilograms  
Tank is 94% full  
Circular Opening Diameter: 2 meters  
Opening is 1.5 meters from tank bottom

Fuente: Office of Emergency Management, EPA.

Y como resultado de la simulación, a una distancia de 2.25 kilómetros se tiene una concentración de 1,100 partes por millón (ppm), a los 5 kilómetros se tiene 160 ppm y a una distancia de 9.2 kilómetros se tiene una concentración de 30 ppm.

Los resultados anteriores también pueden ser mostrados sobre un mapa del lugar donde se encuentra la fuente de emisión del amoníaco. Para esta ocasión se empleó el programa MARPLOT, el cual puede ser descargado de la siguiente liga: <https://www.epa.gov/cameo/marplot-software> (figura 22).

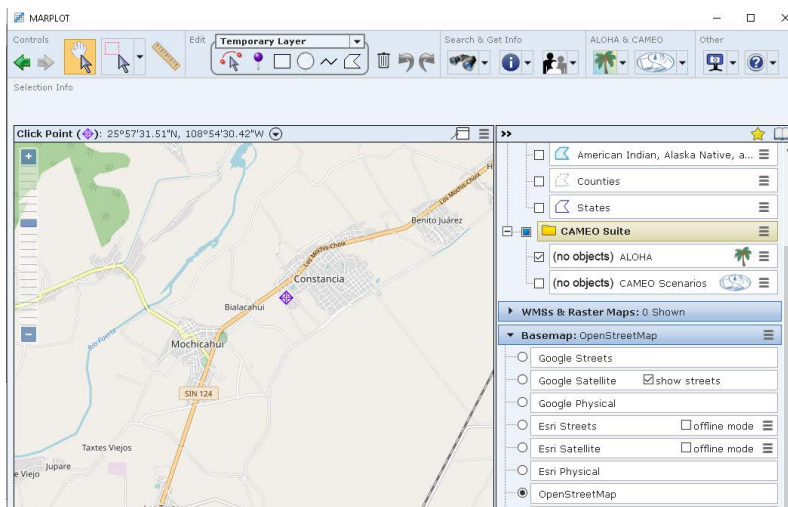
Figura 22. Opciones para mostrar los resultados sobre un mapa.



Fuente: Office of Emergency Management, EPA.

MARPLOT, muestra cualquier ubicación por lo que se puede colocar las zonas de amenazas que se consideren para la simulación que se está realizando (figura 23).

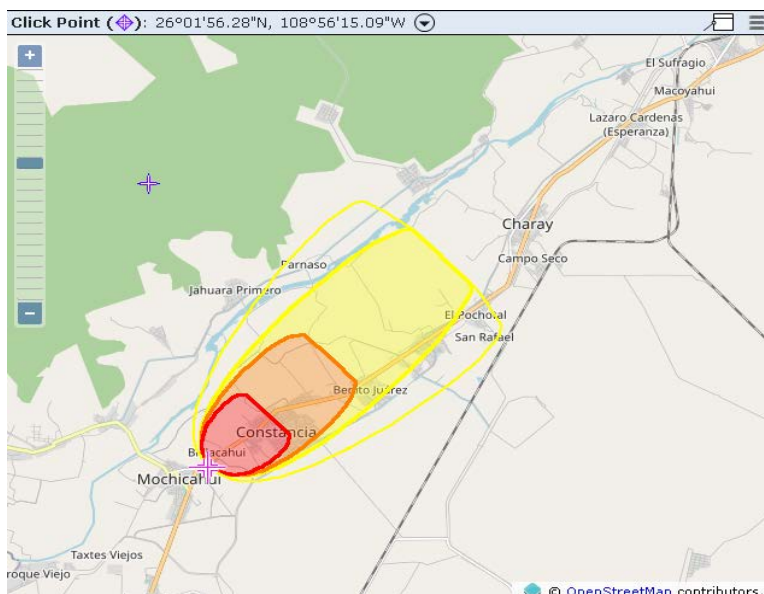
Figura 23. Mapa en Marplot para mostrar los resultados.



Fuente: Office of Emergency Management, EPA.

De esta manera, se puede visualizar de una mejor forma las zonas a las cuales llegará el amoníaco y en qué concentraciones (figura 24).

Figura 24. Resultados mostrados en mapa.



## Discusión

Con el uso del programa Aloha se pueden simular diferentes eventos en lugares donde se emplean sustancias químicas peligrosas, desde el nivel de radiación térmica de un charco de fuego para el caso de derrame de un combustible, o una explosión por expansión del vapor de un líquido en ebullición (BLEVE) para el caso de almacenamiento a presión de algún gas inflamable, o un dardo de fuego igualmente para un gas inflamable almacenado a presión en un tanque pero que es liberado por una tubería o una pequeña abertura circular y también la simulación de liberación de una sustancia toxica no inflamable.

De acuerdo con la simulación mostrada en los resultados anteriores, se puede decir que la liberación de 45,000 kilogramos de amoníaco de manera casi instantánea se desplazara hasta un poco más de nueve kilómetros teniendo una concentración a esta distancia de 30 ppm. Exposiciones de amoníaco a un nivel de 100 ppm provocan irritaciones de las mucosas de la nariz y garganta.

Y concentraciones por encima de 700 ppm causan severas irritaciones en los ojos, hemorragias e hinchazón, y si no es tratado inmediatamente, puede llevar a pérdidas parciales o totales de la vista. Las personas que se encuentren en dirección del viento a aproximadamente 5 kilómetros desde la fuente donde se originó la liberación podrían presentar irritación de las mucosas de la nariz y garganta y las que se encuentren aproximadamente a tres kilómetros podrían perder la vista.

Para el caso del trabajo de Soto (2022) donde se liberaron 25,000 kilogramos de nitrato de amonio ( $\text{NH}_4\text{NO}_3$ ), la zona de amenaza resultó de 5 kilómetros, aunque el radio de afectación fue el mismo para los dos casos, se emplearon diferentes masas y distintas sustancias químicas. Cruz (2013), realizó una simulación de una fuga de amoniaco empelando 4,400 kg de amoniaco, teniendo como resultado 330 metros como área de amenaza, existe en proporción semejanza con los resultados obtenidos en este trabajo.

Dependiendo de la cantidad de masa que es liberada las sustancias químicas, las condiciones atmosféricas y las características de la fuente es la distancia que recorrerá el gas en el aire y por lo tanto los efectos que representará a la salud y medioambiente.

Para obtener un mejor análisis de los eventos que se pueden presentar dentro de una empresa que maneja sustancias químicas peligrosas, existe la opción de realizar más una simulación al emplear distintos escenarios, por ejemplo, utilizando condiciones atmosféricas que predominan en invierno y otra simulación con datos que predominan en verano.

## Conclusión

De acuerdo con la legislación ambiental mexicana, las empresas que manejen, transporten, almacenen o hagan uso de sustancias químicas peligrosas deberán someter a evaluación un estudio e riesgo simulando la liberación de estas sustancias, con el objetivo de determinar las zonas de afectación y por lo tanto que lugares serán afectados y en qué grado. El programa Aloha, al igual que otros programas son una herramienta valiosa que ayuda a la realización de estudios de riesgo y que de esta manera las empresas cumplan con el deber de informar a las autoridades y sociedad el alcance de afectación de se tendrá en caso de la liberación de una sustancia peligrosa, y de esta manera la autoridad competente en materia de riesgo ambiental pueda determinar con estudios cuantitativos si el lugar propuesto para realizar una obra que maneje sustancias químicas peligrosas es viable de llevarse a cabo.

## Referencias

- Cruz, L., A. (2013). *Análisis de riesgo por fuga de amoniaco en tanque de almacenamiento empleando modelo computacional*. [Tesis de licenciatura. Universidad Autónoma Agraria Antonio Narro]. Repositorio Digital del Centro de Información y Documentación.
- Del Sole, M., B. (2019). *Simulación de incidentes en el transporte de materiales peligrosos en Gran la Plata*. [Tesis, Licenciatura. Facultad de Ciencias Exactas, Universidad Nacional de la Plata]. Recuperado de: [http://sedici.unlp.edu.ar/bitstream/handle/10915/81818/Documento\\_completo.pdf-PDFA2u.pdf?sequence=1&isAllowed=y](http://sedici.unlp.edu.ar/bitstream/handle/10915/81818/Documento_completo.pdf-PDFA2u.pdf?sequence=1&isAllowed=y).
- Duche, C., P. (2021). *Evaluación cuantitativa de explosión por riesgo químico en sistemas presurizados de la unidad separadora de aire de Linde Ecuador S.A.* [Tesis de Licenciatura, Universidad de Guayaquil]. Repositorio de la Universidad de Guayaquil.
- EPA. 2022. ALOHA Software. *United State Enviromental Protection Agency*. Recuperado de: <https://www.epa.gov/cameo/aloha-software>.
- Fluideco. 2019. *¿Qué es el amoniaco, ¿cómo se obtiene y para qué se utiliza?* Recuperado de: [https://fluideco.com/amoniaco-utilidades\\_usos/#:~:text=El%20amoniaco%20es%20un%20producto,se%20evapora%20muy%20de%20prisa](https://fluideco.com/amoniaco-utilidades_usos/#:~:text=El%20amoniaco%20es%20un%20producto,se%20evapora%20muy%20de%20prisa).
- Gobierno de la Ciudad de México. (s.f). *Guía para el manejo seguro del gas LP*. Recuperado de: <https://www.proteccioncivil.cdmx.gob.mx/storage/app/uploads/public/61f/184/ab0/61f184ab0bb26681477942.pdf>
- Gutierrez, M. (2014). *Salud Ocupacional- Inhalacion Masiva de Amoniaco*. Obtenido de salud ocupacional- inhalacion masiva de amoniaco.
- Lana Sarrate. (2020). *Hablando de gases: el gas cloro*. Recuperado de: <https://www.lanasarrate.es/blog/deteccion-gas-cloro/#:~:text=El%20cloro%20se%20utiliza%20principalmente,de%20aguas%2C%20pl%C3%A1sticos%2C%20etc>.
- León, A. (2022). *Modelización de accidentes en una refinería de Tarragona usando ALOHA*. [Tesis de Maestría. Dpto. Ingeniería Química y Ambiental Escuela Técnica Superior de Ingeniería Universidad de Sevilla]. Recueprado de: <https://biblus.us.es/bibing/proyectos/abreproy/72347/fichero/TFM-2347+Le%C3%B3n+Sierra%2C+A..pdf>
- Martinez, J. C. (Viernes de agosto de 2006). Fuga de gas de amoniaco mata a dos en Sinaloa. *EL UNIVERSAL*. Recuperado de <https://archivo.eluniversal.com.mx/estados/62312.html>

- Parra Abril, C. A., Ortiz Rey, J. F., & Meneses Velosa, S. L. (2015). Uso de herramientas de ingeniería para la gestión de riesgo tecnológico en una central de abastecimiento agrícola con aglomeraciones permanentes en Bogotá. *Avances Investigación en Ingeniería*, 12(1). <https://doi.org/10.18041/1794-4953/avances.2.277>
- National Academy of Science. (1983). *Risks Assessment in Federal Government Managing the Process*. National Academy Press. Whashington D.C.
- NJHealth. (2016). *Hoja Informativa sobre Sustancias Peligrosas-Amoniac*. <https://www.nj.gov/health/eoh/rtkweb/documents/fs/0084sp.pdf>
- Pelayo, J. (2001). Curso-Taller de Mecánica Automotriz. *Facultad de Ingeniería UNAM*. Recuperado de: [http://www.ptolomeo.unam.mx:8080/jspui/bitstream/132.248.52.100/14676/1/decd\\_2897.pdf](http://www.ptolomeo.unam.mx:8080/jspui/bitstream/132.248.52.100/14676/1/decd_2897.pdf)
- SEMARNAT. (2008). Actividades Altamente Riesgosas. *Compendio de Estadísticas Ambientales 2008*. Recuperado de [https://apps1.semarnat.gob.mx:8443/dgeia/informe\\_2008/compendio\\_2008/compendio2008/10.100.8.236\\_8080/ibi\\_apps/WFServlet90e3.html](https://apps1.semarnat.gob.mx:8443/dgeia/informe_2008/compendio_2008/compendio2008/10.100.8.236_8080/ibi_apps/WFServlet90e3.html)
- SEMARNAT. (2018). Sistema Nacional de Información Ambiental y de Recursos Naturales. *Riesgo Ambiental*. Recuperado de: <https://apps1.semarnat.gob.mx:8443/dgeia/informe18/tema/cap7.html#tema5>
- Soto, S., A. Alcantara M., E. (2022). Determinación de consecuencias ocasionadas por la explosión de nitrato de amonio en Celemania, Coahuila. *Repositorio Institucional Zaloamati*. Universidad Autónoma Metropolitana.
- Suroeste Grupo. (2023). *Usos del Gas Natural*. Recuperado de: <https://grupo-suroeste.es/blog/usos-del-gas-natural/>.
- Torres, E. (2018). Evacuan a 3 mil 604 personas tras fuga de amoníaco en Guasave. *Luz Noticias*, Recuperado de: <https://www.luznoticias.mx/multimedia/evacuan-a-3-mil-604-personas-tras-fuga-de-amoniaco-en-guasave/53283>.
- Valdés L., Ortega B., Consuegra G. (2019) “Technological risks map associated with dangerous goods management in ports of Barranquilla”, *Prospectiva*, Vol 17, N° 1, 45-50, 2019.



# Capítulo 7

---

## **Aplicaciones de redes neuronales recurrentes en el análisis cuantitativo de series temporales para la gestión sostenible de recursos naturales y medioambientales**

*Gilberto Bojórquez-Delgado*

*Jesús Bojórquez-Delgado*

*Adalid Graciano-Obeso*

## Introducción

La gestión eficiente y sostenible de los recursos naturales y medioambientales es uno de los mayores retos que enfrenta la humanidad en el siglo XXI (Rodríguez Morales et al., 2011). Este desafío se ve exacerbado por la complejidad inherente a los sistemas naturales, los patrones climáticos en constante cambio y la necesidad urgente de adaptar nuestras prácticas a una realidad ambiental que evoluciona rápidamente (Jaquenod De Zsögön, 2019). En este panorama, el análisis de series temporales emerge como una herramienta crítica, esencial para entender, predecir y manejar fenómenos ambientales con una precisión sin precedentes. Aquí, la justificación para centrarse en las Redes Neuronales Recurrentes (RNR) se hace evidente, ya que estas poseen una capacidad única para procesar y aprender de datos temporales, ofreciendo un enfoque innovador y más efectivo para enfrentar estos desafíos ambientales complejos.

Los antecedentes del análisis cuantitativo en la ciencia medioambiental tradicionalmente han dependido de modelos estadísticos y matemáticos que, aunque útiles, a menudo se ven limitados por supuestos de linealidad y estacionariedad (Quiguiri Daquilema, 2023). Estas limitaciones son particularmente problemáticas en el contexto de fenómenos naturales, donde la no linealidad y los patrones temporales complejos son la norma (Alonso, 2022). Las RNR, que se benefician de la capacidad de las redes neuronales para aprender de los datos sin la necesidad de supuestos previos, representan una evolución natural de estas metodologías, proporcionando un marco más flexible y robusto para el análisis medioambiental (Vega Moreno, 2021).

Dentro del espectro de las RNR, las Redes de Memoria a Largo Plazo (LSTM por sus siglas en inglés) se destacan por su capacidad para evitar el problema del desvanecimiento del gradiente (Guamán Pachacama y Segura Muñoz, 2021), lo que les permite no solo capturar relaciones a corto plazo, sino también aprender dependencias de largo plazo (Torres, 2021). Esta característica las hace particularmente adecuadas para modelar y predecir series temporales ambientales que requieren la comprensión de efectos acumulativos y tendencias a largo plazo (Salazar Molina, 2023) por lo que los gobiernos del mundo entero se han comprometido a disminuir estas emisiones para lograr el objetivo 7 del desarrollo sostenible 2030 de las Naciones Unidas: Energía

asequible y no contaminante. [Resumen truncado para la entrada .bib]”, “genre”: “Doctoral thesis”, “publisher”: “Universidad Internacional Iberoamericana México”, “title”: “Metodología para la Evaluación en Proyectos de Energía Solar Fotovoltaica”, “URL”: “https://repositorio.unini.edu.mx/id/eprint/6010”, “author”: [{“family”: “Salazar Molina”, “given”: “Erick Alonso”}], “issued”: [{“date-parts”: [[“2023”, 2]]}], “schema”: “https://github.com/citation-style-language/schema/raw/master/csl-citation.json”} .

La arquitectura única de las LSTM les permite retener información por periodos prolongados, haciendo uso de estructuras llamadas “celdas de memoria” (Lindemann et al., 2021). Estas celdas regulan el flujo de información permitiendo que la red recuerde u olvide información de manera selectiva. Este mecanismo es fundamental cuando se trata de datos ambientales, donde la relevancia de la información puede variar drásticamente a lo largo del tiempo (Nguyen et al., 2021).

Dado el complejo entramado de datos que caracteriza al medioambiente, se hace imprescindible recurrir a herramientas computacionales avanzadas que puedan manejar y procesar eficientemente la información. Python emerge como una opción preeminente (Bobadilla, 2021), avalado por su rica colección de bibliotecas científicas y de aprendizaje automático como TensorFlow y Keras, que complementan y potencian la estructura de las LSTM (Pang et al., 2020). La magnitud y la complejidad de los datos medioambientales exigen un enfoque automatizado, dado que la capacidad de procesamiento manual es insuficiente. La implementación de Python en este contexto no solo facilita la gestión de grandes volúmenes de datos, sino que también optimiza la capacidad de las redes neuronales para aprender de patrones temporales extensos y variables (Joseph et al., 2021).

Además, la necesidad de iterar modelos, ajustar parámetros y validar resultados requiere de una plataforma que ofrezca no solo flexibilidad y potencia computacional sino también una comunidad activa y recursos de aprendizaje.

La elección de Python como herramienta para el análisis de RNR responde a estos requisitos, proporcionando un entorno propicio para el desarrollo rápido y eficiente de modelos predictivos robustos (Wang y Dowling, 2022) calibrating, and validating said science-based models often remains an art in practice. Model-based design of experiments (MBDoE). La simplicidad sintáctica de Python, junto con su capacidad para manejar operaciones de alto nivel, permite a los investigadores centrarse en la solución de problemas complejos sin la carga de detalles de programación de bajo nivel. Este enfoque permite trasladar los esfuerzos del análisis manual a uno automatizado, lo cual es esencial para manejar la escala y complejidad de los desafíos que presenta la

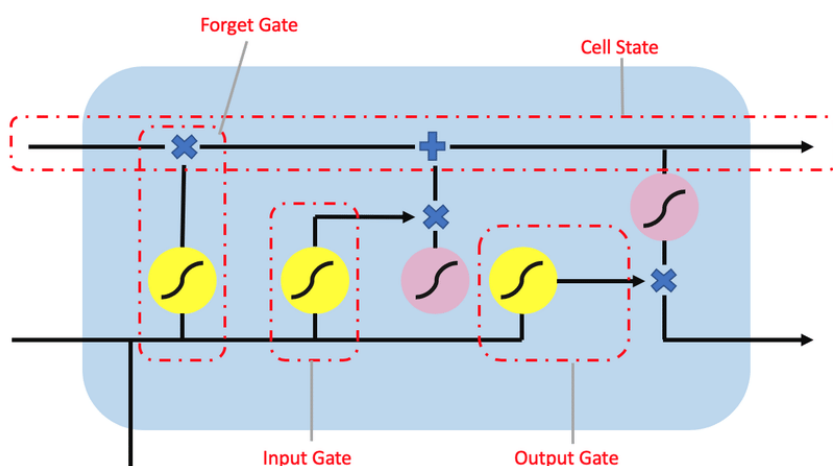
gestión sostenible de recursos naturales y medioambientales (Landázuri Páez et al., 2023). El objetivo de este capítulo es, por lo tanto, doble: en primer lugar, demostrar cómo las RNR y las LSTM pueden ser aplicadas para interpretar y predecir con precisión los patrones en los datos medioambientales. Esto se logra a través de una exploración detallada de su teoría subyacente, arquitectura y aplicabilidad práctica en escenarios reales. En segundo lugar, el capítulo busca inspirar y equipar a los profesionales medioambientales, científicos de datos y responsables políticos con el conocimiento y las herramientas necesarias para integrar estas tecnologías en la toma de decisiones relativas a la sostenibilidad.

## Metodología

### *Fundamentos teóricos*

La arquitectura de las RNR, específicamente las LSTM, es fundamental para modelar series temporales, gracias a su capacidad para aprender y recordar información a lo largo de periodos extensos (Veléz et al., 2021). La Figura 1 ilustra la arquitectura interna de una celda LSTM, la cual es esencial para la comprensión profunda del funcionamiento de las RNR aplicadas al análisis ambiental (Rengasamy et al., 2020).

Figura 1. Arquitectura de una Celda LSTM.



Fuente: (Rengasamy et al., 2020).

*Puertas de Olvido (Forget Gates):* La función de estas puertas es seleccionar la información del estado de la celda anterior que ya no es relevante para la secuencia actual y debe ser descartada. Esto es crítico en datos ambientales donde algunos factores, como los datos estacionales, pueden dejar de ser influyentes después de cierto tiempo.

*Puertas de Entrada (Input Gates) y Estado de la Celda (Cell State):* Las puertas de entrada deciden qué nueva información es relevante y debe ser añadida al estado de la celda. El estado de la celda actúa como un transportador de información relevante a lo largo de la secuencia de tiempo, siendo actualizado en cada paso.

*Puertas de Salida (Output Gates):* Estas puertas determinan qué información del estado de la celda contribuirá a la salida y será utilizada para las predicciones o decisiones del modelo.

Esta dinámica de puertas y estados permite a las LSTM manejar y ajustarse a la naturaleza no estacionaria de los datos ambientales, en los cuales los patrones y relaciones pueden cambiar o evolucionar con el tiempo. Al aplicar esta metodología, se puede capturar tanto la influencia de eventos recientes como la de aquellos acontecimientos más antiguos que todavía tienen un impacto en el estado actual del sistema ambiental estudiado.

#### *Ecuación de Actualización del Estado de la Celda:*

El estado de la celda en las LSTM es actualizado en cada paso de tiempo para mantener la información relevante a lo largo de la secuencia. La actualización se realiza mediante la siguiente ecuación:

$$C_t = f_t \cdot C_{\{t-1\}} + i_t \cdot \tilde{C}_t$$

Donde:

$C_t$  es el estado de la celda en el tiempo actual.

$f_t$  es la activación de la puerta de olvido, determinando qué parte del estado anterior  $C_{\{t-1\}}$  debe conservarse.

$C_{\{t-1\}}$  es el estado de la celda en el paso de tiempo anterior.

$i_t$  es la activación de la puerta de entrada, que decide qué nueva información se añadirá al estado de la celda.

$\tilde{C}_t$  es el estado de la celda candidato, generado a partir de la información actual, que podría añadirse al estado de la celda si la puerta de entrada así lo permite.

### *Implementación con Python y Google Colab:*

La implementación de las LSTM para el análisis de series temporales ambientales se realiza mediante Python, aprovechando bibliotecas de aprendizaje automático como TensorFlow y Keras. Google Colab ofrece un entorno de computación en la nube que facilita la ejecución de modelos complejos de LSTM. Los investigadores pueden colaborar en tiempo real, compartiendo y ejecutando código en un entorno estandarizado que asegura la reproducibilidad de los resultados.

En el contexto del análisis ambiental, las LSTM se entrenan utilizando grandes conjuntos de datos históricos. Estos datos pueden incluir mediciones de temperatura, concentraciones de gases, niveles de precipitación, entre otros factores relevantes. La formación de estos modelos se orienta a predecir futuros estados ambientales y fenómenos con implicaciones críticas para la gestión de recursos y la planificación sostenible.

## **Preparación de datos**

### *Recolección de datos temporales ambientales:*

La recolección de datos precisos y relevantes es vital para el análisis efectivo y significativo de series temporales, especialmente en estudios ambientales. Estos datos, a menudo heterogéneos y voluminosos, pueden provenir de múltiples fuentes, como sensores remotos, estaciones meteorológicas, observaciones satelitales, y registros históricos.

- **Sensores remotos:** Utilizados para recoger datos a gran escala, como imágenes satelitales y mediciones de radar, que pueden proporcionar información sobre la cobertura terrestre, la temperatura superficial y los patrones climáticos.
- **Registros históricos:** Datos acumulados a lo largo del tiempo, a menudo disponibles a través de bases de datos gubernamentales o investigaciones científicas, que pueden incluir temperaturas pasadas, niveles de precipitación y otras mediciones meteorológicas.
- **Mediciones In Situ:** Mediciones directas tomadas en el campo, como niveles de contaminantes en un lago o concentraciones de CO<sub>2</sub> en un bosque, que proporcionan datos detallados a nivel local.

### *Preprocesamiento de datos:*

El preprocesamiento es un paso crítico que prepara los datos brutos para su análisis, asegurando que sean limpios, normalizados y estructurados adecuadamente para la entrada en el modelo LSTM. Esto implica técnicas como la imputación de datos faltantes, la detección y corrección de anomalías, y la transformación de datos para facilitar la extracción de características temporales relevantes para la predicción (García et al., 2016).

#### *1. Limpieza de datos:*

- Se identifican y manejan valores faltantes, que pueden ser rellenados utilizando métodos como la imputación media o la interpolación, dependiendo del contexto y la naturaleza de los datos.
- Se eliminan outliers o valores atípicos que podrían distorsionar el entrenamiento del modelo.
- Se corrigen errores de entrada y se verifica la consistencia de los datos.

#### *2. Normalización de datos:*

- Se aplican técnicas de normalización para escalar los datos dentro de un rango adecuado, generalmente entre 0 y 1, utilizando métodos como la normalización min-max o la estandarización Z-score.
- Se asegura de que todas las características (variables de entrada) estén en la misma escala para que el modelo no se sesgue hacia características con valores numéricos más altos.

#### *3. Transformación de datos:*

- Se convierten datos no numéricos en formatos numéricos mediante codificación, si es necesario, como la codificación one-hot para variables categóricas.
- Los datos se estructuran en secuencias que las LSTM pueden procesar. Por ejemplo, si se están prediciendo niveles diarios de contaminación, se organizan en ventanas de tiempo, como secuencias de 7 días para predecir el octavo día.

#### 4. División de datos:

- Se dividen los datos en conjuntos de entrenamiento, validación y prueba. Un enfoque común para esta división es el temporal, donde los datos más antiguos se utilizan para el entrenamiento, y los más recientes para la prueba. Este método asegura que el modelo se entrene con datos históricos y sea evaluado en un contexto más actual, lo que puede ser crucial en aplicaciones donde las tendencias y patrones cambian con el tiempo.
- Se realiza la validación del modelo utilizando una porción del conjunto de entrenamiento o aplicando técnicas de validación cruzada. Este proceso es crucial para evaluar la capacidad de generalización del modelo, asegurando que su desempeño sea robusto y confiable frente a datos nuevos y no vistos durante el entrenamiento.

## Diseño del modelo

### *Configuración de la Arquitectura LSTM*

La configuración de una red LSTM es un paso crítico que requiere un enfoque reflexivo y experimental para encontrar la arquitectura óptima que se adapte a la complejidad de los datos y la problemática específica. Este proceso se desglosa en dos componentes esenciales que se describen a continuación:

#### 1. Estructuración de la Red LSTM

- Número de capas y unidades: Se recomienda iniciar con una sola capa LSTM para establecer una línea de base, aumentando progresivamente el número de capas y unidades si es necesario para capturar complejidades adicionales en los datos.
  - Comando Python: `keras.layers.LSTM(units=50, return_sequences=True)` para una capa LSTM con 50 unidades y salida secuencial.
- Funciones de activación: La función `tanh` es comúnmente usada para las activaciones internas de la celda LSTM, mientras que la función `sigmoid` es típica para las puertas de olvido, entrada y salida.
  - Comando Python: `keras.layers.LSTM(activation='tanh', recurrent_activation='sigmoid')`.

## 2. Configuración de parámetros

- Tasa de aprendizaje (Learning rate): Una tasa de aprendizaje más pequeña puede ser necesaria para entrenamientos más estables, aunque puede hacer que el entrenamiento sea más lento.
  - Comando Python: `keras.optimizers.Adam(learning_rate=0.001)` para configurar el optimizador Adam con una tasa de aprendizaje de 0.001.
- Número de épocas: El número de épocas debe ser lo suficientemente grande para permitir que la red aprenda de los datos, pero no tan grande que cause sobreajuste.
  - Comando Python: `model.fit(X, y, epochs=100)` para entrenar el modelo durante 100 épocas.
- Tamaño de los Lotes (Batch Size): Un tamaño de lote más pequeño a menudo proporciona una estimación del gradiente más ruidoso, lo que puede ayudar a evitar mínimos locales, pero también puede hacer que el entrenamiento sea más variable.
  - Comando Python: `model.fit(X, y, batch_size=32)` para entrenar el modelo con un tamaño de lote de 32.

### Consideraciones adicionales:

Además de la estructuración básica y la configuración de parámetros en una red LSTM, hay varias consideraciones adicionales que pueden influir significativamente en el rendimiento y la eficiencia del modelo. Estas prácticas avanzadas son esenciales para refinar aún más la red y asegurar su robustez frente a desafíos comunes como el sobreajuste. A continuación, se detallan algunas de estas consideraciones clave:

- Regularización: Para evitar el sobreajuste, se debe considerar usar técnicas como dropout o regularización L1/L2 en tu red.
  - Comando Python: `keras.layers.LSTM(units=50, dropout=0.2)` para aplicar un dropout del 20 % en la capa LSTM.
- Inicialización de Pesos: Se considera una práctica recomendada la inicialización adecuada de los pesos de la red para mejorar el entrenamiento.
  - Comando Python: `keras.layers.LSTM(units=50, kernel_initializer='glorot_uniform')` para usar la inicialización de Glorot/Xavier.

Al diseñar el modelo LSTM, es esencial realizar una búsqueda de hiperparámetros, ya sea manualmente o mediante técnicas automatizadas como la

búsqueda en cuadrícula (grid search) o la optimización bayesiana. Además, es esencial asegurar el monitoreo de métricas de rendimiento clave y la visualización del entrenamiento para realizar ajustes informados a la arquitectura y configuración del modelo. Este proceso involucra ajustar aspectos como el número de capas LSTM, la tasa de aprendizaje, y el número de épocas, para maximizar la precisión de las predicciones. La interpretación de los resultados y la retroalimentación iterativa entre los pasos de preprocesamiento y modelado son fundamentales para refinar el modelo y alcanzar resultados óptimos.

## **Entrenamiento y validación**

El proceso de entrenamiento es fundamental para la construcción de un modelo preciso y fiable. La manera en que se entrena una red LSTM es crucial para su capacidad de generalización y rendimiento en datos no vistos previamente.

Una vez establecido el marco general del proceso de entrenamiento, es crucial abordar aspectos específicos como la división de datos y las estrategias para prevenir el sobreajuste, que son fundamentales para garantizar la efectividad y precisión del modelo LSTM.

### *1. División de datos*

- Antes del entrenamiento, es importante dividir el conjunto de datos en tres partes: entrenamiento, validación y prueba.
  - Comando Python: `train_test_split()` de la biblioteca `sklearn.model_selection` para realizar la división.
- La porción de entrenamiento se utiliza para ajustar los pesos del modelo, la validación para afinar los hiperparámetros y evitar el sobreajuste, y la prueba para evaluar la capacidad predictiva del modelo en datos nuevos.

### *2. Prevención del sobreajuste*

- Técnicas como la regularización y el dropout son esenciales para prevenir el sobreajuste.
  - Comando Python: `keras.layers.LSTM(units=50, dropout=0.2)` para aplicar dropout a la capa LSTM.
  - Comando Python: `keras.regularizers.l1(0.01)` o `keras.regularizers.l2(0.01)` para aplicar regularización L1 o L2 a los pesos de la red.
- Es vital monitorear la pérdida de entrenamiento y validación durante el entrenamiento para detectar señales de sobreajuste. Si la pérdida de validación

aumenta mientras la de entrenamiento disminuye, se puede estar ante un caso de sobreajuste.

## Evaluación del modelo

### 1. Métricas de rendimiento

- Una vez entrenado el modelo, se utiliza el conjunto de prueba para evaluar su rendimiento mediante métricas estándar.
  - Comando Python: `model.evaluate()` para obtener la pérdida y las métricas del modelo.
- El Error Cuadrático Medio (MSE) mide la media de los cuadrados de los errores, siendo una métrica común para problemas de regresión.
- El Coeficiente de Determinación ( $R^2$ ) indica la proporción de la varianza de la variable dependiente que es predecible a partir de las variables independientes.
  - Comando Python: `sklearn.metrics.mean_squared_error()` y `sklearn.metrics.r2_score()` para calcular MSE y  $R^2$ , respectivamente.

### 2. Validación cruzada

- Para garantizar la robustez del modelo, se puede emplear la validación cruzada, especialmente útil en conjuntos de datos más pequeños o cuando se requiere una evaluación exhaustiva del modelo.
  - Comando Python: `sklearn.model_selection.KFold()` o `sklearn.model_selection.TimeSeriesSplit()` para la validación cruzada, dependiendo de si los datos pueden ser aleatorizados o no (`TimeSeriesSplit` es para series temporales donde el orden es importante).

## Implementación práctica con Python y Colab

Python se ha establecido como uno de los lenguajes de programación más versátiles y ampliamente utilizados en el campo del aprendizaje automático y el análisis de datos. Su simplicidad sintáctica, junto con un extenso ecosistema de bibliotecas y herramientas, lo hacen ideal para implementar y experimentar con modelos de aprendizaje profundo como las redes LSTM.

Google Colab es una plataforma en la nube que facilita la escritura y ejecución de código Python, especialmente útil para el aprendizaje profundo. Se puede configurar rápidamente y es compatible con bibliotecas como TensorFlow y Keras. A continuación, se presentan los pasos clave para iniciar con Colab y la programación en Python.

## 1. Uso de Google Colab

Google Colab es un servicio de nube basado en Jupyter Notebooks que facilita la escritura y ejecución de código Python. Ofrece un entorno colaborativo con acceso gratuito a recursos computacionales, incluyendo GPUs y TPUs, lo que lo hace ideal para entrenar modelos de aprendizaje profundo.

Configuración de Colab:

- Debe iniciarse un nuevo notebook en Colab y cargar las bibliotecas necesarias con el comando `!pip install tensorflow keras`.
- El comando `from google.colab import drive`, seguido de `drive.mount('/content/drive')`, se utiliza para acceder y cargar datos desde Google Drive si es necesario.
- Es importante asegurarse de seleccionar la configuración de hardware adecuada en Colab, como la GPU, para acelerar el entrenamiento del modelo.

## 2. Codificación en Python

El desarrollo de modelos LSTM se realiza con TensorFlow y Keras, dos de las bibliotecas más populares para el aprendizaje profundo, que proporcionan abstracciones de alto nivel para diseñar y entrenar redes neuronales.

- Implementación de un Modelo LSTM:
  - El modelo se define utilizando la clase `Sequential` de Keras, y para agregar una capa LSTM, se utiliza `model.add(LSTM(units=50, return_sequences=True))`, ajustando las unidades según la complejidad del problema.
  - El modelo se compila con el comando `model.compile(optimizer='adam', loss='mean_squared_error')`, seleccionando un optimizador y una función de pérdida adecuada para tu tarea.
- Visualización de resultados:
  - Se emplea `matplotlib` o `seaborn` para visualizar los resultados, como la pérdida durante el entrenamiento, utilizando `plt.plot(history.history['loss'])` y `plt.plot(history.history['val_loss'])`.
  - Para la interpretación de las salidas del modelo se examinan las predicciones en comparación con los datos reales y ajusta el modelo si es necesario. Utiliza técnicas de posprocesamiento para analizar errores, ajustar hiperparámetros y probar diferentes arquitecturas de red.

La implementación práctica de las LSTM mediante Python y Colab facilita un avance significativo en la comprensión y predicción de series temporales

complejas. El proceso meticuloso de recolección y preprocesamiento de datos, seguido por una cuidadosa configuración y entrenamiento del modelo, es esencial en la capacidad de predecir dinámicas temporales con gran relevancia para el sector del campo de estudio.

### *Ejercicio práctico:*

Aplicación de redes neuronales recurrentes para el análisis de series temporales en la gestión sostenible de recursos energéticos renovables.

En esta sección, se aborda el desafío de analizar y predecir el consumo de energía renovable, un componente clave en la gestión sostenible de los recursos naturales. Para ello, se utiliza el conjunto de datos *WorldSustainabilityDataset.csv*, obtenido del reconocido repositorio en línea Kaggle. Este amplio conjunto de datos encapsula métricas de sostenibilidad de 173 países a lo largo de 19 años, reflejando la diversidad y complejidad de los desafíos a nivel mundial en sostenibilidad. Compilado originalmente para el TrueCue Women+Data Hackathon, el dataset incluye 54 variables variadas, proporcionando un espejo de la riqueza y la heterogeneidad de los datos ambientales globales.

El análisis se centra en un subconjunto específico: el consumo de energía renovable en México. La elección de esta área focal representa una oportunidad valiosa para demostrar cómo las metodologías rigurosas de análisis de series temporales, con especial énfasis en las RNR, pueden ser aplicadas para obtener insights significativos en el contexto de la sostenibilidad medioambiental. Los datos analizados presentan desafíos típicos encontrados en entornos profesionales, tales como valores faltantes y registros incompletos. Estos desafíos no solo ponen a prueba la robustez de las herramientas analíticas empleadas, sino que también refuerzan la necesidad de enfoques de análisis cuidadosos y considerados. A través de este caso de estudio, se ilustrará la aplicación práctica de las RNR, destacando su utilidad en el procesamiento y modelado de datos complejos para la toma de decisiones informadas en la gestión de recursos naturales. A continuación, se presentan los pasos para realizar el estudio y predicción del comportamiento del consumo de energía renovable en México. Tomando como base los datos históricos desde el año 2000 hasta 2015, se genera una ventana de tres años para predecir el consumo hasta 2018 y compararlo con los datos reales de esa fecha.

## 1. Importación de librerías y definición de funciones

En Google Colab, se comienza importando todas las librerías necesarias y se definen las funciones que se utilizarán. La función *split\_sequence* convierte los datos de series temporales en un formato que las redes LSTM pueden procesar, tal como se muestra en el siguiente fragmento de código:

```
import pandas as pd
import numpy as np
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dropout, Dense
from tensorflow.keras.callbacks import EarlyStopping
import matplotlib.pyplot as plt

def split_sequence(sequence, n_steps):
    X, y = list(), list()
    for i in range(len(sequence)):
        end_ix = i + n_steps
        if end_ix > len(sequence)-1:
            break
        seq_x, seq_y = sequence[i:end_ix], sequence[end_ix:]
        X.append(seq_x)
        y.append(seq_y)
    return np.array(X), np.array(y)
```

En el código se muestra:

- pandas y numpy son librerías estándar para la manipulación de datos y operaciones matemáticas.
- De tensorflow.keras, se debe importar *Sequential*, *LSTM*, *Dropout*, y *Dense* para construir nuestra red neuronal, y *EarlyStopping* para evitar el sobreajuste.
- matplotlib.pyplot se utiliza para la visualización de datos.

La función *split\_sequence* es crucial para preparar los datos. Convierte una secuencia de valores en un conjunto de pares de entrada/salida que la red puede aprender.

## 2. Carga y filtrado de datos

A continuación, se carga el conjunto de datos, los registros específicos para México son seleccionados, enfocándose en el consumo de energía renovable, en el siguiente código se muestra la carga de los datos.

```
data = pd.read_csv('WorldSustainabilityDataset.csv')
mexico_data = data[(data['Country'] == 'Mexico') & (data['Year'] >= 2000)]
mexico_data = mexico_data[['Year', 'RenewableEnergyConsumption']].dropna()
```

Se utiliza `pd.read_csv()` para cargar los datos desde un archivo CSV. Posteriormente, se filtran los datos para México utilizando operaciones de indexación booleana en Pandas. La función `.dropna()` se emplea para eliminar cualquier fila con valores faltantes, asegurando así la integridad de los datos.

### 3. Preparación de los datos

Se utiliza `pd.read_csv()` para cargar los datos desde un archivo CSV. Luego, los datos se filtran para México empleando operaciones de indexación booleana en Pandas. La función `dropna()` se aplica para eliminar cualquier fila con valores faltantes, con el fin de asegurar la integridad de los datos. A continuación, se transforma la serie temporal en un conjunto de datos adecuado para el entrenamiento de la red LSTM, dividiendo los datos en conjuntos de entrenamiento y prueba, como se muestra en el siguiente fragmento de código:

```
years = mexico_data['Year'].values
consumption = mexico_data['RenewableEnergyConsumption'].values
n_steps = 3
X, y = split_sequence(consumption, n_steps)
n_features = 1
X = X.reshape((X.shape[0], X.shape[1], n_features))
```

Aquí, se preparan las secuencias de datos para el entrenamiento, `n_steps` define el número de pasos temporales que se considerarán para predecir el siguiente valor, `n_features` representa el número de atributos utilizados para la predicción en cada paso de tiempo, que en este caso es solo uno: el consumo de energía renovable. La operación `reshape` modifica la forma de `X` para ajustarse a lo que la capa LSTM espera: un tensor tridimensional de la forma `[samples, timesteps, features]`.

### 4. Construcción del modelo LSTM

El modelo LSTM es creado para ser ajustado a los datos. Se añaden capas de LSTM y Dropout para modelar la secuencia y prevenir el sobreajuste, como se ilustra en el siguiente fragmento de código:

```
model = Sequential()
model.add(LSTM(100, activation='relu', return_sequences=True, input_shape=(n_steps,
n_features)))
model.add(Dropout(0.2))
model.add(LSTM(100, activation='relu'))
model.add(Dropout(0.2))
model.add(Dense(1))
model.compile(optimizer='adam', loss='mse', metrics=['mse', 'mae', 'mape', 'ac-
curacy'])
```

La red se inicia con una capa Sequential, que facilita el apilamiento de capas de red una tras otra. Cada capa LSTM contiene 100 unidades y utiliza la función de activación *relu*. La capa Dropout, con un valor de 0.2, descarta aleatoriamente el 20 % de las características durante el entrenamiento para combatir el sobreajuste. La capa Dense al final emite la predicción de un solo valor continuo. Finalmente, se compila el modelo utilizando el optimizador *adam* y la función de pérdida de MSE.

## 5. Entrenamiento y evaluación del modelo

El modelo es entrenado utilizando los datos, utilizando técnicas de validación y Early Stopping para optimizar el proceso de aprendizaje y prevenir el sobreajuste. Posteriormente, se evalúa su rendimiento, como se muestra en el siguiente fragmento de código:

```
early_stopping = EarlyStopping(monitor='val_loss', patience=20, restore_best_
weights=True)
history = model.fit(X_train, y_train, epochs=200, validation_split=0.2, callbacks=[ear-
ly_stopping], verbose=1)
mse, mae, mape, acc = model.evaluate(X_test, y_test, verbose=0)
```

*EarlyStopping*, una función de *Keras*, se utiliza para interrumpir el entrenamiento cuando una métrica monitoreada, en este caso la pérdida de validación (*val\_loss*), deja de mejorar. La propiedad *patience* establece el número de épocas que se esperan sin mejora antes de detener el entrenamiento. Se emplea *model.fit()* para entrenar el modelo, especificando los datos de entrenamiento y validación, así como el número de épocas. Después del entrenamiento, el modelo se evalúa con *model.evaluate()*, obteniendo métricas como el MSE, el error absoluto medio (MAE), el error porcentual absoluto medio (MAPE), y la precisión.

## 6. Visualización de resultados

Finalmente, se calcula y se muestra el rendimiento del modelo utilizando la biblioteca `matplotlib`, como se muestra en el siguiente fragmento de código:

```
predictions = model.predict(X_test, verbose=0)
test_years = years[n_train+n_steps:]
real_consumption = y_test
predicted_consumption = predictions.flatten()

plt.figure(figsize=(12,6))
plt.plot(years[:n_train], consumption[:n_train], label='Entrenamiento')
plt.plot(test_years, real_consumption, label='Real', color='orange')
plt.plot(test_years, predicted_consumption, label='Predicción', linestyle='--', col-
or='green')
plt.title('Consumo de Energía Renovable en México')
plt.ylabel('Consumo de Energía Renovable')
plt.xlabel('Año')
plt.legend()
plt.show()
```

Aquí, se generan las predicciones del modelo y se preparan los datos de tiempo y consumo para la visualización. Se utiliza `plt.plot()` para trazar los datos de entrenamiento, los datos reales de prueba y las predicciones. Este gráfico ayudará a visualizar cómo las predicciones del modelo se alinean con los valores reales, proporcionando una representación intuitiva de la precisión del modelo.

El ejemplo ilustra meticulosamente la secuencia de pasos cruciales en la aplicación de las RNR para desentrañar el patrón del consumo de energía renovable en México. Se inicia con la configuración de un entorno de programación en Google Colab, seguido de la importación de bibliotecas de análisis de datos y aprendizaje automático. Posteriormente, se realiza la recolección de datos ambientales específicos de México, enfocándose en el consumo de energía renovable, y se prepara el conjunto de datos a través de métodos de limpieza, normalización y transformación para su posterior análisis con la red LSTM.

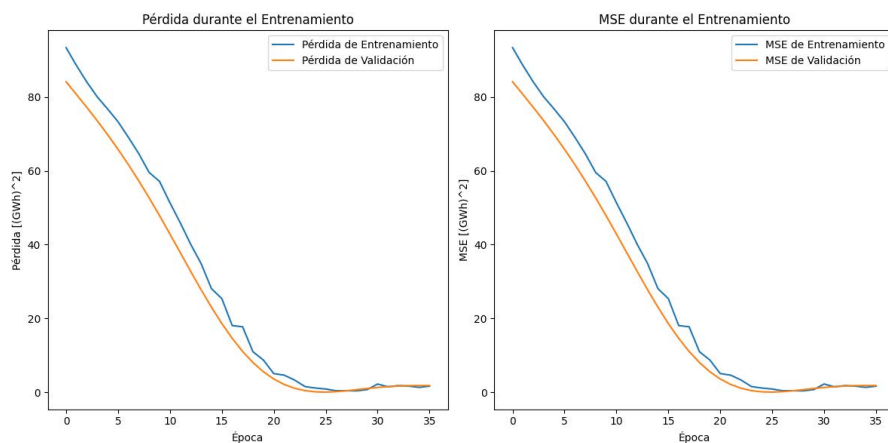
El proceso continúa con la estructuración de los datos en formatos adecuados para las LSTM y la construcción de un modelo de red neuronal que incorpora capas LSTM y Dropout, diseñado para capturar tanto las dependencias a corto como a largo plazo en los datos temporales. La configuración de los parámetros del modelo, la selección de funciones de activación y la inicialización de pesos, se realizan de manera reflexiva para adaptarse a la complejidad y naturaleza de los datos.

## Resultados y discusión

### *Análisis de la evolución del aprendizaje del modelo*

Esta sección presenta una evaluación exhaustiva de los resultados obtenidos a través del modelo propuesto. Se detalla la eficacia del algoritmo, sus patrones de aprendizaje y la relevancia de las métricas obtenidas en el marco del estudio. La figura 2 ilustra la evolución del aprendizaje del modelo a lo largo de las épocas de entrenamiento, proporcionando información crítica sobre la eficacia del proceso de aprendizaje automático y la capacidad de generalización del modelo. Las dos gráficas que componen la figura exhiben la dinámica de la pérdida y el error cuadrático medio (MSE), que son indicadores clave del progreso y la precisión del modelo en la tarea en cuestión.

Figura 2. Evaluación y evolución del rendimiento y aprendizaje del modelo.



Fuente: Elaboración propia.

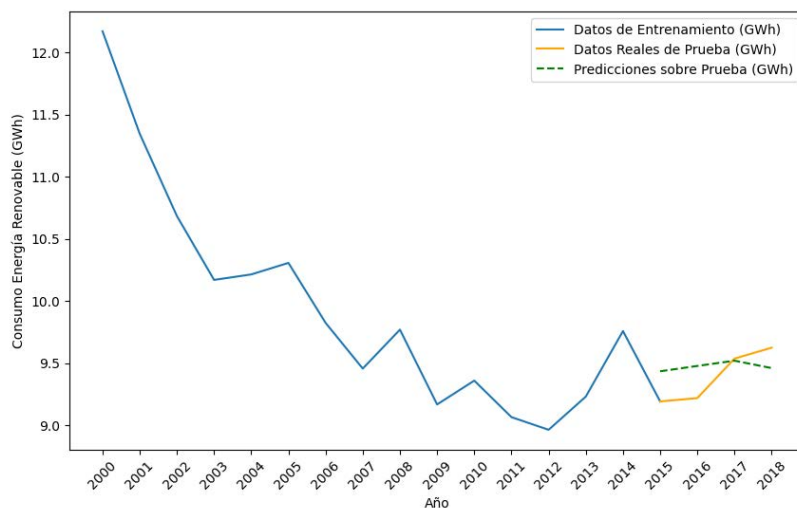
La gráfica izquierda de la figura 2 muestra la curva de pérdida durante el proceso de entrenamiento del modelo. La pérdida de entrenamiento y la pérdida de validación disminuyen conforme se suceden las épocas, una tendencia que es indicativa de un aprendizaje efectivo. La pérdida se mide generalmente como la discrepancia entre las predicciones del modelo y los valores verdaderos de los datos. Si las líneas comenzaran a divergir, con la pérdida de validación aumentando o estancándose mientras la de entrenamiento sigue disminuyendo,

sería indicativo de sobreajuste (overfitting). En la gráfica derecha, se muestra el MSE (Mean Squared Error) durante el entrenamiento, que es una medida cuantitativa del error del modelo. Al igual que con la gráfica de pérdida, una disminución en el MSE para los conjuntos de entrenamiento y de validación es deseable y sugiere un aprendizaje adecuado.

## Comparación de predicciones frente a datos reales

El análisis detallado de las proyecciones del modelo se visualiza en la figura 3, la cual ofrece una comparativa entre las proyecciones del modelo y los datos reales de consumo de energía renovable, poniendo a prueba su precisión predictiva.

Figura 3. Análisis de tendencias de consumo real vs predicciones.



Fuente: Elaboración propia.

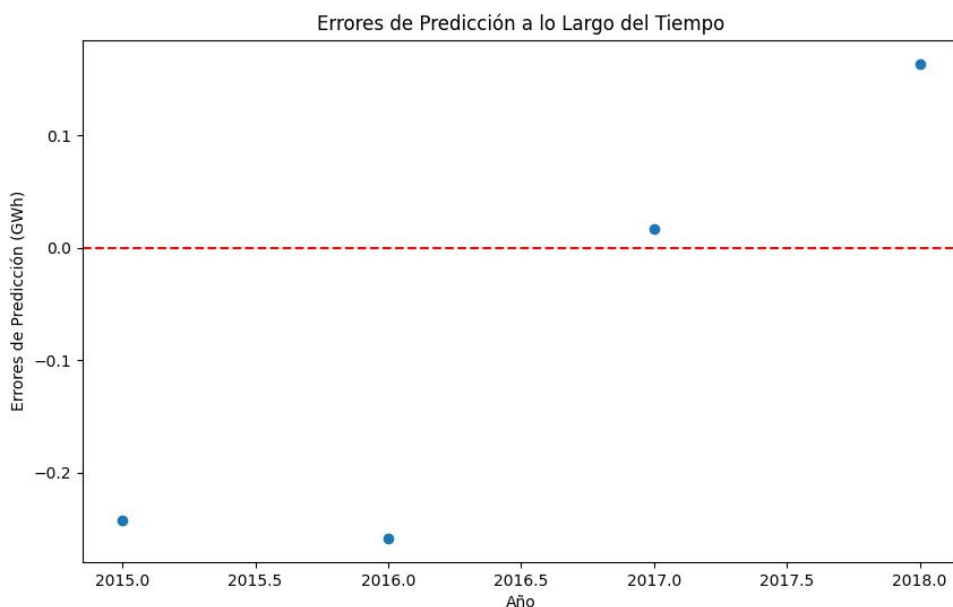
La correspondencia observada entre las predicciones y los valores reales durante el periodo de prueba no solo refleja la tendencia general sino también las variaciones específicas en el consumo a lo largo del tiempo, lo cual demuestra la capacidad del modelo para imitar el comportamiento histórico del consumo energético. Se observa que las predicciones replican un patrón de incremento seguido de un descenso, alineándose con las fluctuaciones inherentes a los datos históricos. Esta coherencia subraya la relevancia del modelo para anticipar tendencias futuras en la gestión de energía renovable. La línea de predicción

discontinua, que se mantiene cercana a los datos reales, sugiere un modelo adecuadamente ajustado. Sin embargo, la presencia de desviaciones en los puntos finales señala la necesidad de investigar la flexibilidad del modelo frente a variaciones inesperadas, lo que es fundamental para su aplicación en escenarios dinámicos de consumo energético.

### **Análisis de los errores de predicción**

Finalmente, la figura 4 detalla los errores de predicción individuales a lo largo del tiempo, lo que nos permite evaluar en qué momentos y en qué medida el modelo se desvía de la realidad.

Figura 4. Errores de predicción del modelo.



Fuente: Elaboración propia.

Aunque hay errores en las predicciones, la mayoría de los puntos están cerca de la línea de error cero, lo que indica que las predicciones son generalmente precisas. Las desviaciones proporcionan una oportunidad para el análisis posterior, potencialmente señalando áreas donde el modelo podría ser refinado para mejorar la precisión.

Basado en las gráficas proporcionadas, el análisis del conjunto de datos *WorldSustainabilityDataset.csv* se centra en la evaluación del consumo de energía renovable en México y cómo este puede ser proyectado en el futuro utilizando un modelo LSTM. El resultado de este análisis ofrece varias interpretaciones significativas en el contexto de la sostenibilidad y la gestión de recursos energéticos:

#### *Interpretación de tendencias históricas:*

La trayectoria descendente en el consumo de energía renovable hasta 2015, evidenciada en los datos de entrenamiento, puede ser el reflejo de múltiples factores interrelacionados, incluyendo variaciones económicas que afectan la viabilidad de las inversiones en tecnologías limpias y la posible influencia de subsidios o incentivos gubernamentales. Además, este patrón podría estar ligado a la conciencia social y la demanda del mercado, que a su vez son impulsados por la percepción pública y la educación sobre las ventajas de las energías renovables.

#### *Evaluación de la capacidad predictiva del modelo:*

Las predicciones generadas por el modelo, aunque no se alinean perfectamente con los datos reales, capturan un cambio significativo en la tendencia que sugiere un aumento en el consumo de energía renovable seguido de una estabilización o disminución. Esto podría reflejar la volatilidad en el mercado de energías renovables o variaciones en la adopción de estas energías que podrían ser críticas para la planificación futura.

#### *Implicaciones para la planificación y política energética:*

Los resultados del análisis aportan una base sólida para la formulación de políticas y la planificación energética. Al comprender mejor las tendencias históricas y las proyecciones futuras, los responsables de la toma de decisiones pueden diseñar estrategias más efectivas que no solo apunten a mejorar la infraestructura existente sino también a incentivar la integración de energías renovables en el tejido energético nacional. Esto podría traducirse en el desarrollo de nuevos proyectos de energía renovable, la actualización de los marcos regulatorios para facilitar la adopción de tecnologías limpias y la implementación de programas de subsidios y créditos fiscales que estimulen tanto la producción como el consumo de energías renovables.

### *Influencia en la inversión en energías renovables:*

Los inversores y empresas del sector energético podrían emplear los datos analizados para discernir tendencias emergentes y periodos críticos en la demanda de energía renovable, lo cual es vital para la toma de decisiones estratégicas. La capacidad de prever con anticipación los ciclos de auge o recesión en la adopción de energías renovables brinda una ventaja competitiva en el mercado, permitiendo a las empresas ajustar sus carteras de inversión y estrategias de crecimiento de manera que se alinee con las proyecciones de demanda y las metas de sostenibilidad. Además, las previsiones derivadas del modelo pueden ser cruciales para la optimización de operaciones logísticas, como la cadena de suministro y la gestión de inventarios, así como para la planificación estratégica de la capacidad productiva y de infraestructura. Esta optimización garantiza que las inversiones estén sincronizadas con los pronósticos de demanda futura, permitiendo una asignación de recursos más eficiente y una respuesta ágil a las fluctuaciones del mercado.

## **Conclusiones**

El análisis minucioso que emplea las RNR con arquitectura de LSTM para desglosar el conjunto de datos “WorldSustainabilityDataset.csv” ha sido revelador en la identificación de patrones complejos del consumo de energía renovable en México. La metodología se centra en la predicción de series temporales y ha probado ser un enfoque robusto y fiable, destacando su capacidad para captar la dinámica subyacente y las tendencias a lo largo del tiempo. Esta capacidad analítica no solo detecta las tendencias existentes sino que también ofrece la posibilidad de proyectar futuros escenarios, permitiendo a los stakeholders anticipar y planificar en consecuencia. Además, este enfoque metodológico podría ser replicado en otros contextos geográficos y sectoriales, lo que sugiere su aplicabilidad y valor en un amplio espectro de desafíos relacionados con la sostenibilidad energética.

La evaluación del modelo reveló que las RNR pueden prever con una precisión razonable la evolución del consumo de energía renovable, una tarea crucial para la planificación de la sostenibilidad energética. Aunque se observaron desviaciones entre las predicciones y los datos reales, estos errores proporcionaron una visión crítica sobre los aspectos del modelo que podrían mejorarse, como la necesidad de refinar los hiperparámetros o incorporar datos adicionales que podrían influir en la precisión predictiva.

Los resultados obtenidos subrayan la utilidad del método LSTM en la interpretación de datos complejos y en la proyección de tendencias futuras en el sector de la energía renovable. El modelo no solo ofreció proyecciones de la demanda de energía renovable, sino que también destacó la posibilidad de cambios en las tendencias de consumo, lo que es esencial para adaptarse a futuras necesidades y para la toma de decisiones basadas en datos.

El estudio confirmó que las técnicas avanzadas de aprendizaje profundo, como las RNR y LSTM, son herramientas valiosas para los analistas de datos y responsables de la formulación de políticas en el ámbito de la sostenibilidad energética. Al aplicar estas técnicas al problema estudiado, se descubrieron patrones significativos y se generaron predicciones útiles que pueden informar estrategias futuras para la gestión sostenible de recursos energéticos renovables. Este enfoque analítico, por tanto, tiene el potencial de contribuir de manera significativa a la optimización de recursos y a la elaboración de políticas energéticas eficientes y sostenibles.

## Referencias

- Alonso, J. I. (2022). *Utilización de redes neuronales recurrentes en la predicción de tendencias del mercado de harina de soja*.
- Bobadilla, J. (2021). *Machine learning y deep learning: Usando Python, Scikit y Keras* (1a ed.). Ediciones de la U; Editorial Ra-Ma.
- García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., & Herrera, F. (2016). Big data preprocessing: Methods and prospects. *Big Data Analytics*, 1(1), 9. <https://doi.org/10.1186/s41044-016-0014-0>
- Guamán Pachacama, J. A., & Segura Muñoz, G. B. (2021). *Red neuronal Long Short-Term Memory (LSTM) aplicada a series temporales para pronosticar consumo energético en edificaciones*. [B.S. thesis]. Universidad de Guayaquil. Facultad de Ciencias Matemáticas y Físicas.
- Jaquenod De Zsögön, S. (2019). *Antropología ambiental*. Dykinson.
- Joseph, F. J. J., Nonsiri, S., & Monsakul, A. (2021). Correction to: Keras and TensorFlow: A Hands-On Experience. En K. B. Prakash, R. Kannan, S. A. Alexander, & G. R. Kanagachidambaresan (Eds.), *Advanced Deep Learning for Engineers and Scientists* (pp. C1–C1). Springer International Publishing. [https://doi.org/10.1007/978-3-030-66519-7\\_12](https://doi.org/10.1007/978-3-030-66519-7_12)
- Landázuri Páez, L. L., Gallo Perugachi, K. M., & Estrella Tapia, D. F. (2023). Sistema de Monitoreo/Control de Consumo de energía eléctrica en el hogar mediante Raspberry Pi y Python.: Home energy Consumption monitoring/

- control System using Raspberry Pi and Python. *Revista Científica Multidisciplinar G-nerando*, 4(2). <https://doi.org/10.60100/rcmg.v4i2.126>
- Lindemann, B., Maschler, B., Sahlab, N., & Weyrich, M. (2021). A survey on anomaly detection for technical systems using LSTM networks. *Computers in Industry*, 131, 103498. <https://doi.org/10.1016/j.compind.2021.103498>
- Nguyen, H. D., Tran, K. P., Thomassey, S., & Hamad, M. (2021). Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management. *International Journal of Information Management*, 57, 102282. <https://doi.org/10.1016/j.ijinfomgt.2020.102282>
- Pang, B., Nijkamp, E., & Wu, Y. N. (2020). Deep Learning With TensorFlow: A Review. *Journal of Educational and Behavioral Statistics*, 45(2), 227–248. <https://doi.org/10.3102/1076998619872761>
- Quiguri Daquilema, C. M. (2023). *Contrastando el Machine Learning y la Econometría en series temporales*. [B.S. thesis]. Quito: EPN, 2023.
- Rengasamy, D., Jafari, M., Rothwell, B., Chen, X., & Figueredo, G. P. (2020). Deep Learning with Dynamically Weighted Loss Function for Sensor-Based Prognostics and Health Management. *Sensors*, 20(3), 723. <https://doi.org/10.3390/s20030723>
- Rodríguez Morales, V., Bustamante Alfonso, L. M., & Mirabal Jean-Claude, M. (2011). La protección del medio ambiente y la salud, un desafío social y ético actual. *Revista Cubana de salud pública*, 37, 510–518.
- Salazar Molina, E. A. (2023). *Metodología para la Evaluación en Proyectos de Energía Solar Fotovoltaica* [Doctoral thesis, Universidad Internacional Iberoamericana México]. <https://repositorio.unini.edu.mx/id/eprint/6010>
- Torres, V. (2021, julio). *La Sustentabilidad Como Estrategia Para El Desarrollo De La Industria*. CONGRESO CIENTÍFICO MULTIDISCIPLINARIO LATAM 2021 Investigación en latinoamérica. <http://cathi.uacj.mx/20.500.11961/19976>
- Vega Moreno, B. D. (2021). *Diseño y desarrollo de un sistema de recomendación basado en filtrado colaborativo utilizando datos secuenciales mediante redes neuronales recurrentes* [B.S. thesis].
- Veléz, A., Mera, C. A., Orduz, S., & Branch, J. W. (2021). Generación de péptidos antimicrobianos mediante redes neuronales recurrentes. *Dyna*, 88(216), 210–219.
- Wang, J., & Dowling, A. W. (2022). Pyomo.DOE: An open-source package for model-based design of experiments in Python. *AIChE Journal*, 68(12), e17813. <https://doi.org/10.1002/aic.17813>

# Capítulo 8

---

## **Cuantificación de recursos naturales pétreos a partir de métodos aerofotogramétricos y análisis en sistemas de información geográfica**

*Román Edén Parra-Galaviz  
Manuel de Jesús Pérez-Valdez  
Ciro de la Concepción Díaz-Hernández*

## Introducción

En la actualidad, es imprescindible hablar acerca de la educación ambiental como punto de partida para la preservación y conservación de los espacios y/o recursos naturales (Chirinos et al., 2017), con el objetivo de lograr una mejor planificación del desarrollo sostenible, el cual se concibe según (Barbera et al., 2019) como un proceso continuo orientado al logro de metas o propósitos comunes, establecidos mediante el diálogo y la comunicación entre los actores de una organización, comunidad, etc., para alcanzar una mejor calidad de vida.

Por lo anterior, es de suma importancia buscar la relación y el trabajo coordinado entre investigadores de distintas universidades públicas, cuya finalidad esté orientada no únicamente en la solución de problemáticas locales, sino que a través de la implementación de diferentes metodologías, se aporte en buena medida a la agenda 2030 y los Objetivos del Desarrollo Sostenible (ODS).

En la última década se ha implementado el uso de drones y barredores laser como LIDAR (Light Detetion and Ranging), para cartografiar parte de la superficie terrestre de manera eficiente (Parra et al., 2018), sin dejar de lado el uso de aeronaves tripuladas, por lo que, en la actualidad se desarrollan numerosos y diversos estudios de sostenibilidad y medioambiente que se basan en la interpretación de fotografías aéreas, tanto satelitales como desde drones y aviones tripulados, las cuales se han convertido en un recurso esencial en el ámbito de la Cartografía y los Sistemas de Información Geográfica (SIG).

Estas imágenes aéreas no solo permiten la observación y el análisis, sino que también son la base para la formulación de hipótesis, la creación de modelos y la generación de productos que brindan una representación precisa de los fenómenos estudiados. A esta ciencia se le conoce como Fotogrametría, la cual se define como el conjunto de métodos y procedimientos mediante los cuales podemos deducir de la fotografía de un objeto, la forma y dimensiones del mismo (Schenk, 2002; Lerma, 2002; Santamaría y Sanz, 2011), en donde cada imagen se modela idealmente como una proyección central (Remondino y Fraser, 2006).

De esta forma, se posibilita la generación de información espacial y la topografía en dos o tres dimensiones a partir de las características y elementos visibles en dos o más imágenes (Westoby et al., 2012) so that data acquisition is often passed on to specialist third party organisations. The high costs of

data collection are, for many applications in the earth sciences, exacerbated by the remoteness and inaccessibility of many field sites, rendering cheaper, more portable surveying platforms (i.e. terrestrial laser scanning or GPS que juntamente con los paquetes de software como Arc3D, Photosynth/Bundler, AutoDesk PhotoFly y Agisoft Photoscan (Agisoft Metashape, en su versión reciente), permiten la reconstrucción de escenas y modelos (Dellepiane et al., 2013; De Reu et al., 2014; Koutsoudis et al., 2014).

Sobre este tema, en México, la Dirección General de Geografía (DGG) ha realizado la captura de aproximadamente 700,000 aerofotografías, desde aeronaves tripuladas con cámaras fotogramétricas, que abarcan todo el territorio nacional, para la creación de cartografía topográfica de diferentes escalas, además de su uso en ciertos proyectos especiales que incluyen el cubrimiento de islas y la realización de tomas para necesidades concretas como situaciones de desastre o proyectos específicos de fotointerpretación de recursos naturales (INEGI, 2000).

En este sentido, el presente capítulo aborda una metodología de tipo experimental con enfoque cuantitativo para inferir en los volúmenes de excavación de recursos naturales pétreos a partir de técnicas de fotogrametría digital y análisis SIG.

En la propuesta, se plantea la toma de fotografías aéreas con cámaras digitales no fotogramétricas (convencionales) en el rango visible, desde aeronaves tripuladas, y el uso de software tanto en el procesamiento digital de imágenes como en el análisis de los modelos obtenidos, con el objetivo de proponer un método cuantitativo que sirva como instrumento de medición y, además, ilustre e invite al lector a conocer el soporte teórico del que dependen modernas tecnologías de estudios geoespaciales, como Drones y LIDAR, adentrándose en el antecedente de la fotogrametría clásica y su formulación elemental.

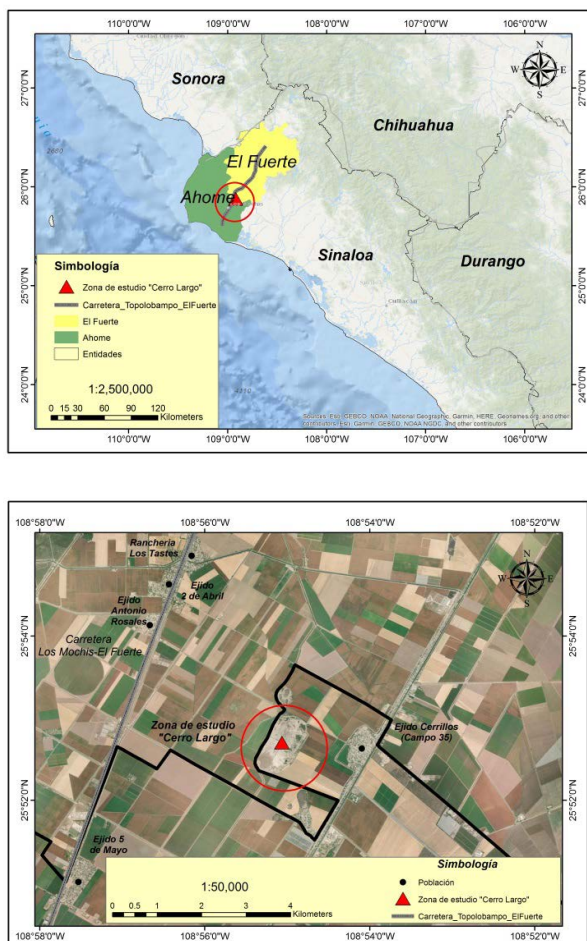
## **Materiales y métodos**

### **2.1 Zona de estudio**

La investigación se realizó en el banco de material pétreo denominado “Cerro Largo”, ubicado a 8 kilómetros al Noreste de la ciudad de Los Mochis, Sinaloa, en las coordenadas Latitud Norte 25°52'26.91" y Longitud Oeste 108°55'4.24", entre los límites de los municipios de Ahome y El Fuerte, del mismo estado, con una superficie aproximada de 300 hectareas (figura 1).

Esta zona está siendo explotada por empresas particulares para fines de la construcción y se ha observado un acelerado ritmo de explotación desde febrero del 2016 a la fecha (2023), debido a la demanda de material pétreo en diferentes obras ingenieriles en ambos municipios colindantes, principalmente.

Figura 1. Zona de estudio: banco de materiales pétreos “Cerro Largo”.



Fuente: Elaboración propia.

## 2.2 Metodología

La reconstrucción de objetos en el espacio como taludes o afloramientos es posible mediante herramientas fotogramétricas a partir de algoritmos computacionales y técnicas denominadas estructuras obtenidas de movimiento (SFM), que permiten obtener Modelos Digitales de Elevación (MDE), mosaicos geo-referenciados tridimensionales, entre otros (Vasuki, et al., 2014), y son precisamente este tipo de insumos los que permiten realizar un análisis minucioso de los elementos cartografiados.

En este sentido, Parra et al., (2021) mencionan que los SIG tienen la capacidad para construir modelos o representaciones del mundo real a partir de las bases de datos digitales y que estos modelos pueden utilizarse para la simulación de los efectos que un proceso de la naturaleza o una acción antrópica origina sobre un determinado escenario en una época específica, realizando proyecciones, a través de un análisis multitemporal de la información geográfica generada.

La metodología utilizada en esta investigación es de tipo experimental con enfoque cuantitativo, y consiste en la generación y validación de insumos geoespaciales de diferentes épocas (multitemporales) del banco de materiales pétreos “Cerro Largo” (como ortofotografías y modelos digitales de elevación de alta resolución), para el monitoreo de la extracción volumétrica de materiales a partir del procesamiento digital de imágenes aéreas tomadas desde aeronaves tripuladas y su análisis en SIG.

Para ello, se utilizaron los siguientes equipos y softwares: Aeronave tripulada Aerotrek A240, cámara digital Hasselblad 39 Mpixel con lente de 35 mm, plataforma portacámara, antenas GPS de doble frecuencia, Software Fotogramétrico Digital (SFD) *Agisoft Metashape*, el programa para dibujo avanzado *AutoCad* y el programa para sistema de información Geográfica ArcGIS.

Con la finalidad de profundizar en la propuesta metodológica, la investigación se dividió en las siguientes fases: a) Cálculo de parámetros de vuelo, b) Señalización y medición de los puntos de control fotogramétricos, c) Vuelo fotogramétrico, d) Procesamiento Digital de Imágenes, e) Análisis en SIG.

### 2.2.1 Cálculo de parámetros de vuelo

El cálculo de los parámetros de vuelo en una misión fotogramétrica es un tema que suele presentarse al estudiar la fotogrametría. En la actualidad, los drones por ejemplo, realizan esta tarea de forma automatizada, relacionando las características del sensor de la cámara fotográfica, altura de vuelo, velocidad de

desplazamiento, traslapes entre fotografías, etc. Sin embargo, se hace relevante en esta investigación, ya que se plantea el uso de aeronaves tripuladas no automatizadas, profundizar en el uso e interpretación de la formulación elemental (Parra et al., 2022; Pérez et al., 2021; Santamaría y Sanz, 2011):

$$H = f * m \quad ; \quad m = H/f \quad (1)$$

Donde  $H$  es la altura de vuelo,  $f$  es la distancia focal de la cámara fotográfica y  $m$  es la escala de la imagen o bien la resolución espacial en cámaras digitales.

$$b = \frac{(L_x)(100-p)}{100} \quad (2)$$

Aquí, se calculó la base de fotografiado  $b$  que indica la cantidad de píxeles entre las coordenadas centrales de dos fotografías consecutivas, en función de la longitud en píxeles del eje  $X$  en la matriz de la cámara ( $L_x$ ) y el porcentaje de recubrimiento o traslape longitudinal  $p$ .

$$B = \frac{(L_x)(100-p)}{100} * m \quad ; \quad B = b * m \quad (3)$$

Con la expresión anterior, se determinó la base de fotografiado  $B$  que caracteriza la distancia o cantidad de metros entre los centros de dos fotografías consecutivas.

$$A = \frac{(L_y)(100-q)}{100} * m \quad (4)$$

Siendo  $A$  la separación o distancia en metros entre líneas de vuelo,  $L_y$  es la longitud en píxeles del eje  $Y$  en la matriz de la cámara y  $q$  el porcentaje de recubrimiento o traslape transversal.

$$I = \frac{3.6 B}{V} \quad (5)$$

Aquí,  $I$  es el intervalo de tiempo en segundos entre dos disparos consecutivos de la cámara,  $V$  es la velocidad de la aeronave, mientras que 3.6 es la constante de conversión entre  $Km/h$  y  $m/s$ .

A partir de esta formulación y las características de la Cámara Hasselblad de 39 Mpix con lente de 35 mm, considerando las alturas 1000, 700 y 500 metros, respecto al objeto de estudio, se confeccionaron los proyectos de levantamientos aéreos (líneas de vuelo o rutas de navegación de la aeronave). El cálculo de los parámetros para cada proyecto se muestra en la tabla 1.

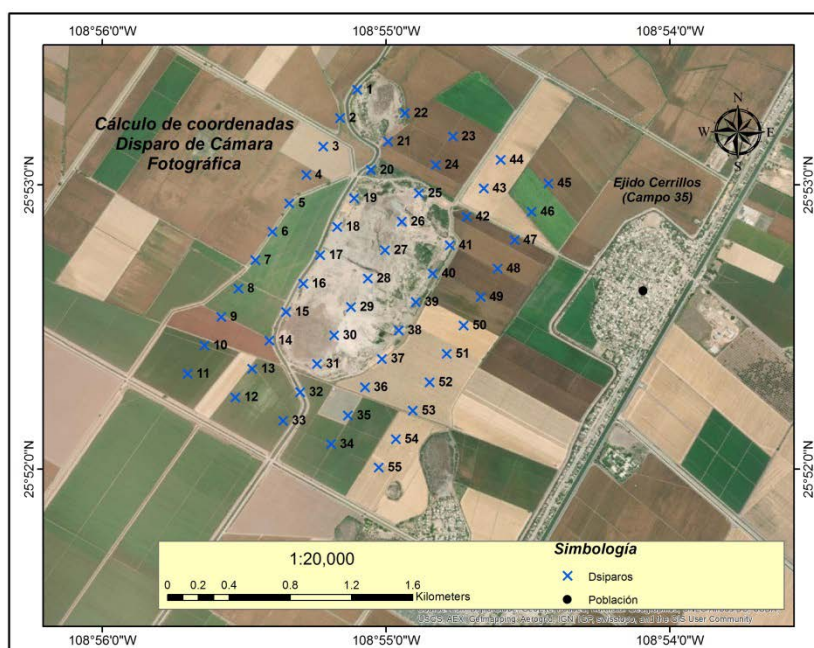
Tabla 1. Cálculo de parámetros de vuelo.

| Parámetros                                                | Proyecto 1 | Proyecto 2 | Proyecto 3 |
|-----------------------------------------------------------|------------|------------|------------|
| f Distancia focal (pix)                                   | 5,158.386  | 5158.386   | 5158.386   |
| Longitud en del eje X en la matriz (pix)                  | 7,212      | 7212       | 7212       |
| Longitud en del eje Y en la matriz (pix)                  | 5,412      | 5412       | 5412       |
| p Traslape longitudinal (%)                               | 70         | 70         | 70         |
| q Traslape transversal (%)                                | 40         | 40         | 40         |
| V Velocidad de la aeronave (km/h)                         | 100        | 100        | 100        |
| H Altura de vuelo (m)                                     | 1,000      | 700        | 500        |
| m Escala de la imagen o Resolución espacial (m)           | 0.194      | 0.136      | 0.097      |
| b Base fotográfica (pix)                                  | 2,163.6    | 2163.6     | 2,163.6    |
| B Base fotográfica (m)                                    | 420        | 295        | 210        |
| A Separación entre líneas de vuelo (m)                    | 630        | 442        | 315        |
| I Intervalo de tiempo entre dos disparos consecutivos (s) | 15.1       | 10.6       | 7.6        |

Fuente: Elaboración propia.

Retomando los parámetros calculados en el proyecto 3 (para ejemplificar), se determinaron las coordenadas de los disparos de la cámara durante la navegación sobre las líneas de vuelo. En total, para el cubrimiento total de la zona estudiada, se trazaron 5 líneas de vuelo con 11 disparos (fotografías) en cada una de ellas, dando un total de 55 imágenes para cubrir la zona de estudio (figura 2).

Figura 2. Proyección de líneas de vuelo y disparos de la cámara fotográfica.



Fuente: Elaboración propia.

### 2.2.2 Señalización y medición de los puntos de control fotogramétricos

Para poder realizar un procesamiento de imágenes aéreas, es de vital importancia la colocación y medición de puntos de control fotogramétricos (PCF), que son puntos del terreno claramente identificados en la fotografía, de los cuales se conocen las coordenadas precisas (Agüera et al., 2017) y permiten unir todas las fotografías en un solo bloque para una perspectiva integral del lugar estudiado. Estos se clasifican en Puntos de apoyo terrestre (GCP) y Puntos de validación (CCP), los primeros garantizan una geometría robusta para el

modelo, mientras que los CCP proporcionan una evaluación independiente de la precisión (James et al., 2020).

En el área de “Cerro Largo”, se colocaron un total de 15 PCF, de los cuales 10 fueron para el procesamiento de las imágenes y los otros 5 para la validación del modelo (10 como GCP y 5 como CCP). Las mediciones de sus coordenadas (tabla 2) se efectuaron con un GPS doble frecuencia (figura 3) en método RTK (Real Time Kinematic).

Tabla 2. Coordenadas de los PCF.

| <b>Coordenadas UTM zona 12 N WGS84</b> |            |             |        |
|----------------------------------------|------------|-------------|--------|
| PCF                                    | X          | Y           | Z      |
| 1                                      | 709286.102 | 2864576.316 | 22.184 |
| 2                                      | 708229.066 | 2862839.432 | 20.147 |
| 3                                      | 708665.120 | 2863737.932 | 41.360 |
| 4                                      | 708105.906 | 2864188.470 | 20.626 |
| 5                                      | 708898.256 | 2863510.596 | 21.729 |
| 6                                      | 709130.054 | 2864299.690 | 21.952 |
| 7                                      | 708860.701 | 2864271.661 | 49.523 |
| 8                                      | 708366.765 | 2863306.259 | 46.061 |
| 9                                      | 708219.947 | 2863218.749 | 21.636 |
| 10                                     | 708158.834 | 2863573.149 | 23.247 |
| 11                                     | 708553.441 | 2864483.411 | 22.293 |
| 12                                     | 708771.040 | 2864850.697 | 21.031 |
| 13                                     | 707921.084 | 2863875.289 | 20.820 |
| 14                                     | 709494.062 | 2864879.220 | 21.579 |
| 15                                     | 709735.164 | 2863984.629 | 21.355 |

Fuente: Elaboración propia.

Figura 3. Mediciones GPS de los puntos de control fotogramétricos.



Fuente: Elaboración propia.

### 2.2.3 Vuelo fotogramétrico

Se realizaron cuatro vuelos fotogramétricos comparativos; los primeros dos en octubre y diciembre de 2015 respectivamente, el tercero en septiembre del 2016 y el último en noviembre de 2023. La elección de épocas de muestreo se determinó para realizar el análisis de extracción multitemporal a corto, mediano y largo plazo en la dinámica de explotación del banco de materiales pétreos.

Las imágenes fueron tomadas con una cámara Hasselblad 39 Mpíxel, usando una lente de 35 mm, adaptada en una plataforma móvil (Figura 4a) sobre una aeronave ultraligera Aerotrek-A240 (figura 4b) con altitudes de 1000, 500, 700 y 500 metros, de acuerdo con el periodo de muestreo, retomando los valores de los parámetros de vuelo calculados en la tabla 1.

Se utilizó un disparador manual, navegando sobre las líneas de vuelo con sistema de posicionamiento global sincronizado con el sistema de navegación del avión (figura 5).

Figura 4. Toma de fotografías aéreas: (a) Cámara Hasselblad 39 Mpix.; (b) Aeronave Aerotrek A-240.



Fuente: Elaboración propia.

Figura 5. Sistema de Navegación de la aeronave.



Fuente: Elaboración propia.

#### 2.2.4 Procesamiento digital de imágenes

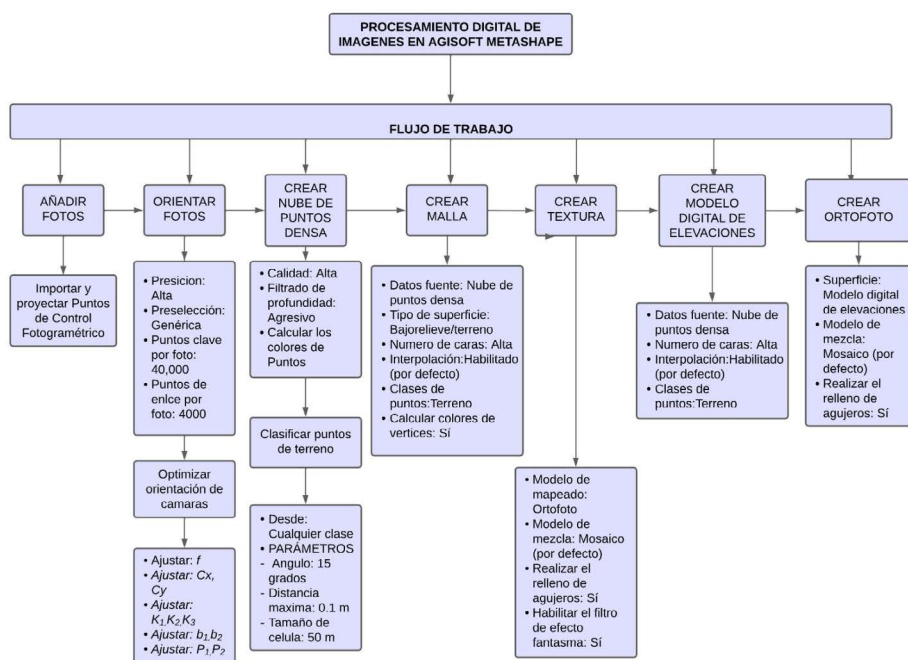
El Procesamiento Digital de Imágenes (PDI) se llevó a cabo en el software fotogramétrico *Agisoft Metashape* (anteriormente *Photoscan*). Para Peinado et al., (2014), este es un programa de fácil manejo, rápido y una herramienta

muy potente, tiene una gran velocidad de cálculo, y corrige automáticamente los errores de las deformaciones fotográficas y distorsiones del objetivo, ya que como lo mencionan Koutsoudis et al., (2014); Fernández et al., (2015), está basado en la adquisición de grandes densidades de puntos a partir de la correlación y utiliza procedimientos aplicados en el campo de la visión computacional y gráficos, como SfM (Structure from Motion) y MVS (Multi View Stereo), con los que se obtiene automáticamente un modelo 3D a partir de múltiples imágenes convergentes (Colica et al., 2021; Gonçalves & Henriques, 2015; Kyriakaki et al., 2014) with a daily average upload of 300 million photos. These photos, apart from documenting personal lives, often relate to experiences in well-known places of cultural interest, throughout several periods of time. Thus from the viewpoint of Cultural Heritage professionals, they constitute valuable and freely available digital cultural content. Advances in the fields of Photogrammetry and Computer Vision have led to significant breakthroughs such as the Structure from Motion algorithm which creates 3D models of objects using their 2D photographs. The existence of powerful and affordable computational machinery enables the reconstruction not only of single structures such as artefacts, but also of entire cities. This paper presents an overview of our methodology for producing cost-effective 4D – i.e. in space and time – models of Cultural Heritage structures such as monuments and artefacts from 2D data (pictures, video).

Las comparaciones entre softwares disponibles comercialmente, han señalado que Agisoft Metashape proporciona mayores tasas de alineación de imágenes, una resolución ortomosaica más fina y distancias de muestreo de suelo mejoradas (Fraser y Congalton, 2018; Przybilla et al., 2019), destacándolo como un paquete de software particularmente común utilizado para la producción de conjuntos de datos 3D a través de IBM (Hill et al., 2014; Jones y Church, 2020; Magnani et al., 2020; Verhoeven, 2011).

Además, tiene la ventaja de que la cámara fotográfica no necesita una calibración previa, por lo que nos brinda la posibilidad de usar cámaras digitales convencionales. Dicho esto y considerando las ventajas ofrecidas, el procedimiento empleado en esta investigación, para el procesamiento digital de imágenes, se basa en el flujo de trabajo Agisoft Metashape (esquema 1).

Esquema 1. Procesamiento digital de imágenes en Agisoft Metashape.

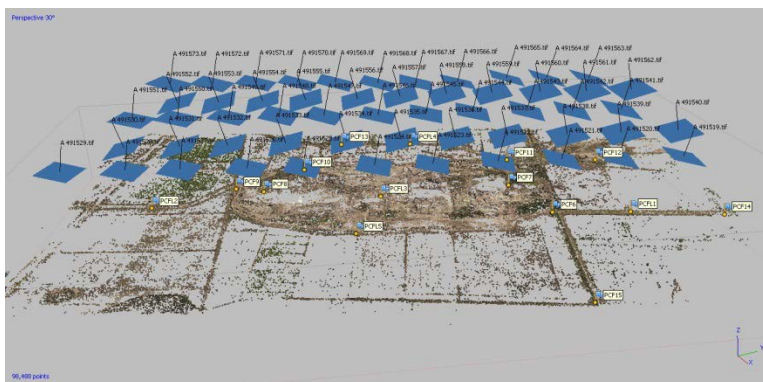


Fuente: Elaboración propia.

La primera etapa consistió en agregar tanto las imágenes de cada vuelo (en proyectos independientes) como las coordenadas de los PCF, marcando y midiendo cada uno de ellos en las imágenes proyectadas, declarando el sistema de coordenadas a emplear (UTM zona 12N WGS84).

Después, se realizó la orientación de las imágenes: la orientación interna, que describe la geometría particular de cada una de las fotografías y permite relacionar las coordenadas foto (mm) con las coordenadas imagen (píxel), la orientación relativa que permite determinar la disposición angular relativa y el desplazamiento de posición entre fotografías consecutivas al momento en que fueron tomadas y la orientación absoluta (a partir de PCF) para corregir las distorsiones geométricas (figura 6). Se efectuó una automatización al proceso de orientación, ajustándose los parámetros principales de la cámara fotográfica (autocalibración).

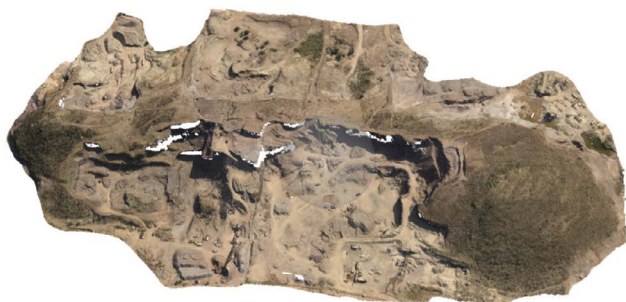
Figura 6. Orientación de Imágenes con PCF.



Fuente: Elaboración propia.

El segundo paso fue generar la nube de puntos densa basada en la imagen estimada. *Metashape* calcula la información de profundidad para cada una de las imágenes y posteriormente se combina una sola nube de puntos (figura 7). Se obtuvo una nube densa con un total de 5'079,440 puntos, lo que equivale aproximadamente a 1 punto por cada 0.5 m<sup>2</sup> en la zona fotografiada. Se utilizó la herramienta de clasificación de puntos de terreno, con el algoritmo automatizado, para detectar los puntos característicos fijados sobre el material pétreo.

Figura 7. Nube de puntos densa.

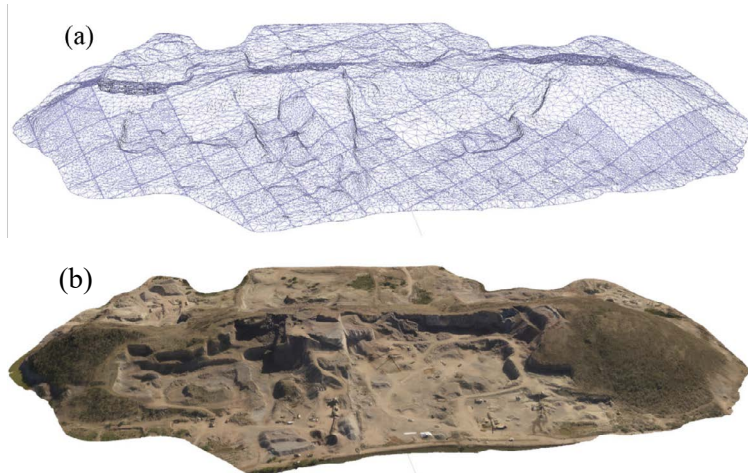


Fuente: Elaboración propia.

El tercer paso consistió en la construcción de red de triángulos irregulares o malla de triángulos 3D (figura 8a) a partir de la interpolación de los puntos en la nube densa, en la que se aplicó la textura en base al mosaico de imágenes

y el relleno de agujeros. Posteriormente se creó el MDT (Modelo Digital de Terreno) que consideró únicamente los puntos característicos del terreno natural, discriminado vegetación o cualquier objeto identificado como punto no perteneciente a esta categoría (figura 8b).

Figura 8. Modelos digitales de elevación: (a) Malla de triangulación 3D; (b) Modelo digital de terreno.



Fuente: Elaboración propia.

### 2.2.5 Análisis en SIG

En palabras de Olaya (2020), el SIG es un sistema de información diseñado para trabajar con datos referenciados mediante coordenadas espaciales o geográficas, o bien, es tanto un sistema de base de datos con capacidades específica para datos georeferenciados, como un conjunto de operaciones para trabajar con esos datos.

Por ello, la utilidad principal de un SIG radica en su capacidad para construir modelos o representaciones del mundo real a partir de las bases de datos digitales y en la utilización de esos modelos en la simulación de los efectos que un proceso de la naturaleza o una acción antrópica origina sobre un determinado escenario en una época específica, pudiendo comparar efectos con otras épocas o proyectar consecuencias.

En la actualidad los modelos constituyen un medio para lograr la representación del relieve muy versátil y funcional, ya que a partir del mismo no solo

se puede conocer la conformación o morfología del terreno sino también los elementos de origen antrópico y la vegetación presente en el mismo.

Para realizar el análisis comparativo entre volúmenes de material extraído en las diferentes épocas de “Cerro Largo”, de los 4 modelos raster generados en la investigación, se utilizó ArcGIS 10.8, un “software” de Sistema de Información Geográfica diseñado por la empresa californiana Enviromental Systems Research Institute (ESRI) para trabajar a nivel multiusuario.

## Resultados y discusión

### 3.1 Precisión y validación del procesamiento digital de imágenes

Se determinó la precisión del PDI mediante el estimador RMSE (Error Medio Cuadrático) de la posición de los GCP a partir de las correlaciones que estos guardan al interior del modelo (tabla 3). Por otra parte, la validación o confiabilidad en las mediciones se realizó mediante la comparativa del modelo orientado respecto a la posición espacial de los CCP y su proyección en las imágenes correspondientes (tabla 4).

$$RMSE (Ex) = \sqrt{\sum (Ex_i - \overline{Ex})^2 / n} \quad (6)$$

$$RMSE (Ey) = \sqrt{\sum (Ey_i - \overline{Ey})^2 / n} \quad (7)$$

$$RMSE (Exy) = \sqrt{RMSE (Ex)^2 + RMSE (Ey)^2} \quad (8)$$

$$RMSE (Ez) = \sqrt{\sum (Ez_i - \overline{Ez})^2 / n} \quad (9)$$

En las fórmulas anteriores,  $RMSE (Ex)$ ,  $RMSE (Ey)$  y  $RMSE (Ez)$  representan los errores medios cuadráticos para las coordenadas X, Y, Z respectivamente.  $RMSE (Exy)$  es el error medio cuadrático del modelo en el sistema cartesiano xy, mientras que  $Ex_i$ ,  $Ey_i$  y  $Ez_i$  son los errores de correlación de los puntos de control al interior del mismo, los cuales se comparan con los respectivos errores promedios  $(\overline{Ex}, \overline{Ey}, \overline{Ez})$ .

Tabla 3. RMSE de los GCP al interior del modelo.

| Nombre | Error X (cm) | Error Y (cm) | Error Z (cm) | Imagen (pix) |
|--------|--------------|--------------|--------------|--------------|
| GCP 1  | -4.913       | 12.839       | -16.605      | 0.540 (4)    |
| GCP 2  | 0.369        | -0.477       | -0.303       | 0.328 (3)    |
| GCP 3  | -3.876       | 8.926        | -1.553       | 0.503 (4)    |
| GCP 4  | 0.019        | 0.019        | 0.002        | 0.325 (3)    |
| GCP 5  | -0.684       | -2.663       | 0.980        | 0.321 (5)    |
| GCP 6  | 0.046        | -0.835       | 0.021        | 0.395 (7)    |
| GCP 7  | 3.471        | 0.712        | 8.795        | 0.554 (5)    |
| GCP 8  | 2.239        | 5.382        | 20.248       | 0.484 (3)    |
| GCP 9  | -0.658       | -1.404       | -0.314       | 0.153 (6)    |
| GCP 10 | 4.294        | -0.310       | -0.460       | 0.354 (3)    |
| RMSE   | 2.751        | 5.330        | 8.757        | 0.409        |

Fuente: Elaboración propia.

Tabla 4. RMSE de los CCP de validación en la precisión del modelo.

| Nombre | Error X (cm) | Error Y (cm) | Error Z (cm) | Imagen (pix) |
|--------|--------------|--------------|--------------|--------------|
| CCP 11 | -1.595       | -0.186       | 0.327        | 0.269 (7)    |
| CCP 12 | 43.797       | 13.232       | 16.505       | 1.119 (4)    |
| CCP 13 | -0.042       | -0.096       | -0.113       | 0.130 (3)    |
| CCP 14 | 0.666        | -0.258       | -0.506       | 0.035 (2)    |
| CCP 15 | 1.044        | 0.549        | -0.492       | 0.153 (3)    |
| RMSE   | 19.6074      | 5.92482      | 7.38984      | 0.545        |

Fuente: Elaboración propia.

Se puede observar que el *RMSE* ( $E_{xy}$ ) calculado a partir de los puntos de validación CCP es de 21.7 cm, mientras que el *RMSE* ( $E_z$ ) fue de 7.3 cm, suficiente precisión para considerar que los insumos obtenidos (en este caso de los MDT) tiene un alto grado de confiabilidad y se asemejan a las obtenidas con métodos de la topografía moderna (estaciones totales y sistemas GPS).

3.2 Insumos obtenidos: modelos digitales de terreno

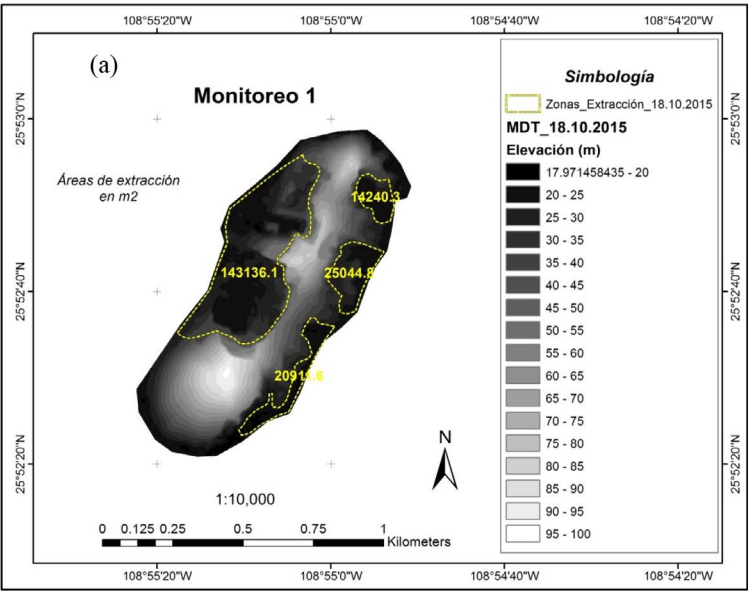
Se obtuvieron 4 MDT, de buena precisión, sobre el banco de materiales pétreos “Cerro Largo”, uno para cada monitoreo, clasificados por la altura, fecha de vuelo y sus respectivos parámetros geométricos; tamaño del raster y resolución espacial (tabla 5). Los modelos multitemporales fueron exportados de *Agisoft Metashape*, mediante el formato Raster “GEO TIFF”, para ser analizados en *ArcGis* y determinar las zonas de extracción de material (figura 9a, 9b, 9c y 9d).

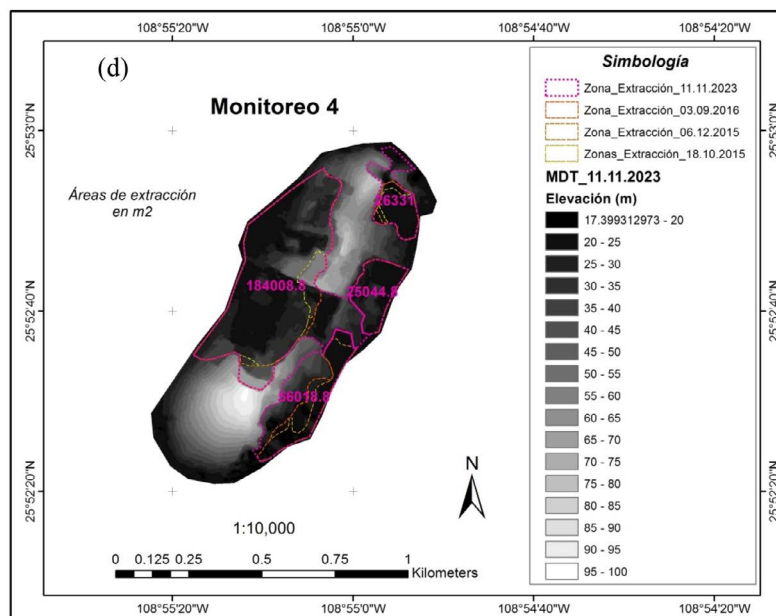
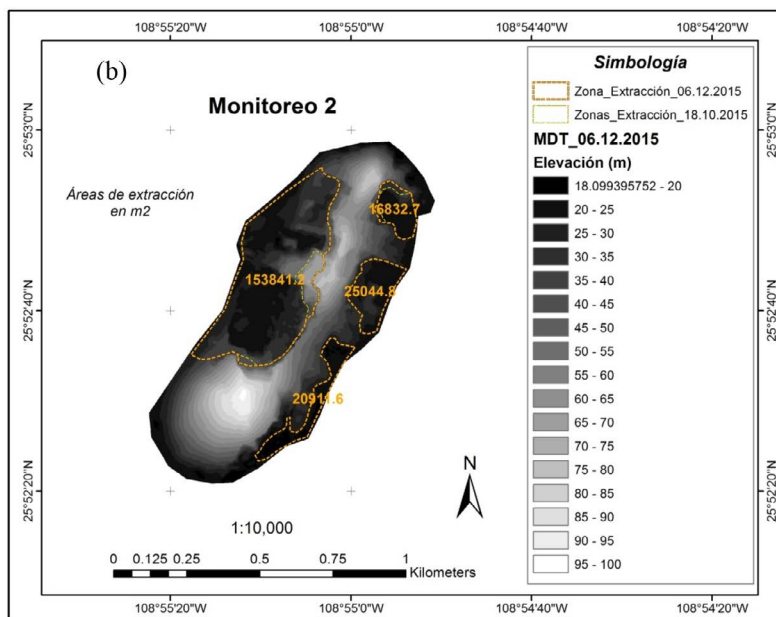
Tabla 5. Parámetros geométricos de los MDT.

| Monitoreo # | Altura de vuelo (m) | Fecha (dd/mm/aaaa) | Tamaño del Raster (pix) | Resolución espacial (cm/pix) |
|-------------|---------------------|--------------------|-------------------------|------------------------------|
| 1           | 1,000               | 18/10/2015         | 1,114*1,504             | 72                           |
| 2           | 500                 | 06/12/2015         | 2,239*3,021             | 38                           |
| 3           | 700                 | 03/09/2016         | 1,561*2,107             | 55                           |
| 4           | 500                 | 11/11/2023         | 4,061*5,481             | 21                           |

Fuente: Elaboración propia.

Figura 9. MDT multitemporales: (a) Monitoreo 1, MDT\_18/10/2015; (b) Monitoreo 2, MDT\_06/12/2015; (c) Monitoreo 3, MDT\_03/09/2016; (d) Monitoreo 4, MDT\_11/11/2023.





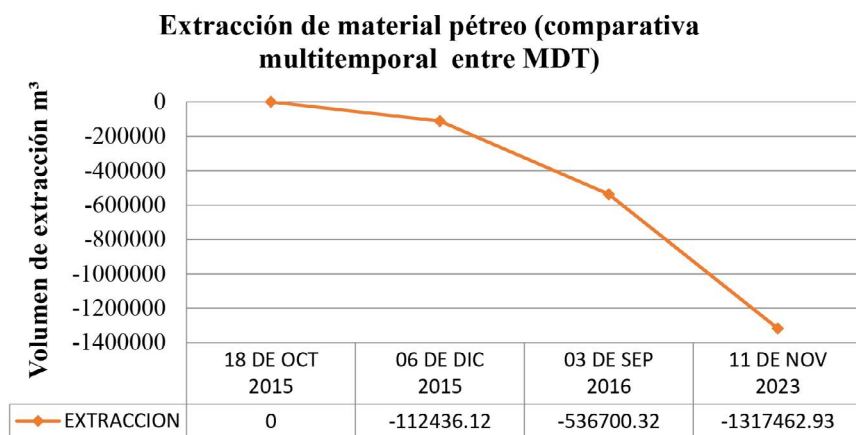
En este software SIG, se realizaron operaciones de análisis de superficies mediante las extensiones *3D analyst*, *Network analyst* y *Spatial analyst*, específicamente con la herramienta *area and volumen*, para obtener los datos volumétricos del objeto de estudio en las cuatro épocas de monitoreo. La separación elegida entre estratos o cortes altitudinales del terreno fue de cinco metros, lo que permitió realizar un análisis detallado de los parámetros geométricos de cada MDT (tabla 6) e inferir en las pérdidas de material de acuerdo a los diferentes periodos (gráfica 1).

Tabla 6. Parámetros geométricos de los MDT.

| Elevación del MDT (m) | Volumen de material pétreo (m³) |                |                |                |
|-----------------------|---------------------------------|----------------|----------------|----------------|
|                       | MDT_18/10/2015                  | MDT_06/12/2015 | MDT_03/09/2016 | MDT_11/11/2023 |
| 95                    | 466.15                          | 267.05         | 238.18         | 233.07         |
| 90                    | 14,793.53                       | 13,056.02      | 12,548.93      | 12,524.26      |
| 85                    | 53,719.36                       | 52,683.84      | 49,081.36      | 48,925.27      |
| 80                    | 129,688.77                      | 120,989.22     | 118,583.88     | 113,144.42     |
| 75                    | 273,240.25                      | 242,152.6      | 226,749.22     | 211,148.96     |
| 70                    | 496,255.11                      | 441,797.21     | 396,727.7      | 356,601.48     |
| 65                    | 807,734.68                      | 742,354.27     | 673,861.57     | 590,475.65     |
| 60                    | 1,222,253.86                    | 1,154,836.68   | 1,064,132.4    | 925,468.24     |
| 55                    | 1,754,755.09                    | 1,688,904.95   | 1,578,544.47   | 1,378,524.33   |
| 50                    | 2,421,839.55                    | 2,351,081.21   | 2,212,873.48   | 1,945,198.34   |
| 45                    | 3,231,173.79                    | 3,149,040.92   | 2,978,935.13   | 2,618,265.17   |
| 40                    | 4,178,883.73                    | 4,087,041.38   | 3,885,097.58   | 3,418,546.69   |
| 35                    | 5,294,737.07                    | 51,960,55.16   | 4,951,988.43   | 4,378,977.14   |
| 30                    | 6,645,339.54                    | 6,540,250.33   | 6,231,659.27   | 5,557,740.03   |
| 25                    | 8,367,082.93                    | 8,245,350.4    | 7,884,383.38   | 7,109,771.49   |
| 20                    | 10,665,095.99                   | 10,552,659.87  | 10,128,395.67  | 9,347,633.06   |

Fuente: Elaboración propia.

Gráfica 1. Volumen de extracción de material pétreo.



Fuente: Elaboración propia.

Considerando el monitoreo 1 como punto de partida para el cálculo de pérdida de volumen pétreo se determinó que, del día 18 de octubre de 2015 al 6 de diciembre del mismo año existe una pérdida de 1 % en un lapso de tiempo corto (un mes y medio).

Si comparamos el monitoreo 1 contra el 3 (del 18 de octubre de 2015 al 3 de septiembre de 2016) se calculó una pérdida del 5 % en un lapso de tiempo medio de 1 año. Por último, entre el primer monitoreo y actual (11 de noviembre de 2023) se calcula una extracción del 12.4 % en el lapso de tiempo largo, de ocho años.

Se observa que en los ocho años de monitoreo de la dinámica de explotación, el primer ciclo (del 2015-2016) mostró un acelerado ritmo de extracción de más de medio millón de metros cúbicos, en gran parte por las nuevas construcciones de fraccionamientos del sur de la ciudad de Los Mochis y las remodelaciones del puente a desnivel “El trébol” ubicado en la carretera México 15 con entronque en la misma ciudad.

Consideramos que el tiempo de pandemia 2020-2022, periodo crítico, ayudó en gran parte en la disminución de la dinámica observada en los estudios realizados a corto y mediano plazo. Sin embargo, en la actualidad, sigue siendo objeto de explotación.

## Conclusiones

Se destacan algunas ventajas importantes de los vuelos fotogramétricos como: la panorámica de conjunto ofrecida por las imágenes al momento de realizar un análisis integral de los recursos naturales, la calidad técnica en la obtención de resultados, la notoria reducción de tiempos y el alto nivel de detalle en las características de las zonas de estudio.

Los RMSE obtenidos en el procesamiento digital de imágenes, (Exy) y (Ez), calculados a partir de los puntos de validación CCP fueron de 21.7 y 7.3 cm respectivamente, que indican una precisión elevada en la generación de Modelos Digitales analizados para la cuantificación volumétrica. Además, se obtuvieron resoluciones espaciales entre 21 y 72 cm/pix, lo que confirma la importancia de los productos obtenidos, ya que estas resoluciones no se encuentran disponibles para usuarios SIG de forma gratuita.

Actualmente, la metodología presentada en la investigación está siendo utilizada por parte de profesores, investigadores, estudiantes de licenciatura y posgrado, en diferentes proyectos de investigación a nivel regional y estatal que requieren de análisis espacial, por ejemplo: para estudios de mapeos de gran cobertura como es el caso de la actualización cartográfica en la ciudad de Los Mochis y Guasave en Sinaloa, en cálculos de volumen de vegetación del cerro de la memoria, estudios para delimitación de zonas federales, generación de insumos de alta resolución para el estudio de inundaciones, análisis de afectación en comunidades rurales por avenidas máximas de agua del Río Fuerte, entre otras.

Los resultados expuestos en la presente investigación, evidencian que, al pasar de los años, la fotogrametría clásica no deja de ser una herramienta factible en la captura de información geoespacial y en este caso permitió de manera rápida, precisa y a un bajo costo, inferir en la dinámica de explotación de materiales pétreos. Por todo esto, podemos concluir que la metodología empleada, en cuanto a la combinación de la aerofotogrametría y el uso de software automatizados tanto para el procesamiento de imágenes como el análisis SIG, ofrece buenos resultados al momento de cuantificar recursos naturales.

En investigaciones futuras se pretende analizar la metodología empleada mediante el uso de Vehículos aéreos no tripulados (drones), donde se analizarán cálculos automatizados, sensores, software para la ejecución de planes de vuelo, entre otras características, ya que consideramos relevante experimentar, validar y exponer diferentes alternativas de métodos cuantitativos para recursos naturales.

Por último, los autores, invitan a estudiantes y profesores de áreas afines (en el uso y manejo de tecnologías geoespaciales), así como a nuevos convenios con cuerpos académicos, a realizar proyectos colaborativos mediante el uso de esta herramienta metodológica (se pone a disposición para trabajo interdisciplinario) en la solución de problemas, principalmente en el área de sostenibilidad y medioambiente, que impacten positivamente en los Objetivos del desarrollo Sostenible de la Agenda 2030.

## Referencias

- Agüera, F., Carvajal, F., & Martínez, P. (2017). Assessment of photogrammetric mapping accuracy based on variation ground control points number using unmanned aerial vehicle. *Measurement*, 98, 221–227. <https://doi.org/10.1016/j.measurement.2016.12.002>
- Barbera, N., Chirinos, Y., & Posada, M. (2019). La planificación del desarrollo sostenible desde los sistemas complejos adaptativos en Colombia. In Y. Chirinos, A. Ramírez, N. Barbera, & D. Rojas (Eds.), *Tendencias en la investigación universitaria. Una visión desde Latinoamérica. Volumen V* (pp. 24–42). Fondo Editorial Universitario Servando Garcés de la Universidad Politécnica Territorial de Falcón Alonso Gamero / Alianza de Investigadores Internacionales S.A.S.
- Chirinos, Y., González, C., Meriño, V., & Martínez, C. (2017). La educación ambiental en el proyecto educativo integral comunitaria: una reflexión desde la participación en comunidades escolares. In Y. Chirinos, A. García, E. Rodríguez, C. Pérez, & L. Camejo (Eds.), *Tendencias en la investigación universitaria. Una visión desde Latinoamérica. Volumen I* (pp. 3–31). Fondo Editorial Universitario Servando Garcés de la Universidad Politécnica Territorial de Falcón Alonso Gamero / Alianza de Investigadores Internacionales S.A.S.
- Colica, E., D'Amico, S., Iannucci, R., Martino, S., Gauci, A., Galone, L., Galea, P., & Paciello, A. (2021). Using unmanned aerial vehicle photogrammetry for digital geological surveys: case study of Selmun promontory, northern of Malta. *Environmental Earth Sciences*, 80(17), 551. <https://doi.org/10.1007/s12665-021-09846-6>
- De Reu, J., De Smedt, P., Herremans, D., Van Meirvenne, M., Laloo, P., & De Clercq, W. (2014). On introducing an image-based 3D reconstruction method in archaeological excavation practice. *Journal of Archaeological Science*, 41, 251–262. <https://doi.org/10.1016/j.jas.2013.08.020>

- Dellepiane, M., Dell'Unto, N., Callieri, M., Lindgren, S., & Scopigno, R. (2013). Archeological excavation monitoring using dense stereo matching techniques. *Journal of Cultural Heritage*, 14(3), 201–210. <https://doi.org/10.1016/j.culher.2012.01.011>
- Fernández, J., González, D., Rodríguez, P., & Mancera, J. (2015). Image-Based Modelling from Unmanned Aerial Vehicle (UAV) Photogrammetry: An Effective, Low-Cost Tool for Archaeological Applications. *Archaeometry*, 57(1), 128–145. <https://doi.org/10.1111/arc.12078>
- Fraser, B., & Congalton, R. (2018). Issues in Unmanned Aerial Systems (UAS) Data Collection of Complex Forest Environments. *Remote Sensing*, 10(6), 908. <https://doi.org/10.3390/rs10060908>
- Gonçalves, J. A., & Henriques, R. (2015). UAV photogrammetry for topographic monitoring of coastal areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 104, 101–111. <https://doi.org/10.1016/j.isprsjprs.2015.02.009>
- Hill, A., Rowan, Y., & Kersel, M. (2014). Mapping with Aerial Photographs: Recording the Past, the Present, and the Invisible at Marj Rabba, Israel. *Near Eastern Archaeology*, 77(3), 182–186. <https://doi.org/10.5615/nea-reastarch.77.3.0182>
- INEGI. (2000). *Sistema Nacional de Fotografía aérea (SINFA)* (Issue 1). Instituto Nacional de Estadística, Geografía e Informática.
- James, M., Antoniazza, G., Robson, S., & Lane, S. (2020). Mitigating systematic error in topographic models for geomorphic change detection: accuracy, precision and considerations beyond off nadir imagery. *Earth Surface Processes and Landforms*, 45(10), 2251–2271. <https://doi.org/10.1002/esp.4878>
- Jones, C. A., & Church, E. (2020). Photogrammetry is for everyone: Structure-from-motion software user experiences in archaeology. *Journal of Archaeological Science: Reports*, 30, 102261. <https://doi.org/10.1016/j.jasrep.2020.102261>
- Koutsoudis, A., Vidmar, B., Ioannakis, G., Arnaoutoglou, F., Pavlidis, G., & Chamzas, C. (2014). Multi-image 3D reconstruction data evaluation. *Journal of Cultural Heritage*, 15(1), 73–79. <https://doi.org/10.1016/j.culher.2012.12.003>
- Kyriakaki, G., Doulamis, A., Doulamis, N., Ioannides, M., Makantasis, K., Protopapadakis, E., Hadjiprocopis, A., Wenzel, K., Fritsch, D., Klein, M., & Weinlinger, G. (2014). 4D Reconstruction of Tangible Cultural Heritage Objects from Web-Retrieved Images. *International Journal of Heritage in the Digital Era*, 3(2), 431–451. <https://doi.org/10.1260/2047-4970.3.2.431>
- Lerma, J. (2002). *Fotogrametría moderna: analítica y digital*. Universidad Politécnica de Valencia.

- Magnani, M., Douglass, M., Schroder, W., Reeves, J., & Braun, D. R. (2020). The Digital Revolution to Come: Photogrammetry in Archaeological Practice. *American Antiquity*, 85(4), 737–760. <https://doi.org/10.1017/aaq.2020.59>
- Olaya, V. (2020). *Sistemas de Información Geográfica*. <https://pcsitna.navarra.es/archivo/Documents/Sistemas.de.Informacion.Geografica.pdf>
- Parra, R., Díaz, C., Medina, S., Torres, J., & Felix, R. (2018). Monitoreo en la dinámica de explotación de bancos de materiales pétreos a partir de levantamientos aerofotográficos tripulados. In J. Rodríguez, H. Piña, & E. Lara (Eds.), *Estudios de Desarrollo Sustentable* (pp. 39–72). Universidad Autónoma Indígena de México.
- Parra, R., Pérez, M., Díaz, C., & Rodríguez, M. (2022). Generación de ortofotografías de alta resolución en la ciudad de Guasave, sinaloa, a partir de vuelos tripulados y el usos de software automatizado para su procesamiento. In H. Ayala, A. Ramírez, & Y. Martínez (Eds.), *Tendencias actuales en las Ciencias de la Computación* (pp. 205–228). Universidad Autónoma Indígena de México/Astra Editorial.
- Parra, R., Pérez, M., Rodríguez, J., Piña, H., & Castro, C. (2021). Implementación de fotografías aéreas y sistemas de información geográfica para la determinación de áreas y volúmenes de corteza vegetal. In Y. Chirinos, A. Ramírez, R. Godínez, N. Barbera, & D. Rojas (Eds.), *Tendencias en la investigación universitaria. Una visión desde Latinoamérica. Volumen XIII* (pp. 66–85). Fondo Editorial Universitario Servando Garcés de la Universidad Politécnica Territorial de Falcón Alonso Gamero / Alianza de Investigadores Internacionales S.A.S. <https://doi.org/10.47212/tendencias2020vol.xiii.5>
- Peinado, Z., Fernández, A., & Hernández, L. (2014). Combinación de fotogrametría terrestre y aérea de bajo coste: el levantamiento tridimensional de la iglesia de San Miguel de Ágreda (Soria). *Virtual Archaeology Review*, 5(10), 51. <https://doi.org/10.4995/var.2014.4210>
- Pérez, M., Parra, R., Macías, R., Díaz, C., & Félix, R. (2021). Análisis geoespacial para la delimitación de la zona federal marítimo terrestre en áreas naturales protegidas, caso playa El Maviri, Sinaloa, México. *Ra Ximhai*, 405–424. <https://doi.org/10.35197/rx.17.03.2021.17.mp>
- Przybilla, H.-J., Lindstaedt, M., & Kersten, T. (2019). INVESTIGATIONS INTO THE QUALITY OF IMAGE-BASED POINT CLOUDS FROM UAV IMAGERY. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W13(June), 539–545. <https://doi.org/10.5194/isprs-archives-XLII-2-W13-539-2019>
- Remondino, F., & Fraser, C. (2006). Digital Camera Calibration Methods: Considerations and Assumptions. In *International Archives of Photogram-*

- metry* (Vol. 36, Issue 5, pp. 266–272). <http://www.mendeley.com/research/digital-camera-calibration-methods-12/>
- Santamaría, J., & Sanz, T. (2011). Fundamentos De Fotogrametría. In *Material didactico* - Ingenierías 16. Universidad de la Rioja.
- Schenk, T. (2002). Fotogrametría digital. Instituto Cartográfico de Catalunya.
- Verhoeven, G. (2011). Taking computer vision aloft – archaeological three-dimensional reconstructions from aerial photographs with photoscan. *Archaeological Prospection*, 18(1), 67–73. <https://doi.org/10.1002/arp.399>
- Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). ‘Structure-from-Motion’ photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179, 300–314. <https://doi.org/10.1016/j.geomorph.2012.08.021>

## **Síntesis curriculares**

*Dr. Adalid Graciano Obeso*

Profesor investigador de tiempo completo titular B del Tecnológico Nacional de México, Campus Guasave, doctor en Ciencias en Estudios para la Sostenibilidad y Medio Ambiente por la Universidad Autónoma Indígena de México. Forma parte del Cuerpo Académico en Formación ITESGUA-CA-1 del Tecnológico Nacional de México, miembro del Sistema Sinaloense de Investigadores y Tecnólogos del estado de Sinaloa, pertenece a la Red Estatal de Divulgadores de la Ciencia y la Tecnología en Sinaloa, así como a la Red de Promotores ODS y a la Red Iberoamericana de Academias de Investigación A.C. Sus líneas de investigación son Tecnologías Emergentes e Innovación Agroalimentaria y Sistemas de Producción Agrícola Sustentable. Contacto: 6873668606, [adalid.go@guasave.tecnm.mx](mailto:adalid.go@guasave.tecnm.mx)

*Dr. Gilberto Bojórquez Delgado*

Profesor investigador de tiempo completo Titular B del Tecnológico Nacional de México, Campus Guasave, ingeniero Electrónico del Instituto Tecnológico de Los Mochis, obtuvo su Maestría en Ciencias en Ingeniería Electrónica por el Instituto Tecnológico de Chihuahua y su doctorado en Ciencias por la Universidad Autónoma Indígena de México. Como miembro del Sistema Nacional de Investigadores de Conacyt Nivel II, se especializa en tecnologías avanzadas para sistemas agroindustriales, con énfasis en IoT, Inteligencia Artificial y Big Data. Su trabajo se enfoca en promover la autosuficiencia tecnológica y el desarrollo regional sostenible, con numerosas publicaciones y proyectos de investigación aplicada. Contacto: 6871026729, [gilberto.bd@guasave.tecnm.mx](mailto:gilberto.bd@guasave.tecnm.mx)

*Dr. Lennin Enrique Amador Castro*

Doctor en Ciencias en Estudios para la Sostenibilidad y Medio Ambiente por la Universidad Autónoma Indígena de México. Maestro en Ciencias en Electrónica con orientación en Telecomunicaciones por el Centro de Investigación Científica y de Estudios Superior de Ensenada e ingeniero en Electrónica por el Tecnológico Nacional de México/ IT de Los Mochis. Investigador integrante del Sistema Sinaloense de Investigadores y Tecnólogos (SSIT). Actualmente

pertenece al Sistema Nacional de Investigadores e Investigadoras (SNII) como candidato a investigador. Correo electrónico: [lennin.ac@mochis.tecnm.mx](mailto:lennin.ac@mochis.tecnm.mx)  
ORCID: <https://orcid.org/0000-0001-6773-711X>

*Dr. Jesús Ramón Rodríguez Apodaca*

Doctor en Ciencias por la Facultad de Ciencias Químicas e Ingeniería de la Universidad Autónoma de Baja California, Campus Tijuana. Maestro en Ciencias en Química por el Centro de Graduados e Investigación en Química del Instituto Tecnológico de Tijuana, Campus Otay e ingeniero Químico por el Instituto Tecnológico de Los Mochis. Investigador integrante del Sistema Sinaloense de Investigadores y Tecnólogos (SSIT). Actualmente Profesor-Investigador de la Universidad Autónoma Indígena de México (UAIM). Cuenta con Reconocimiento a Perfil Deseable del Programa para el Desarrollo Profesional Docente (PRODEP). Integrante del Cuerpo Académico Biodiversidad y Estrategias Comunitarias de Desarrollo Sostenible UAIM-CA-13. Miembro de la Red Temática Nacional Patrimonio Biocultural del CONAHCyT e integrante del Sistema Nacional de Investigadores e Investigadoras Nivel I. Correo electrónico: [jramon@uaim.edu.mx](mailto:jramon@uaim.edu.mx) ORCID: <https://orcid.org/0000-0003-2679-8594>

*Esther Graciela Lizárraga Mata*

Maestra en Administración de Negocios de la Universidad TecMilenio y licenciada en Contaduría por el Tecnológico Nacional de México/ IT de Los Mochis. Docente del departamento de Ciencias Económico Administrativas de la Universidad Autónoma de Occidente. Correo electrónico: [esther.lizarraga@uadeo.mx](mailto:esther.lizarraga@uadeo.mx) ORCID: <https://orcid.org/0009-0008-3673-781X>

*Dr. Juan Héctor Alzate Espinoza*

Profesor investigador de tiempo completo del Tecnológico Nacional de México, Campus Guasave, doctor en Administración por el Centro Universitario de Ciencias e Investigación Líder del Cuerpo Académico en Formación ITESGUA-CA-4 del Tecnológico Nacional de México, miembro del Sistema Sinaloense de Investigadores y Tecnólogos del estado de Sinaloa, pertenece a la Sociedad Mexicana de Ingeniería Mecánica, así como al Colegio Mexicano de Ingenieros en Gestión Empresarial. Sus líneas de investigación son Desa-

rollo Regional Innovación tecnológica y Calidad, Y Diseño y optimización de productos y procesos. Contacto: 6681644473, [juan.ae@guasave.tecnm.mx](mailto:juan.ae@guasave.tecnm.mx)

*Dra. Grace Erandy Báez Hernández*

Profesor investigador de tiempo completo titular B del Tecnológico Nacional de México, Campus Guasave, candidata a doctora en Ciencias en Desarrollo Sustentable de Recursos Naturales por la Universidad Autónoma Indígena de México. Líder del Cuerpo Académico en Formación ITESGUA-CA-1 del Tecnológico Nacional de México, miembro del Sistema Sinaloense de Investigadores y Tecnólogos del estado de Sinaloa, pertenece a la Red Estatal de Divulgadores de la Ciencia y la Tecnología en Sinaloa, así como la Red Iberoamericana de Academias de Investigación A.C., y presidente de la Red de Investigación de Ergonomía del Noroeste. Sus líneas de investigación son: Desarrollo Regional Innovación tecnológica y Calidad, Y Diseño y optimización de productos y procesos. Contacto: 6681420529 [grace.bh@guasave.tecnm.mx](mailto:grace.bh@guasave.tecnm.mx)

*Dr. Marco Arturo Arciniega Galaviz*

Doctor en Desarrollo Sustentable de Recursos Naturales. Profesor de carrera de tiempo completo del Departamento Académico de Ingeniería y Tecnología de la Universidad Autónoma de Occidente. Líneas de investigación: Prevención y control ambiental, Manejo sustentable de recursos biológicos. ORCID ID: <https://orcid.org/0000-0001-8532-7130>. Correo electrónico: [marco.arciniega@uadeo.mx](mailto:marco.arciniega@uadeo.mx), Celular: 6681463982

*M. C. Jesús Bojórquez Delgado*

Graduado en Sistemas Computacionales de la Universidad de Occidente, y de la maestría en Sistemas Computacionales y actualmente estudia un doctorado en la misma especialidad. Es docente en el Instituto Tecnológico Superior de Guasave. Además, es miembro del cuerpo académico Tecnologías estratégicas de vanguardia aplicadas para el desarrollo sustentable de sistemas agroindustriales de precisión. Sus áreas de interés se centran en sistemas inteligentes, contribuyendo al avance y la aplicación de estas tecnologías en el contexto agroindustrial. Contacto: Teléfono - 6871046387, Correo Electrónico – [jesus.bd@guasave.tecnm.mx](mailto:jesus.bd@guasave.tecnm.mx)

*Dr. Román Edén Parra Galaviz*

Obtuvo su doctorado en Ciencias en Desarrollo Sustentable de Recursos Naturales por la Universidad Autónoma Indígena de México (UAIM) en 2018 y licenciatura en Ingeniería Geodésica por la Universidad Autónoma de Sinaloa (UAS) en 2007. Profesor e investigador en la Facultad de Ingeniería Mochis de la UAS, miembro del Sistema Nacional de Investigadores Nivel C en el área IX (Interdisciplinaria), y del Sistema Sinaloense de Investigadores y Tecnólogos (SSIT) con distinción honorífica. Colaborador del Cuerpo Académico UAS-CA-274 y en proyectos de investigación PROFAPI. Ha impartido clases y dirigido tesis a nivel de licenciatura y posgrado, con publicaciones en revistas indexadas y capítulos de libros, participación en congresos nacionales e internacionales en áreas de interés; Geodesia, Fotogrametría, Cartografía, SIG y Desarrollo Sostenible. Correo electrónico: roman.parra@uas.edu.mx, celular: 668-1246479, ORCID: 0000-0003-2679-8594

*Manuel de Jesús Pérez Valdez*

Profesor e investigador de tiempo completo titular B, en la Universidad Autónoma de Sinaloa (UAS). Doctorado en Ciencias Técnicas con especialidad en Fotogrametría, otorgado en el año 2016, por la Universidad Estatal de Geodesia y Cartografía de Moscú, Rusia (MIIGAik). Licenciado en ingeniería geodésica por la Universidad Autónoma de Sinaloa en la Facultad de Ingeniería Mochis, desde el año 2003. Actividad docente como profesor de nivel licenciatura, maestría, en el área de la geodesia, fotogrametría, geomática y medioambiente. Participación como miembro del grupo de investigación: Cuerpo académico Geosistemas UAS-CA-274, en proyectos de investigación (PROFAPI), dentro las líneas de aplicación y Generación de Conocimiento: Geodesia y Geomática. Dirección de proyectos de tesis de nivel licenciatura en el área de Geodesia y la Fotogrametría. Perfil PROMEP, Secretaría de Educación Pública, Programa de Mejoramiento del Profesorado, 2022-2025. Miembro del Sistema Sinaloense de Investigadores y Tecnólogos (SSIT). e-mail: manuel.perez@uas.edu.mx, celular: 668-1202101, ORCID: 0000-0003-4919-489X

*Dr. Ciro de la C. Díaz Hernández*

Obtuvo su doctorado en Ciencias técnicas de Geodesia Aplicada en la Habana, Cuba y licenciatura en Ingeniería Geodésica por la Universidad Estatal de

Geodesia y Cartografía de Moscú, Rusia (MIGAik). Profesor e investigador de la Facultad de Ingeniería Mochis de la UAS, líder del Cuerpo Académico UAS-CA-274 y en proyectos de investigación PROFAPI. Ha impartido clases y dirigido tesis a nivel de licenciatura y maestría, con publicaciones en revistas indexadas y capítulos de libros, participación en congresos nacionales e internacionales en áreas de interés; Geodesia, Fotogrametría, Cartografía, SIG. e-mail: cirodiaz@uas.edu.mx, celular: 668-1622872, ORCID: 0000-0001-5402-1438

*Dra. Elvia Nereyda Rodríguez Saucedo*

Doctora en Ciencias en Desarrollo Sustentable de los Recursos Naturales por la Universidad Autónoma Indígena de México. elviaro@uaim.edu.mx

*Dra. Raquel Rodríguez Saucedo*

Doctora en Ciencias en Desarrollo Sustentable de los Recursos Naturales por la Universidad Autónoma Indígena de México. raquelrguez@uaim.edu.mx

*Dr. Adán Guillermo Ramírez García*

Doctor en Geografía por la Universidad Nacional Autónoma de México. Profesor Investigador Universidad Autónoma Chapingo. aguillermoramirezgarcia@gmail.com

*Dr. Jesús Manuel Díaz Gaxiola*

Profesor de tiempo completo titular C del Instituto Tecnológico de Los Mochis (TECNM) desde 1994, en la academia de Biología. Licenciatura en Biología con especialidad en Biología Acuática por el Instituto Tecnológico de Los Mochis. Doctor en Ciencias en Desarrollo Sustentable de Recursos Naturales por la UAIM. Jefe del laboratorio de Ecología y de proyectos de investigación del departamento de Ingeniería Química y Bioquímica. Área de investigación de recursos naturales en zonas marinas y costeras. Correo electrónico: jesus.dg@mochis.tecnm.mx teléfono móvil: 6681633171

*Jesús Octavio Escobedo Urquidez*

Egresado de la carrera de licenciatura en Biología del Instituto Tecnológico de Los Mochis (TECNM). Correo electrónico: [escobedo.octavio004@gmail.com](mailto:escobedo.octavio004@gmail.com). Teléfono móvil: 6684 649818

*Dra. Claudia Selene Castro Estrada*

Doctora en Ciencias en Desarrollo Sustentable de Recursos Naturales. Integrante del Sistema Sinaloense de Investigadores y Tecnólogos. Coordinadora General de Investigación y Posgrado en la Universidad Autónoma Indígena de México. Profesora-Investigadora de la Universidad Autónoma Indígena de México. Integrante del Sistema Nacional de Investigadores como Candidata. Teléfono: 6681690318. Correo electrónico: [draclaudiacastro@uaim.edu.mx](mailto:draclaudiacastro@uaim.edu.mx) ORCID: <https://orcid.org/0000-0003-4461-9633>

*Dra. Alma Lorena Quintero Romanillo*

Licenciada en Biología por el Instituto Tecnológico de Los Mochis. Maestra y doctora en Ciencias en Desarrollo Sustentable de Recursos Naturales por la Universidad Autónoma Indígena de México. Profesora adscrita a la carrera de Ingeniería Forestal en la Universidad Autónoma Indígena de México. Integrante del Cuerpo Académico Biodiversidad y Estrategias Comunitarias de Desarrollo Sostenible UAIM-CA-13. Tel: 6681068621. Correo electrónico: [lorenaquintero@uais.edu.mx](mailto:lorenaquintero@uais.edu.mx) ORCID: <https://orcid.org/0000-0002-4738-3024>

*Métodos cuantitativos aplicados a recursos naturales, medioambiente y la sostenibilidad.*

Se terminó de imprimir en marzo de 2024

en los talleres de **Astra Ediciones**

Av. Acueducto No. 829

Colonia Santa Margarita, C. P. 45140

Zapopan, Jalisco, México.

33 38 34 82 36

E-mail: [edicion@astraeditorial.com.mx](mailto:edicion@astraeditorial.com.mx)

[www.astraeditorialshop.com](http://www.astraeditorialshop.com)

Impresión digital con interiores en papel bond 75g

El tiraje consta de 300 ejemplares.

**En este libro se exponen de manera magistral y a través de escenarios reales, los beneficios del análisis cuantitativo, ya sea en el mar, la tierra o el aire, y de los recursos naturales que estos poseen y sus servicios ambientales que brindan, para realizar aprovechamientos sostenibles y generar conocimientos actuales y de un valor inestimable dado las afectaciones que hemos realizado como humanidad.**

***Dr. Jesús Manuel Díaz Gaxiola***

ISBN: 978-84-10215-06-1

