

# MATEMÁTICAS, ESTADÍSTICA Y MEDIO AMBIENTE



## EDITORES

Hortensia Josefina Reyes Cervantes

Manuel Ibarra Contreras

Nahun Israel Loya Monares

# 2013



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA  
FACULTAD DE CIENCIAS FÍSICO - MATEMÁTICAS

Matemáticas Estadística y Medio Ambiente

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA  
Dirección de Fomento Editorial  
2013

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

Mtro. José Alfonso Esparza Ortíz

Rector

Dr. J. Ignacio Morales Hernández

Secretario General

Fernando Santiesteban Llaguno

Vicerrector de Extensión y Difusión de la Cultura

José Ramón Enrique Arrazola Ramírez

Director de la Facultad de Ciencias Físico Matemáticas

Carlos Contreras Cruz

Director Editorial

2013 Primera Edición

ISBN: 978-607-487-571-3

©Benemérita Universidad Autónoma de Puebla

Dirección de Fomento Editorial

2 Norte 1404, CP 7200,

Puebla, Pue.

Teléfono y fax 01-222-2468559

Impreso y hecho en México

Printed and made in México

## Prólogo

La contaminación ha alterado el estado de equilibrio de los ecosistemas entre otras cosas, por no tener cuidado al usar una tecnología responsable y cuidadosa del medio ambiente. Actualmente el problema de la contaminación es más agudo que en épocas pasadas, porque gran parte de los desechos tienen origen inorgánico ya que la alteración de las cantidades de ciertos elementos y compuestos químicos tiene una repercusión en los procesos naturales. Tratar de encontrar una salida rápida y exitosa es difícil debido al incremento demográfico, el vertiginoso desarrollo industrial y por la falta de conciencia o interés.

En este contexto surge el Segundo Encuentro Internacional de Medio Ambiente (2EIMA) que es posible gracias a la participación entusiasta de investigadores, docentes, alumnos y público en general que se interesa por la modelación matemática de problemas ambientales y a su vez, que tiene la entera disposición de indagar nuevas tendencias, perspectivas y enfoques que resulten de este encuentro multidisciplinar.

El objetivo de este 2EIMA es transitar por la problemática del medio ambiente que puede ser vista desde distintos y variados enfoques. Intentar amalgamar la visión particular de estadísticos, matemáticos, químicos, biólogos, ingenieros y físicos, es algo importante porque cada investigador trata de resolver y hacer contribuciones con técnicas diferentes que pueden ayudar a tener respuestas concretas.

Este libro consta de doce capítulos, cada uno de ellos desarrollados por investigadores de amplio reconocimiento nacional y que pertenecen a instituciones como: Centro de Ciencias de la Atmósfera de la Universidad Autónoma de México, Instituto Nacional de Salud Pública, Universidad Autónoma de Guerrero, Centro de Investigación Científica de Yucatán, Instituto de Ecología, Colegio de Postgraduados, Universidad Veracruzana, Universidad de la Habana y algunos Institutos y Facultades de la Benemérita Universidad Autónoma de Puebla. Los temas que se presentan son muy diversos, pero se pueden clasificar en aquellos que tratan de Modelación en diversos ambientes de seres vivos y los que hacen alguna Evaluación o Cuantificación de contaminantes en zonas específicas.

Agradecemos a los investigadores José Agustín García Reynoso, Mario Cantú Sifuentes y David Martínez Moreno, el habernos apoyado con sus revisiones y sugerencias en los extensos elegidos para integrarse a este libro. Además, agradecemos a Cintli

Estefanía Ibarra Reyes el diseño de la carátula de este material.

Esperamos que esta publicación sea de interés para la comunidad universitaria y público en general.

### **Los Editores**

**Hortensia J. Reyes Cervantes, Manuel Ibarra Contreras y  
Nahun Loya Monares**

# Índice

1	Análisis de varianza de la Contaminación en hojas de <i>Fraxinus uhdei</i> en la Zona Metropolitana del Valle de México.	1
	SÁNCHEZ LANDERO LUZ AMELIA	
2	Cuantificación de los metales pesados Plomo y Cadmio en pan de fabricación industrial.	10
	LÓPEZ OLIVARES GUADALUPE	
3	Diagnóstico de cáncer de mama mediante una red neuronal usando meta-heurísticas.	17
	RIVERA MARTÍNEZ MARCELA	
4	Diagnóstico de los residuos solidos urbanos generados en la colonia Atlahuayan, Yautepec Morelos 2012	27
	DOMÍNGUEZ ORTÍZ MARCO G	
5	El modelo de regresión logística: una aplicación al problema de hipertensión arterial, sobrepeso u obesidad de personas que acuden a los Centros de Salud de Chilpancingo. Gro.	37
	MORENO SANDOVAL MA. ISABEL	
6	Evaluación del riesgo por contaminantes tóxicos en la Ciudad de México	47
	GARCÍA REYNOSO JOSÉ AGUSTÍN	
7	<i>Inga jinicuil</i> (Jinicuil) Como árbol frutal nativo: valor comercial, calidad de frutos y valor cultural como determinantes para permanecer o ser eliminados.	54
	PULIDO-SALAS TERESA PATRICIA	
8	Modelación del Ozono en la estación Agua Santa de la Ciudad de Puebla, usando L-moments	66
	CRUZ JUÁREZ JUAN ANTONIO	
9	Predicción de la precipitación media para el estado de Puebla.	76
	MARCIAL CASTILLO LUIS RENÉ	
10	Recuperación de aceite quemado de motor aplicando dos métodos: físico-químico y biológico como medio preventivo de la contaminación ambiental	86
	POZOS GUTIÉRREZ ARTURO	
11	Teledetección del carbono orgánico en suelos de Tetelilla, Teziutlán, Puebla.	98
	TORRES T. EDGARDO	

12 Una aplicación del índice Dcova a la diversidad de escarabajos del bosque mesófilo de montaña	108
ROSAS NAVA RUBÍ	

# Análisis de varianza de la Contaminación en hojas de *Fraxinus uhdei* en la Zona Metropolitana del Valle de México.

Sánchez Landero Luz Amelia<sup>1</sup>      Hernández González Sergio<sup>2</sup>

de la Isla de Bauer María de Lourdes<sup>1</sup>      López López Miguel Ángel<sup>1</sup>

Gómez Guerrero Armando <sup>1</sup>      González Ávalos Eugenio<sup>3</sup>

Olguín Palacios Eugenia J. <sup>4</sup> Línea de investigación: Estadística

<sup>1</sup>Colegio de Postgraduados

*Carretera México-Texcoco Km. 36.5*

*Montecillo, Texcoco, Estado de México, 56230, México*

*amelia@colpos.mx* <sup>2</sup>Universidad Veracruzana, Fac. de Estadística e Informática

*Av. Xalapa Esq. Manuel Ávila Camacho s/n*

*Xalapa, Veracruz, 91020, México*

<sup>3</sup>Instituto Mexicano del Petróleo

*Eje Central Norte Lázaro Cárdenas No. 152, Col. San Bartolo Atpehuacan*

*México, D.F., 07730, México*

<sup>4</sup>Instituto de Ecología

*Carretera antigua a Coatepec 351*

*El Haya, Xalapa, Veracruz, 91070, México*

## Resumen

Se realizó un biomonitoreo de partículas PM 10 de plomo (Pb), zinc (Zn), cobre (Cu), cobalto (Co), cromo (Cr), níquel (Ni) y cadmio (Cd); en tres bosques urbanos de la ciudad de México, los cuales fueron seleccionados en función de los vientos que predominan la mayor parte del año y que van de N-SW. La especie que se seleccionó para biomonitoreo fue *Fraxinus uhdei* mejor conocida como Fresno, la cual predomina en la mayor parte de la reforestación urbana en la ciudad. Para realizar el estudio se colectaron hojas de 45 árboles ubicados en: parque de Naucalli, Bosque de Chapultepec, y Bosque de Tlalpan. Las hojas se cortaron a una altura de 8 a 10 m una vez durante otoño e invierno del 2011. Para la extracción de las PM 10, se utilizó un lavado con ácido etilendiaminotetraacético (EDTA) tipo disódico y posteriormente se sometieron las muestras a un proceso de digestión con ácido nítrico (HNO<sub>3</sub>). Todas la lecturas se realizaron con ayuda de un equipo ICP 3000 (Inductively Coupled Plasma). Para el análisis estadístico se aplicó una prueba de *Kruskal Wallis* y se encontró que existen diferencias significativas entre las concentraciones de material particulado que se deposita sobre las hojas por parque y ciclo estacional, además en la separación de medias se observó que el bosque de Chapultepec es diferente en sus

contenidos de metales de los otros dos bosques urbanos en los dos ciclos estacionales.

**Palabras clave:** Contaminación de hojas de *Fraxinus uhdei*, Análisis de varianza, Biomonitorio.

## 1. INTRODUCCIÓN

La contaminación aérea es un serio problema en muchas áreas densamente pobladas e industrializadas en el mundo (Kambezidis et al., 1996). El tráfico que se genera por los vehículos automotores es la principal fuente de contaminación aérea contribuyendo con aproximadamente un 57-75 % del total de emisiones (WHO, 2006). Para el caso de la ciudad de México, la calidad del aire se considera como uno de los principales objetos de preocupación entre los distintos problemas ambientales a los que se enfrentan la ciudad, ya que se tienen datos de que es uno de los núcleos urbanos con mayor población a nivel mundial; y la elevada concentración de industrias, vehículos y el uso doméstico de la energía provocan el deterioro de éste (Vázquez-Cruz, 2009). Para tal caso se sabe que hoy en día bosques y árboles en ambientes urbanos pueden mejorar la calidad del aire (Pourkhabbaz et al., 2010); al actuar como captadores de gases y partículas mejorando con ello los problemas de contaminación (Beckett et al., 1998, 2000). Sin embargo, la capacidad de intercepción del polvo en las plantas, depende de la geometría de su superficie, y de las características externas de la hoja como pelos, cutícula, altura, así como de la copa de los árboles e incluso de la sensibilidad que la especie pueda presentar (Alcalá et al., 2008). El material particulado daña a las hojas de las plantas, afecta al crecimiento de éstas, limita su productividad, según la sensibilidad que tengan a los contaminantes (St. Clair et al., 2002); puede ocasionar clorosis, necrosis y epinastia (Katiyar y Dubey, 2000). Además de una disminución del contenido de clorofila y aumento en el contenido de ácido ascórbico, el cual es un antioxidante que neutraliza a los oxidantes naturales de las hojas (Alcalá et al., 2008).

El análisis químico foliar es usado como biomonitorio al funcionar como una especie de seguimiento y de diagnóstico en el sector forestal así como estudios del medio ambiente. Se ha utilizado para deficiencias de nutrientes y estimación de la toxicidad (Morrison., 1974, Van den Driessche 1974). En bosques urbanos sirve para monitorear la calidad del aire en las grandes ciudades (Pourkhabbaz et al., 2010). A través de dicho análisis se pueden conocer las concentraciones de partículas PM<sub>10</sub>, las cuales son llamadas así por su diámetro aerodinámico equivalente o menor a 10<sup>-6</sup>m generadas por la combustión de materiales carburantes fósiles o de procesos industriales en los que se trabaja con altas temperaturas que generan condensación de gases, humos y vapores (Pourkhabbaz et al., 2010). Estas partículas impactan en el medio ambiente, ya que actúan como catalizadores sobre superficies metálicas, favoreciendo su oxidación, se consideran perjudiciales para la salud y su efecto depende de la composición química que éstas puedan tener (Marcazzan et al., 2003).

El presente trabajo muestra los resultados preliminares de un estudio de biomonitorio de tres bosques urbanos de la Zona Metropolitana del Valle de México en donde se realizaron colectas de hojas en árboles de *Fraxinus uhdei* (Fresno) con la finalidad de determinar cualitativa y cuantitativamente el material particulado que se deposita sobre éstos árboles así como obtener información sobre la distribución y comportamiento de partículas (me-

tales) y determinar si existe una relación entre la cantidad de metales depositado sobre hojas que afecte el crecimiento de los árboles.

## 2. MATERIALES Y MÉTODOS

El estudio se realizó en tres bosques urbanos de la Zona Metropolitana del Valle de México, los cuales se seleccionaron en función de la dirección de los vientos que a traviesan la mayor parte del año a la ciudad de México, (N- SW). Dichos parques son: Parque Naucalli, el cual cuenta con una superficie de 42 ha y se localiza en la parte norte, el Bosque de Chapultepec con una superficie de 686.01 ha localizado al poniente de la Ciudad de México y finalmente el Bosque de Tlalpan ubicados en la parte sur y con una superficie de 252 ha (Figura 1).

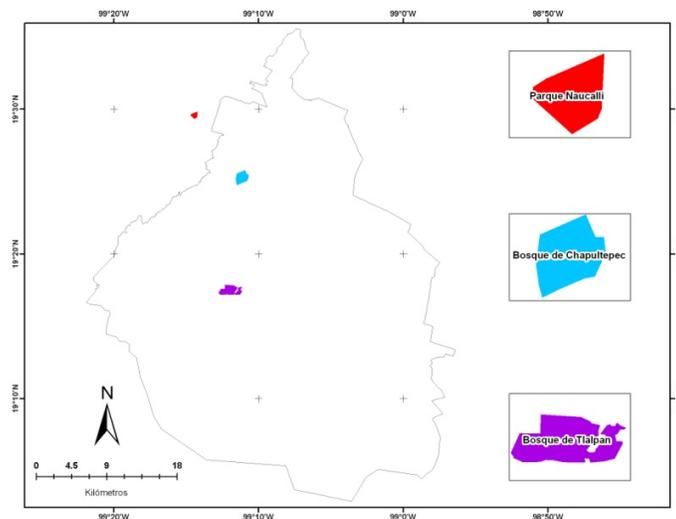


Figura 1: Localización de los bosques urbanos Naucalli, Chapultepec y Tlalpan.

De acuerdo a los criterio de Wells et al., 1985 se realizó la selección de árboles, considerando 15 por parque y delimitado a través de un transecto en donde se procuró que éste colindara con las principales avenidas y que se contara la especie *Fraxinus uhdei*. Las hojas se cortaron a una altura de 8-10 m aproximadamente con ayuda de una garrocha, realizando una colecta en los ciclos estacionales de otoño e invierno.

Una vez cortadas las hojas se depositadas en bolsas de plástico y fueron colocadas en una hielera para su transporte al laboratorio en donde se procedió a separar tallos de los foliolos y se pesaron 20g de éstas muestras foliares en fresco para someterse a un proceso de lavado con una solución EDTA 0.01 M, en donde los 20g de muestra fueron depositaron en recipientes de plásticos con 200mL de esta solución y se agitaron en una placa de vaivén por 60 min; al finalizar este proceso el extracto obtenido de la agitación fue colocado en recipientes de 100mL y se agregó 0.5mL de ácido nítrico para evitar precipitación de metales. La hojas fueron sometidas nuevamente a un segundo proceso

de agitación con 200mL de agua desionizada por otros 60min y al finalizar este segundo lavado el extracto obtenido se almacena igual que en el caso del primer lavado (Figura 2).

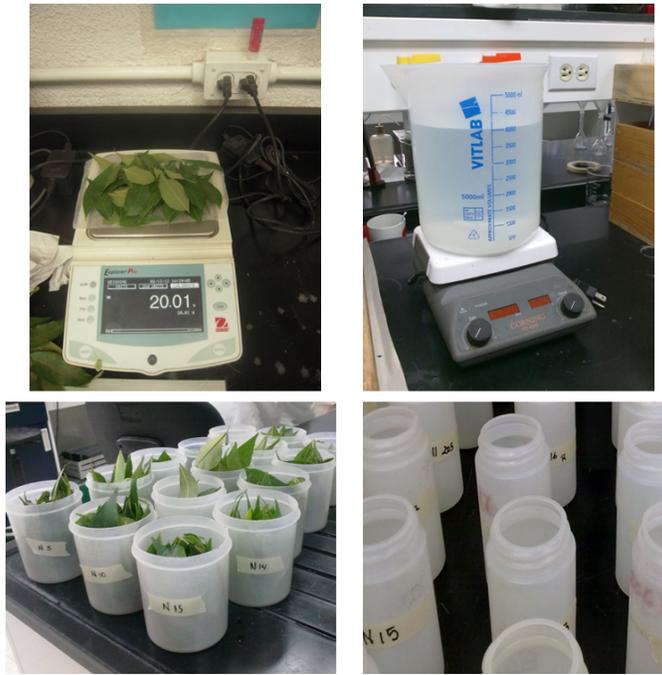


Figura 2: Procesamiento de muestras en laboratorio.

Finalizado el proceso de lavado, las muestras fueron sometidas al secado en estufa a 70° C durante 3 días, o hasta que tomaran peso constante, para posteriormente pulverizarse en un molino eléctrico marca GE; y después con ayuda de una balanza analítica se pesa 0.25 g de muestra pulverizada y someterlas a un proceso de digestión con 6mL de (HNO<sub>3</sub>) en un digestor tipo micro ondas por espacio de una hora, una vez que fueron retiradas del digestor se aforó a 25mL y llevo a lectura en un Espectro de Masa de Plasma Acoplado (ICP) al igual que los extractos obtenidos de los lavados, con la finalidad de verificar la presencia o ausencia de los metales pesados de interés y la cantidad de ellos (Pb, Co, Cr, Ni, Cu, Zn, Ni). Con los datos obtenidos se aplicó una prueba de *Kruskal Wallis* para ver si existen diferencias significativas entre contenidos y además se hizo una separación de medias.

### 3. RESULTADOS

Los datos fueron analizados con el paquete estadístico MINITAB® Release 14.1. Y se encontró que existen diferencias significativas entre los concentraciones de metales por bosques urbanos y por ciclo estacional (cuadro 1 y 2). Se observó que las concentraciones son significativas cuando el valor de  $p > 0.05$ ; por lo que para el periodo de verano los contenidos de zinc, cromo y cobre que son los elementos que se manifestaron en las hojas

<b>Periodo: Verano</b>		
Elemento	Valor p	
Plomo(Pb)	N/D	Los contenidos eran mínimos y no se alcanzaron a detectar con el ICP
Zinc(Zn)	0.029	Diferencias significativas
Cobalto(Co)	0.303	No existen diferencias significativas
Cromo(Cr)	0.000	Diferencias significativas
Cadmio(Cd)	0.269	No existen diferencias significativas
Cobre(Cu)	0.006	Diferencias significativas
Níquel(Ni)	N/D	Los contenidos eran mínimos y no se alcanzaron a detectar con el ICP

Cuadro 1: Resultados de prueba de *Kruskal Wallis* verano 2011

<b>Periodo: Invierno</b>		
Elemento	Valor p	
Plomo (Pb)	0,027	Diferencias significativas
Zinc (Zn)	0,320	No existen diferencias significativas
Cobalto (Co)	0,278	No existen diferencias significativas
Cromo (Cr)	0,322	Los contenidos eran mínimos y no alcanzaron a detectar con el ICP
Cadmio (Cd)	0,250	No existen diferencias significativas
Cobre (Cu)	0,277	No existen diferencias significativas
Níquel (Ni)	N/D	Los contenidos eran mínimos y no se alcanzaron a detectar con el ICP

Cuadro 2: Resultados de prueba de *Kruskal Wallis* invierno 2011

de los árboles, varían de un bosque urbano a otro, mientras que para el periodo de invierno las concentraciones de zinc y cobre son iguales en los tres bosques urbanos y el plomo diferente en sus concentraciones.

Además los resultados indican que el metal que se concentra en mayor cantidad para los tres bosques urbanos es el zinc, lo cual ocurrió para ambos periodos, seguido por el cobre (Figura 3).

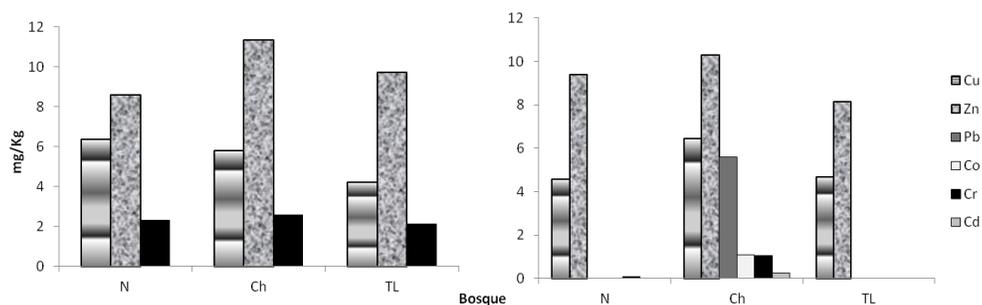


Figura 3: Concentración de metales por ciclo estacional. a) Periodo otoño, b) Periodo invierno.

Otro aspecto que resalta son los contenidos de plomo, ya que para el periodo de verano las concentraciones eran mínimas o casi nulas y no se pudieron detectar con el equipo, sin embargo para el invierno se logra apreciar que existe depósito de este metal sobre las hojas. En general se puede observar que en el periodo de invierno es cuando se logró encontrar a la mayor cantidad de metales aunque las concentraciones son parecidas a las de periodo verano.

Además de lo anterior se realizó una separación de medias por elemento y en la mayoría de los casos se observó que el bosque de Chapultepec fue el que obtuvo las mayores concentraciones de metales depositados tanto en el ciclo de otoño como en el ciclo de invierno (Figuras 4-9).

### Cobre (Cu)

#### Otoño

Bosque	N<=	N>	Median	Q3-Q1	Individual 95.0% CIs
N	4	11	6.36	4.43	(-----*-----)
Ch	6	9	5.79	4.93	(-----*-----)
TL	13	2	4.20	1.49	(-----*-----)

4.5                      6.0                      7.5

#### Invierno

Bosque	N<=	N>	Median	Q3-Q1	Individual 95.0% CIs
N	8	7	4.58	2.97	(-----*-----)
Ch	5	10	6.46	4.39	(-----*-----)
TL	10	5	4.67	2.35	(-----*-----)

4.5                      6.0                      7.5

Figura 4: Separación de medias para cobre



## Cromo (Cr)

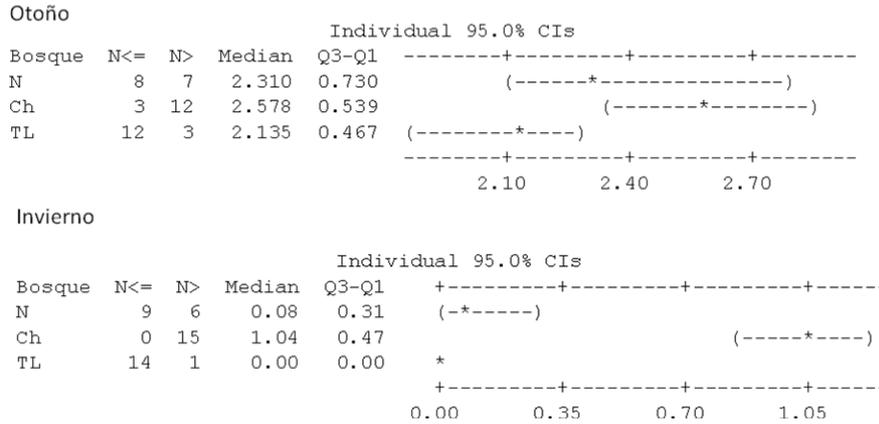


Figura 8: Separación de medias para cromo.

## Cadmio (Cd)

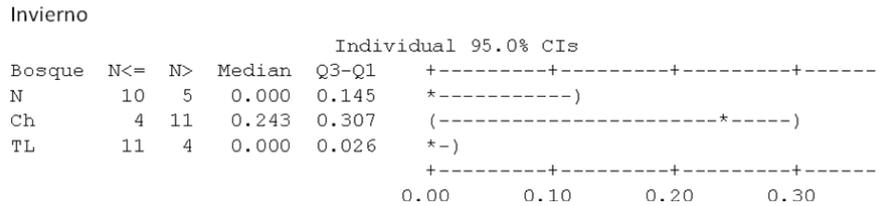


Figura 9: Separación de medias para cadmio

## 4. CONCLUSIONES

Los resultados indican que las condiciones urbanas afectan la concentración de metales que se depositan sobre las hojas de árboles urbanos. Los elementos que presentan una mayor variabilidad en cuanto a sus contenidos son cobalto y cadmio en todos los parques. En cuanto a la concentración de los elementos, se observó que existen diferencias significativas por parque lo que indica que no son iguales las concentraciones de estos depositados sobre las hojas de los árboles, sin embargo en la separación de medias en la mayoría de los casos el bosque de Chapultepec es diferente de Naucalli y Tlalpan.

## Referencias

- [1] Alcalá, J.; Sosa, M.; Moreno, M.; Quintana, C.; Quintana, G.; Miranda, S. y Rubio A. (2008) “Metales pesados en vegetación arbórea como indicador de la calidad am-

- biental urbana: ciudad de Chihuahua, México.”, *Latin American Journal of Natural Resources* **17**(1): 39-54.
- [2] Beckett, K. P., P. H. Freer-Smith, y G. Taylor. (1998) “ Urban woodlands: Their role in reducing the effects of particulate pollution”, *Environ. Pollut.* **99**: 347-360.
- [3] Beckett, K. P., P. H. Freer-Smith, y G. Taylor (2000) “ Particulate pollution capture by urban trees: Effect of species and windspeed”, *Glob. Change Biol.* **6**: 995-1003.
- [4] Kambezidis, H.D, A.D. Adamopoulos., y C. Gueymard (1996) “ Total NO<sub>2</sub> column amount over Athens”, *Atmospheric Research* **57**: 1-8. Greece in 1996 - 1997.
- [5] Katiyar, V. y Dubey, P.S. (2000) “ Growthbehaviour of two cultivors of Maize in responseto SO<sub>2</sub> and Nox.”, *Environmental Biology* **21**(1): 317-323.
- [6] Marcazzan, G.M., M.Ceriani, G.Valli, y R.Vecchi. (2003) “ Source apportionment of PM 10 y PM 2.5 in Milan (Italy) using receptor modeling”, *The Science of Total Environment* **317**(1): 137-147.
- [7] Morrison, I.K. (1974) “ Mineral nutrition of conifers with special reference to nutrient status interpretation: a review of literature. Department of the Environment”, *Canadian Forestry Service* **1**: 73.
- [8] St. Clair, S. B., L. L. St. Clair, N. F. Mangelson, y D. J. Weber. (2002) “ Influence of growth form on the accumulation of airborne copper by lichens, Atmos”, *Environ* **36**: 5637 - 5644..
- [9] Pourkhabbaz, A., NayerahRastin, Olbrich, A., Langenfeld-Heysen, R., Polle, A. (2010) “ Influence of Environmental Pollution on Leaf Properties of Urban Plane Tree, *Platanus orientalis* ”, *L. Bulletin of Environmental Contamination and Toxicology* **85**(3): 251-255.
- [10] Van DEN Driessche, R. (1974) “ Prediction of mineral nutrient status of trees by foliar analysis”, *Experimental Botany* **40**: 347-394.
- [11] Vázquez-Cruz C. (2009) *Estudio del comportamiento dinámico de la contaminación por material particulado PM10 en la zona metropolitana del Valle de México*. Tesis de Maestría, Universidad Autónoma de México., México.
- [12] Wells, C. y A. Lee. (1985) “ A Loblolly Pine Management Guide Whmand Whem to Apply Fertilizer. Department of Agriculture Forest Service ”, *General Technical Report* **36**: 32.
- [13] WHO (2006) “ Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulphur dioxide ”, *Global update 2005. WHO Regional Office for Europe, Copenhagen, Denmark*

# Cuantificación de los metales pesados Plomo y Cadmio en pan de fabricación industrial

López Olivares Guadalupe<sup>1</sup> Castro Caballero Leopoldo<sup>1</sup>  
Cruz Ortega Norma Judith<sup>1</sup>  
Meléndez Balbuena Lidia<sup>1</sup>  
Castro Lino Alejandra<sup>1</sup>  
Soto López Ismael<sup>1</sup>

Línea de investigación: Ciencias Químicas.

<sup>1</sup>Benemérita Universidad Autónoma de Puebla.

*14 Sur y Av. San Claudio, Col. San Manuel, C.U,  
Puebla, Puebla, CP 74570, México.  
e-mail:guadamax@yahoo.com.mx*

## Resumen

El presente trabajo reúne los datos de la cuantificación por Absorción Atómica de los metales pesados plomo y cadmio presentes en muestras de pan de fabricación industrial distribuido en tiendas, establecimientos comerciales y de autoservicio de la ciudad de Puebla. Tanto el plomo como el cadmio son metales tóxicos que no deben estar presentes en los alimentos o estar dentro de las especificaciones de la NOM-247-SSA1-2008, se realizó un análisis estadístico de las concentraciones encontradas que mucho tienen que ver con el proceso de fabricación y con la procedencia de la materia prima en la que muy probablemente se encuentren también estos metales debido a la contaminación ambiental.

**Palabras clave:** Metales pesados tóxicos, fabricación industrial, Ciudad de Puebla.

## 1. Introducción

Los metales pesados son un grupo de elementos químicos que presentan una densidad relativamente alta y cierta toxicidad para el ser humano. Muchos de los metales que tienen una densidad alta no son especialmente tóxicos y algunos son elementos esenciales en el ser humano, independientemente de que a determinadas concentraciones puedan ser tóxicos en alguna de sus formas. Sin embargo, hay una serie de elementos que en alguna de sus formas pueden representar un serio problema medioambiental y es común referirse a ellos con el término genérico de "metales pesados"(1).

La peligrosidad de los metales pesados es mayor al no ser química ni biológicamente degradables. Una vez emitidos, principalmente debido a la actividad industrial y minera,

pueden permanecer en el ambiente durante cientos de años, contaminando el suelo y acumulándose en las plantas y los tejidos orgánicos. Además, su concentración en los seres vivos aumenta a lo largo de la cadena alimentaria (2).

Los metales pesados tóxicos más conocidos son el mercurio, el plomo y el cadmio. El plomo es un contaminante medioambiental natural, muy usado en el pasado en tuberías de agua, pinturas y gasolina. Los alimentos y el agua de bebida, junto con el aire, son la principal fuente de exposición humana al plomo. El plomo se acumula en el cuerpo y afecta principalmente al desarrollo del sistema nervioso central en los niños. En adultos se ha asociado su presencia con la aparición de hipertensión y enfermedades renales crónicas. En 2006 la Agencia Internacional para la Investigación del Cáncer (IARC) clasificó el plomo inorgánico como probable carcinógeno para los humanos (3).

El Cadmio es un componente de la corteza terrestre que se encuentra en pequeñas cantidades asociado a minerales de cinc, cobre o plomo. Su contenido medio es minoritario y está estimado en torno a 0.10-0.15 mg/Kg. Su presencia en el medio ambiente es, por tanto, de origen natural, a través de la erosión, pero su nivel puede verse incrementado por la acción del hombre, debido principalmente a actividades como la minería o también derivado de su uso en distintas aplicaciones (pinturas, baterías, etc.) (4).

Se trata de un elemento que presenta numerosos efectos tóxicos, siendo la disfunción renal el principal efecto por una exposición prolongada, de ahí que tenga un impacto negativo sobre la calidad del alimento y esté sujeto a medidas, a nivel internacional, dirigidas a minimizar su presencia en los alimentos(1).

La absorción del cadmio en el aparato digestivo es baja (5-10%), sin embargo, el cadmio se acumula en el organismo, principalmente en el hígado y riñón, estando su vida media estimada en unos 20-30 años en humanos(5).



Figura 1. Imágenes de algunos panes de dulce estudiados en la investigación.

## 2. Procedimiento

### 2.1. Objetivo

Determinar las concentraciones de los metales pesados Plomo y Cadmio en muestras de pan de fabricación industrial realizando su cuantificación por Absorción Atómica.

## 2.2. Metodología

Para realizar las determinaciones se utilizó el equipo de absorción atómica modelo AAAnalyst-100 de Perkin Elmer, para poder realizar las determinaciones se le dio un tratamiento previo a las muestras, consistente en calcinación a 400°C en un horno de alta temperatura y su posterior digestión ácida (Figura 1).

### Fundamento del método

El método de absorción atómica se basa en hacer pasar un haz de luz monocromática de una frecuencia tal que puede ser absorbido por el analito que se encuentra presente en forma de vapor atómico. La medida de la intensidad luminosa antes y después de su paso por el vapor atómico permite determinar el porcentaje de absorción. La cantidad de absorción aumenta con la concentración de los átomos en el medio absorbente, es decir, la medida de la absorción aumenta con la concentración del elemento en la muestra, ya sea que esté en su condición original o sujeta a pretratamiento.

### Tratamiento de la muestra

Se puede realizar el tratamiento por dos vías que se explican a continuación.

#### 1. Digestión por vía húmeda

Pesar con precisión de  $\pm 0.1\text{mg.}$ , una cantidad apropiada de muestra. Para la determinación por el método de absorción por flama pesar como máximo 20 g. de alimentos que contengan del 50 al 75 % de agua y 10 g. de alimentos sólidos o semisólidos. Límite del contenido de grasa o aceite a un máximo de 4 g. y el total de materia orgánica a 5 g.

Añadir 10 ml. de ácido nítrico concentrado y dejar reposar toda la noche o iniciar directamente la digestión. Usar matraz de Kjeldhal o matraz conectado al sistema de refrigerantes. Calentar suavemente. Digerir la muestra 3 horas o más tiempo si es necesario (algunas muestras requieren la adición de mayor cantidad de ácido nítrico) hasta la aparición del color traslúcido, si queda ámbar, adicionar peróxido de hidrógeno gota a gota con agitación continua (reacción exotérmica). Enfriar. Recuperar, filtrar y llevar a un volumen conocido en matraz volumétrico. Correr un blanco de reactivos y muestra fortificada por cada serie de digestión. Leer en el aparato de elección (espectrómetro de absorción atómica por flama u horno de grafito).

#### 2. Digestión por vía seca

Para la determinación por el método de absorción por flama pesar como máximo, 20 g. de alimentos que contengan del 50 al 75 % de agua y 10 g. de alimentos sólidos y semisólidos. Límite del contenido de grasa o aceite a un máximo de 4 g. y el total de materia orgánica a 5 g. Añadir 10 ml. de ácido nítrico concentrado y dejar reposar toda la noche o iniciar directamente la digestión. En productos con alta concentración de proteínas adicionar una solución de nitrato de magnesio al 7.0 % p/v y mezclar completamente, llevar a sequedad aproximadamente durante 6 horas en estufa a una temperatura de 90 a 95°C. Colocar la muestra en una mufla y elevar la temperatura lentamente de 2 a 4°C por minuto hasta 350°C. Mantener la temperatura hasta que cesen los humos. Elevar gradualmente la temperatura de 500

a 550°C para evitar que la muestra se incinere y mantener esa temperatura durante 16 horas o toda la noche. Apagar la mufla y dejar enfriar.

Un segundo paso de calcinación puede ser requerido para remover algunos residuos de carbón, mediante el siguiente procedimiento: Lavar las paredes del crisol con 2 ml. de ácido nítrico al 50%. Colocar la muestra en una placa de calentamiento puesta a 120°C para remover el exceso de ácido. Colocar la muestra en una mufla fría y elevar la temperatura gradualmente de 500 a 550°C, manteniéndola por

Pesar con precisión de  $\pm 0.1mg.$ , una cantidad apropiada de muestra.

el tiempo necesario. Repetir este procedimiento cuantas veces sea necesario hasta que quede libre de carbón remanente. Disolver las cenizas completamente en 5 ml. de ácido clorhídrico 1N, transferir la muestra disuelta a un tubo de propileno o a un matraz de volumen conocido, enjuagar el crisol con dos alícuotas de 5 ml. de ácido clorhídrico 1 N y transferir al mismo tubo o matraz para obtener un volumen de 15 ml. en el primero y llevar al aforo en el segundo, tapar y mezclar, si existe presencia de partículas o materia insoluble, filtrar en papel Whatman No. 2, antes de la determinación. Correr un blanco de reactivos y muestra fortificada por cada serie de digestión. Leer en el espectrómetro de absorción atómica.

### 2.3. Resultado

Las especificaciones se muestran en la Tabla 1, de acuerdo a la Norma Oficial Mexicana NOM-247-SSA1-2008, Productos y servicios. Cereales y sus productos. Cereales, harinas de cereales, sémolas o semolinas. Alimentos a base de: cereales, semillas comestibles, de harinas, sémolas o semolinas o sus mezclas. Productos de panificación. Disposiciones y especificaciones sanitarias y nutrimentales. Métodos de prueba.

Tabla 1. Niveles de referencia límite máximo

Metales pesados	límite máximo mg/Kg.
Plomo	0.5
Cadmio	0.1

En la Tabla 2 se muestran las condiciones de trabajo del equipo

Tabla 2. Condiciones del equipo

Plomo estándares: 1.0,2.0 y 3.0 mg/L.	Cadmio estándares: 1.0,2.0 y 3.0 mg/L.
Longitud de onda (nm):283.3 Slit:0.7 Acetileno/Aire:1.3 Autocalibración	Longitud de onda (nm):228.8 Slit:0.7 Acetileno/Aire:1.3 Autocalibración

En la Tabla 3 se muestran los resultados cuantificados de acuerdo a las lecturas de cada muestra respectivamente, donde el término ND (significa, No detectado)

Tabla 3. Resultados en mg/Kg. de acuerdo a las lecturas obtenidas de cada metal.

<b>Muestra</b>	<b>Plomo mg/Kg.</b>	<b>Cadmio mg/Kg.</b>	
1)Donas	1.000	ND	<b>Fecha de muestreo</b> 12/Enero/2012
2)Tostado	0.365	ND	
3)Conchas	0.380	ND	
4)Blanco	0.370	ND	
5)Integral	ND	ND	
6)Roles	ND	ND	
7)Panque	2.75	ND	
1)Donas	0.692	0.046	<b>Fecha de muestreo</b> 10/Febrero/2012
2)Tostado	0.360	ND	
3)Conchas	1.625	ND	
4)Blanco	0.600	0.050	
5)Integral	0.536	0.074	
6)Roles	6.045	ND	
7)Panque	0.560	0.048	
1)Donas	0.408	0.018	<b>Fecha de muestreo</b> 12/Marzo/2012
2)Tostado	0.424	ND	
3)Conchas	0.378	ND	
4)Blanco	6.045	ND	
5)Integral	0.560	0.048	
6)Roles	0.408	0.018	
7)Panque	0.424	ND	
1)Donas	0.378	ND	<b>Fecha de muestreo</b> 18/Abril/2012
2)Tostado	3.55	ND	
3)Conchas	6.470	ND	
4)Blanco	0.192	ND	
5)Integral	0.322	0.018	
6)Roles	0.368	0.012	
7)Panque	0.452	0.036	
1)Donas	0.496	0.022	<b>Fecha de muestreo</b> 07/Mayo/2012
2)Tostado	0.488	0.024	
3)Conchas	0.422	0.006	
4)Blanco	0.498	0.028	
5)Integral	0.908	0.162	
6)Roles	2.276	0.615	
7)Panque	0.378	ND	

Tabla 3. Continuación de resultados en mg/Kg. de acuerdo a las lecturas obtenidas de cada metal.

<b>Muestra</b>	<b>Plomo mg/Kg.</b>	<b>Cadmio mg/Kg.</b>	
1)Donas	0.578	ND	<b>Fecha de muestreo</b> 08/Junio/2012
2)Tostado	0.542	ND	
3)Conchas	0.884	ND	
4)Blanco	0.798	ND	
5)Integral	0.656	ND	
6)Roles	0.800	ND	
7)Panque	0.368	0.012	
1)Donas	0.452	0.036	<b>Fecha de muestreo</b> 10/Julio/2012
2)Tostado	0.496	0.022	
3)Conchas	0.488	0.024	
4)Blanco	0.422	0.006	
5)Integral	0.498	0.028	
6)Roles	0.908	0.162	
7)Panque	2.276	0.615	
1)Donas	0.378	ND	<b>Fecha de muestreo</b> 06/Agosto/2012
2)Tostado	0.578	ND	
3)Conchas	0.542	ND	
4)Blanco	0.884	ND	
5)Integral	0.798	ND	
6)Roles	0.656	ND	
7)Panque	0.800	ND	
1)Donas	0.582	0.026	<b>Fecha de muestreo</b> 10/Septiembre/2012
2)Tostado	0.838	0.072	
3)Conchas	0.800	ND	
4)Blanco	0.60	ND	
5)Integral	5.20	ND	
6)Roles	1.00	ND	
7)Panque	0.08	ND	
1)Donas	8.28	0.50	<b>Fecha de muestreo</b> 11/Octubre/2012
2)Tostado	0.45	0.04	
3)Conchas	0.15	0.01	
4)Blanco	0.49	0.02	
5)Integral	4.03	ND	
6)Roles	2.02	ND	
7)Panque	0.16	ND	

Tabla 4, los promedios obtenidos para cada muestra y cada metal.

<b>Muestra</b>	<b>Plomo mg/Kg.</b>	<b>Cadmio mg/Kg.</b>
1)Donas	1.324	0.020
2)Tostado	0.809	0.016
3)Conchas	1.214	0.004
4)Blanco	1.090	0.010
5)Integral	1.351	0.033
6)Roles	1.448	0.081
7)Panque	0.825	0.071

### 3. Conclusiones

De acuerdo a los resultados obtenidos, en cuanto al contenido de plomo en las muestras, todas se encuentran fuera de la especificación que indica la NOM-247-SSA1-2008 lo cual es alarmante ya que las muestras tratadas son de un producto comercial que se consume en grandes cantidades todos los días por la mayoría de la población tanto adulta como infante. Respecto al contenido de cadmio todas las muestras se encuentran dentro de especificación.

### Referencias

- [1] S., Manahan. (2007). Introducción a la química Ambiental. Editorial Revert'e.
- [2] Norma Oficial Mexicana NOM-247-SSA1-2008. (2008), Productos y servicios. Cereales y sus productos. Cereales, harinas de cereales, sémolas o semolinas, Alimentos a base de: semillas comestibles, de harinas, sémolas o sus mezclas. Productos de panificación. Disposiciones y especificaciones sanitarias y nutrimentales. Métodos de prueba.
- [3] Modificación a la Norma Oficial Mexicana NOM-028-FITO-1995. (1995), Establecimiento de los requisitos fitosanitarios y especificaciones para la importación de granos y semillas excepto para la siembra.
- [4] NOM-086-SSA1-1994. (1994). Bienes y servicios. Alimentos y bebidas no alcohólicas con modificaciones en su composición. Especificaciones nutrimentales.
- [5] NOM-120-SSA1-1994. (1994). Bienes y servicios. Prácticas de higiene y sanidad para el proceso de alimentos, bebidas no alcohólicas y alcohólicas.

# Diagnóstico de cáncer de mama mediante una red neuronal usando meta-heurísticas

Rivera Martínez Marcela<sup>1</sup> Marcial Castillo Luis René<sup>1</sup>  
Sandoval Solís María de Lourdes<sup>1</sup>  
Allende Alonso Sira<sup>2</sup>

Línea de investigación: Ciencias Médicas.

<sup>1</sup>Benemérita Universidad Autónoma de Puebla.

<sup>2</sup>Universidad de la Habana, Habana, Cuba.

*14 Sur y Av. San Claudio, Col. San Manuel, C.U,  
Puebla, Puebla, CP 74570, México.*

*e-mail:cmr@cs.buap.mx*

## Resumen

En este trabajo, se implementa una red neuronal para decidir si un paciente tiene, o no, cáncer de mama. Se utiliza una red neuronal artificial multicapa y se entrena con las meta-heurísticas: colonia de abejas y búsqueda armónica. Se comparan los resultados de las meta-heurísticas usando la base de datos de la Universidad de Wisconsin.

**Palabras clave:** Cáncer, colonia de abejas, búsqueda armónica, red neuronal artificial.

## 1. Introducción

El cáncer de mama constituye un problema de salud pública y es la primera causa de incidencia y mortalidad por cáncer en la mujer adulta en Latinoamérica y en el ámbito mundial. En México el 40% de 36 casos diagnosticados de cáncer de mama corresponden a mujeres menores de 50 años. El cáncer mamario dejó de estar circunscrito a los países desarrollados y a mujeres con mayores recursos económicos, incluso ha desplazado al cáncer cervicouterino en varias regiones de Latinoamérica como primera causa de muerte por neoplasias malignas femeninas. El examen clínico y la mamografía, son las principales herramientas para el diagnóstico de cáncer mamario, ya que se ha demostrado que el diagnóstico temprano incrementa el tiempo de supervivencia y reduce la mortalidad para este tipo de cáncer [1]. En este trabajo, se implementa una red neuronal artificial multicapa para decidir si un paciente tiene o no cáncer de mama, para encontrar los pesos óptimos de la red neuronal se utilizan dos meta-heurísticas: colonia de abejas y búsqueda armónica. Las implementaciones realizadas se prueban usando la base de datos de la Universidad de Wisconsin [2]. En la siguiente sección se presenta el concepto de red neuronal, el uso de éstas en la clasificación y el aprendizaje de una red neuronal, en la sección tres se

explica el funcionamiento de la meta-heurística colonia de abejas artificiales y se muestra el algoritmo de dicha meta-heurística, la meta-heurística búsqueda armónica se explica en la sección cuatro, la sección cinco muestra las pruebas que se realizaron así como la discusión de los resultados, las conclusiones aparecen en la sección seis, y finalmente se presentan las referencias que se utilizaron en el desarrollo de éste trabajo.

## 2. Redes neuronales

Una red neuronal artificial es un modelo que emula el comportamiento de las redes neuronales biológicas, describiéndolas como una unidad de proceso que actúa en paralelo, que consta de unas Entradas  $x_i$  equivalentes a las dendritas en su paralelo biológico. Dichas entradas están ponderadas por pesos  $w_i$ , que corresponden a la forma en que se evalúan los impulsos eléctricos y se combinan con la función de red, la cual define el nivel de potencial de la neurona. La salida de la función de la red es evaluada en la función de activación que da lugar a la salida de la unidad de proceso. La manera por la cual las redes neuronales artificiales realizan el proceso de cambio de la ponderación de las entradas mediante el cambio de pesos se denomina aprendizaje de la red, el modo de aprendizaje se dictamina de acuerdo a las condiciones de la misma red o a su configuración, entendiéndose esto último como la forma en que se organizan las neuronas con otras unidades, estableciendo la arquitectura de la red neuronal [3, 4].

### 2.1. Clasificación

El problema de clasificación es un problema actual de gran interés debido a que se utiliza en muchas áreas tales como medicina, biología, finanzas, etc [5, 6].

Un acercamiento al reconocimiento de patrones asume que el mundo está dividido en un número de categorías discretas. Dado un conjunto de datos de entrada, el dispositivo para el problema de clasificación debe decidir qué categoría es la más apropiada. Supóngase que se tiene un conjunto de datos de entrada, en el contexto de las redes neuronales, esto correspondería a un conjunto de patrones, y existe un número de categorías de salida. Esta estructura está basada en observaciones sobre categorías del mundo real. Las cosas que se encuentran muy cercanas entre sí en el espacio de estados, tienen muchas probabilidades de encontrarse en la misma categoría. Claramente la clasificación hace algo más que una simple categorización, acopla una rica estructura asociativa con la flexibilidad. El beneficio que aportan las redes neuronales a este problema es que con un conjunto pequeño de datos, la red puede aprender, para poder clasificar en un momento dado cualquier dato aunque no se le haya presentado nunca antes, es por ello que una parte importante de las redes neuronales es el aprendizaje. En este trabajo, se utiliza una clasificación binaria para decidir si un paciente tiene cáncer o no lo tiene. [3, 4].

### 2.2. Aprendizaje

El aprendizaje en una red neuronal, es la forma en la cual la red aprenderá a hacer la tarea para la que se está creando, por ejemplo en clasificación, es la manera en la cual la red neuronal podrá clasificar cualquier objeto que se le presente a partir de un entrenamiento previo, es decir, la red neuronal clasifica a través de información adecuada de los objetos a clasificar, en esta fase, los objetos de entrenamiento pueden ser pocos.

Cuando se habla de aprendizaje supervisado, se necesita de un conjunto de datos llamado *conjunto de entrenamiento*, este conjunto está formado por unos cuantos datos para los cuales se conoce la salida correcta, nuevamente, si se habla del problema de clasificación, se tiene un conjunto de datos para los cuales se conoce a qué clase pertenecen. Cabe mencionar que es común que la fase de entrenamiento tarde minutos o incluso horas, dependiendo de la estructura y dimensión del problema, sin embargo, después de esta fase, la clasificación es muy rápida tomando sólo segundos. Lo que se trata de hacer con el aprendizaje o entrenamiento de la red, es que a partir de un conjunto de datos de entrada, del cual se conoce la salida correcta, la red neuronal sea capaz en otra etapa posterior, de responder por sí sola a situaciones diferentes a las aprendidas, es decir, se trata de que con el aprendizaje, la red neuronal vaya ajustando adecuadamente todos y cada uno de los pesos asociados a cada arista con el fin de minimizar el error cometido entre la salida obtenida por la red neuronal y la salida deseada para lograr una respuesta correcta ante una situación totalmente nueva [3, 4]. Existen diferentes formas de minimizar este error, entre ellas tenemos a: máximo descenso, gradientes conjugados, búsqueda lineal inexacta, base radial, backpropagation, entre otras, sin embargo, estos algoritmos sólo son eficientes para encontrar mínimos locales, en este trabajo se propone utilizar las meta-heurísticas de colonia de abejas y búsqueda armónica las cuales son eficientes en encontrar mínimos globales para funciones multimodales [7, 8].

### 2.3. Arquitectura

La arquitectura de una red es una estructura o patrón de conexión. En ésta arquitectura las neuronas se organizan en una secuencia de capas formando un patrón de interconexiones entre los diferentes elementos dependiendo de la posición de las capas. En un perceptrón multicapa existen tres tipos de capas: de entrada, salida y ocultas. La capa de entrada está compuesta por neuronas que reciben información proveniente del entorno. La capa de salida es la que proporciona la respuesta de la red neuronal. La capa oculta es la que no tiene relación alguna con el exterior. Este tipo de capa agrega a la red, grados de libertad adicionales, con los cuales es posible mejorar la capacidad del modelo (ver Figura 1) [3, 4].

### 2.4. Diagnóstico de cáncer de mama

En este caso, se utiliza una clasificación binaria para decidir si un paciente tiene cáncer o no. Dada una serie de datos representados por vectores  $m$ -dimensionales de atributos que constituye el espacio de entradas  $X$ , cada dato posee una etiqueta (clase), entre dos únicas opciones, que pueden ser representadas por las etiquetas 1 y -1, estas etiquetas forman el dominio de salida  $Y$ . La clasificación consiste en determinar una función a partir de un conjunto de entrenamiento, formado por pares de datos patrón-clase, que pronostique la clase a la cual pertenezca un nuevo patrón de entrada.

En este trabajo los datos se tomaron de la base de datos de la Universidad de Wisconsin, la cual contiene 749 muestras, cada una de ellas presenta 9 características, las cuales fueron calculadas a partir de una imagen digitalizada de una masa en la mama. Las características de los núcleos de las células son las siguientes: radio (medida de la distancia desde el centro hasta el perímetro), la textura (desviación estándar de los valores en la escala de grises), área, grado de suavidad (variación local en longitudes de radio), compactamiento,

concauidad (severidad de porciones cóncavas del contorno), puntos de concauidad (número de porciones cóncavas del contorno), simetría y dimensión fractal [2].

### 2.5. Arquitectura de la red neuronal para el diagnóstico de cáncer de mama

Una arquitectura de red neuronal es una 4-upla  $(I, N, O, E)$  donde  $I$  es un conjunto de sitios de entrada,  $N$  es un conjunto de unidades de cálculo,  $O$  es un conjunto de sitios de salida y  $E$  es un conjunto de arcos dirigidos con pesos. Un arco ponderado es, a su vez, una 3-upla  $(u, v, w)$  donde  $u \in I \cup N$ ,  $v \in N \cup O$  y  $w \in R$  [4]. La red neuronal usada en éste trabajo, para realizar el diagnóstico, es una red multicapa con 9 sitios de entrada que corresponden al número de características del cáncer de mama de acuerdo a la base de datos de la Universidad de Wisconsin, más el umbral para cada neurona de la primera capa, además se tienen dos capas ocultas, la primera con 10 neuronas y la segunda con otras 10, el número de nodos de la capa de salida es 2 que corresponde a las dos posibles salidas: tiene cáncer o no lo tiene. Los autores de este trabajo también han utilizado una arquitectura similar a la propuesta presentada en este trabajo, para la clasificación de dígitos manuscritos [9].

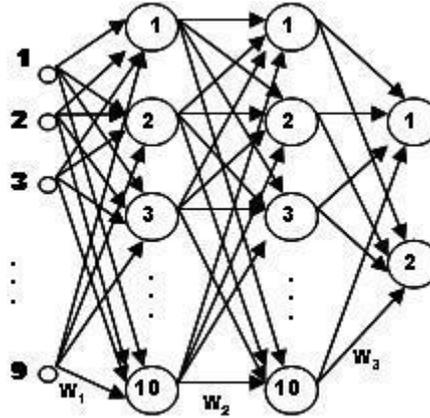


Figura 1. Arquitectura de la red neuronal

## 3. Colonia de abejas artificiales

Un enjambre de abejas es un colectivo inteligente, en el cual los miembros del mismo asignan las tareas de manera dinámica, además, el enjambre se adapta en respuesta a los cambios de su medio ambiente. Las abejas poseen memoria fotográfica y aeroespacial, toman decisiones en grupo durante el proceso de selección de sus nuevos sitios para anidar, realizan tareas como almacenar, distribuir la miel y el polen, estas características han sido los incentivos de los investigadores para modelar el comportamiento inteligente de las abejas.

Las abejas tienen varias características que se modelan para los algoritmos entre éstas tenemos: la danza, la alimentación, selección de tareas, toma de decisiones de manera colectiva y selección del nido entre otras. La danza de la abeja, en general, sirve para la comunicación, existen diferentes tipos de ésta dependiendo de lo que se quiera comunicar, por ejemplo, una vez que una abeja ha encontrado una fuente de alimento ésta realiza ciertos movimientos (danza) con el fin de compartir su información acerca de la dirección y la distancia de la fuente de alimento, así como si ésta es abundante y de calidad.

Existen muchos algoritmos de abejas, de los cuales, el más popular en la actualidad es el que Karaboga y Basturk introdujeron en 2006, llamado colonia de abejas artificiales (ABC), el cual está desarrollado en base a la inspección del comportamiento real de las abejas en la búsqueda de néctar y en compartir la información de las fuentes de alimento a las abejas en la colmena. El objetivo de estas abejas es localizar las fuentes de alimento con mayor néctar y almacenarlo en el panal con una división eficiente del trabajo [10]. El modelo de colonia de abejas artificial conocido como ABC se compone de tres tipos de abejas: las abejas empleadoras, las abejas observadoras y las abejas exploradoras. Las abejas empleadoras son las responsables de explotar las fuentes de néctar y explorar antes de dar información a las abejas observadoras, las abejas observadoras esperan en la colmena y deciden explotar una fuente de néctar y esto lo hacen en base a la información dada por las abejas empleadoras, las abejas exploradoras realizan una búsqueda en el medio ambiente con el fin de encontrar una nueva fuente de alimento en función de una motivación interna, posibles pistas externas o bien mediante una búsqueda al azar. Siendo  $Cfa$  la cantidad de fuentes de alimento, el algoritmo básico es el siguiente:

1. Inicializar de modo aleatorio las fuentes de alimento  $x_i$ :

$$x_i \sim (li, ui), \text{ para } i \text{ de } 1 \text{ a } Cfa.$$

Donde  $li < ui$ ,  $x_i$  toma valores aleatorios en el rango de  $li$  a  $ui$ , y  $Cfa$  es la cantidad de fuentes de alimento.

2. Cada abeja empleadora produce nuevas fuentes de alimento  $x_n$  buscando en la vecindad de la fuente de alimento  $x_i$  seleccionada de forma aleatoria, es decir:

$$x_n \leftarrow x_i + \varphi * (x_i - x_k), k \sim [1, Cfa], k \in Z \text{ y } \varphi \sim (-1, 1).$$

Donde  $k$  se selecciona aleatoriamente en el rango de 1 a  $Cfa$  y  $\varphi$  se selecciona de modo aleatorio en el intervalo abierto de -1 a 1.

Si  $x_n$  es mejor que la fuente de alimento  $x_i$  entonces  $x_i \leftarrow x_n$ .

3. Cada una de las abejas observadoras selecciona una fuente de alimento y dependiendo de la calidad, explota la fuente. Se elige la fuente de alimento mediante una ruleta, en donde se toma en cuenta la probabilidad dada por la ecuación siguiente:

$$p_i = \frac{f(x_i)}{\sum_{i=1}^{Cfa} f(x_i)}.$$

La probabilidad  $p_i$  se calcula como el cociente de la evaluación de la fuente de alimento  $x_i$  en la función de costo  $f(x_i)$  y la suma de la evaluación de los costos para todas las fuentes de alimento.

Si la fuente de alimento encontrada es mejor se actualiza la fuente.

4. Determinar la fuente de alimento que será abandonada. Si una fuente de alimento no se ha actualizado en una cantidad de ciclos, la abeja exploradora la cambia por una nueva generada aleatoriamente.
5. Memorizar la mejor fuente de alimento.  $Se \leftarrow$  mejor fuente de alimento.
6. Si el criterio de paro se cumple salir, en otro caso ir al paso 2.

El criterio de paro usado en este trabajo es por tiempo.

## 4. Búsqueda Armónica

La búsqueda armónica se puede explicar con más detalle con la ayuda de la discusión del proceso de la improvisación de un músico. Cuando un músico improvisa, tiene tres opciones: 1) reproducir cualquier pieza famosa de música exactamente de memoria; 2) tocar algo similar a la pieza más conocida; 3) redactar notas nuevas o al azar. Si se formalizan estas tres opciones para la optimización, tenemos tres correspondientes componentes: el uso de la memoria de la armonía, el ajuste de tono, y el azar.

La estimación estética se lleva a cabo uniendo los tonos que puede ofrecer cada instrumento, y los sonidos pueden mejorar con la práctica, análogamente en el caso de optimización la evaluación de la función objetivo se realiza por el conjunto de los valores producidos por las variables, y éstos pueden mejorar en cada iteración. En la optimización real, cada músico puede ser sustituido por cada variable de decisión, y sus campos de sonido de su preferencia pueden ser sustituidos por los valores preferidos de cada variable, es por ello que este algoritmo puede manejar problemas mixtos, por ejemplo, variables discretas y continuas en un mismo problema [11].

El uso de la memoria de la armonía es importante ya que es similar a elegir el mejor individuo, esto garantizará que la mejor armonía se trasladará a la memoria de la nueva armonía. Con el fin de utilizar esta memoria con mayor eficacia se asigna un parámetro  $rac \in [0, 1]$  llamado tasa de aceptación de la armonía, si esta tasa es demasiado baja, sólo algunas de las mejores armonías son seleccionadas lo cual dará como resultado una convergencia lenta, si por el contrario la tasa es muy alta (cercana a uno), casi todas las armonías se utilizarán en la memoria, por lo general  $rac = 0.7 \sim 0.95$ .

El ajuste del tono ( $rpa$ ) que es el segundo componente se realiza de manera lineal y está definido por:

$$x_{actual} = x_{anterior} + b_p(2 * rand - 1), \quad (1)$$

donde  $rand$  es un número aleatorio de una distribución uniforme entre  $[0, 1]$ ,  $b_p$  es el ancho de banda que controla el rango local del ajuste del tono. Por lo general  $rpa$  está entre 0.1 y 0.5. El algoritmo general en pseudocódigo es el siguiente:

Definir la razón de ajuste del tono( $rpa$ )

Definir la razón de aceptación( $rac$ )

Generar las armonías iniciales.

**Mientras** ( $t <$  número máximo de iteraciones)

**Mientras** ( $i \leq$  número de variables)

**Si** ( $rand < rac$ )

            Elegir un valor armónico para la variable  $i$

```

Si ( $rand < rpa$ )
    Ajustar los valores
Fin (si)
Si no
    Elegir un valor aleatorio
Fin (si)
Fin (mientras)
    Actualizar la nueva armonía
Fin (mientras)

```

## 5. Resultados y discusión

Las implementaciones fueron realizadas bajo Octave [12]. En el algoritmo de colonia de abejas se utilizaron 8 fuentes de alimento, 4 abejas empleadoras, 4 observadoras y una exploradora, en el algoritmo de búsqueda armónica se utilizó una tasa de aceptación de 0.9, un ajuste de tono de 0.3 y una memoria de 8. La computadora donde se realizaron los experimentos es una computadora con procesador INTEL Core 2 QUAD a 3.0Ghz y 2GB de memoria RAM. Se realizaron dos experimentos, el primer experimento consta de tres pruebas, en la prueba 1 se tomaron los primeros 100 elementos de la base de datos de la Universidad de Wisconsin y se entrenó durante 20 minutos, en la segunda prueba los primeros 150 elementos de la base de datos fueron tomados y se entrenaron en un lapso de 30 minutos, finalmente en la tercer prueba se entrenó la red por un lapso de 30 minutos y para ello se tomaron los primeros 350 elementos de la base de datos, los resultados del porcentaje de diagnóstico son los mostrados en el Cuadro 1.

CANTIDAD DE DATOS	ABEJAS	B. ARMÓNICA.	TIEMPO
100	99.00	94.00	20 min.
150	99.33	92.00	30 min.
350	92.57	84.85	30 min.

Cuadro 1. Porcentajes de diagnóstico en la fase de entrenamiento.

Se observa en el cuadro anterior, que cuando se tomaron 100 datos para realizar el entrenamiento durante 20 minutos, la meta-heurística de colonia de abejas pudo diagnosticar de manera correcta el 99 %, mientras que la meta-heurística de búsqueda armónica sólo logra diagnosticar el 94 % de los 100 datos. En 30 minutos de entrenamiento la meta-heurística de colonia de abejas diagnostica el 99.33 % contra 92 % que logra diagnosticar la meta-heurística de búsqueda armónica de 150 datos. Cuando se toman los primeros 350 datos, la meta-heurística de colonia de abejas diagnostica de manera correcta el 92.57 %, lo cual pareciera si se compara con el caso anterior de 150 datos que el porcentaje del diagnóstico disminuye, lo cual no es cierto, si tomamos en cuenta que el tiempo de entrenamiento sigue siendo el mismo (30 minutos) y la cantidad de datos se ha incrementado a un poco más del doble.

En la figura 2 se muestra la función de error durante la fase de entrenamiento con 100 datos para las dos meta-heurísticas utilizadas en este trabajo. Se puede observar que aún cuando el algoritmo de búsqueda armónica decrementa muy rápido la función de error, antes de 2 minutos, el algoritmo de colonia de abejas es el que logra finalmente obtener un mejor error, es decir, logra que éste sea más pequeño.

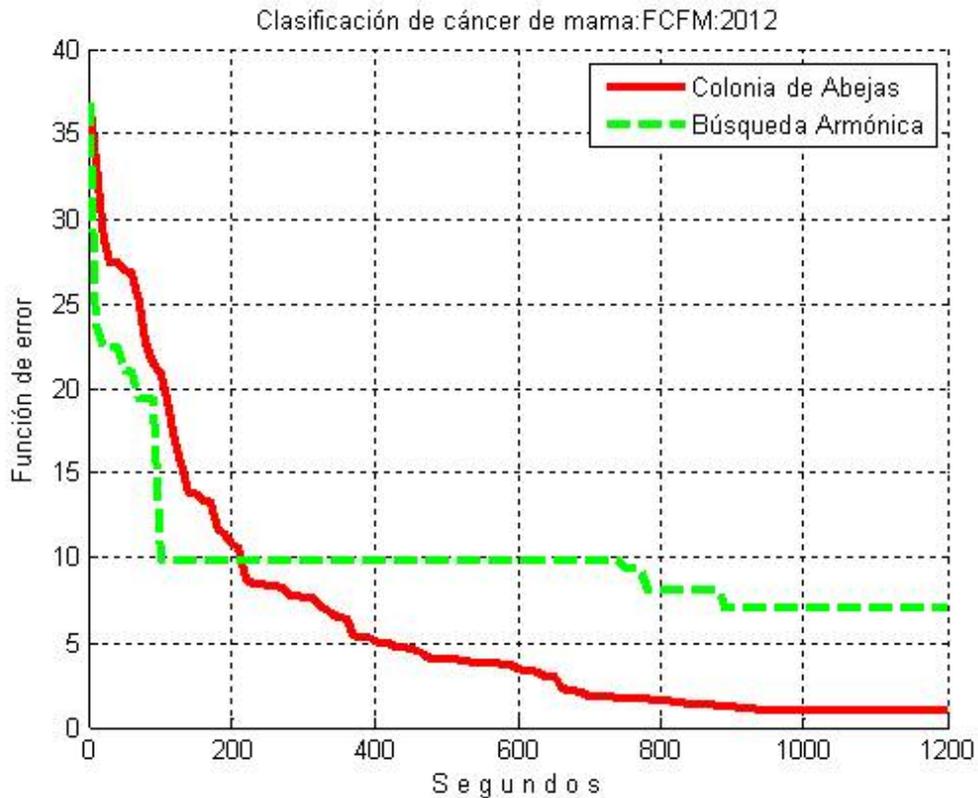


Figura 2. Función de error durante el entrenamiento

En el segundo experimento, en donde se llevaron a cabo 3 pruebas, se eliminan en cada prueba, los elementos entrenados en el experimento 1, por ejemplo, en la primera prueba se toman sólo 649 elementos de la base de datos porque en el primer experimento en la prueba 1 se entrenó a la red con los primeros 100 elementos, los cuales se omiten en esta ocasión. De igual manera, se procedió con los experimentos restantes, se eliminaban los elementos que se utilizaron en la fase de entrenamiento y se probó la red con el resto de los datos. Los resultados del porcentaje de diagnóstico se muestran en el Cuadro 2.

CANTIDAD DE DATOS	ABEJAS	B. ARMÓNICA
649	83.66	80.89
599	93.82	78.96
399	93.98	88.25

Cuadro 2. Porcentajes de diagnóstico en la fase de prueba.

Como puede observarse en el cuadro anterior, la red neuronal que utiliza el algoritmo de colonia de abejas como método de aprendizaje supervisado, obtiene un muy buen diagnóstico de los datos, pues aún cuando se tiene el caso en que sólo se haya entrenado con 100 datos, logra diagnosticar el 83.66 por ciento correctamente. En la figura 3 se muestra la función de error obtenida en la fase de prueba con 399 datos. Se puede observar que el algoritmo de colonia de abejas es el que decrementa más y mejor la función de error.

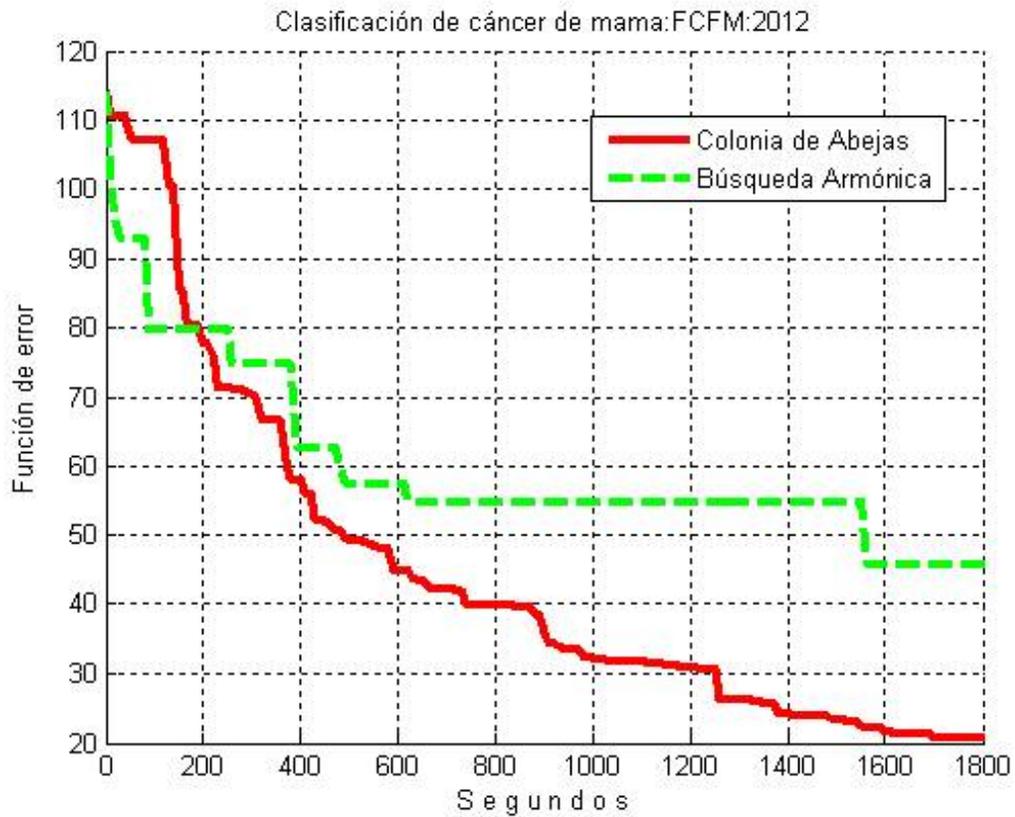


Figura 3. Función de error durante la fase de prueba

## 6. Conclusiones

Las conclusiones son las siguientes:

1. La arquitectura propuesta, esto es, una red neuronal multicapa con dos capas ocultas cada una de ellas con 10 neuronas, es adecuada para resolver el problema de diagnóstico de cáncer de mama.
2. El uso de las meta-heurísticas en el entrenamiento de la red neuronal, se refleja en el alto porcentaje correcto de diagnóstico que se realiza en ambas fases, en la de entrenamiento y en la fase de prueba.
3. Para todos los casos en los dos experimentos, el algoritmo de colonia de abejas es el mejor, al tener un alto porcentaje de diagnóstico correcto de cáncer de mama tanto en la fase de entrenamiento así como en la fase de prueba.

## Referencias

- [1] F. Marie, López L. et al. (2009). Cáncer de mama: un reto para la sociedad y los sistemas de salud, Salud Pública de México, Vol. 51, Supl. 2.
- [2] [www.archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](http://www.archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)) (última fecha de acceso: Agosto de 2012).
- [3] Anderson J. (2007). Redes Neurales, Alfaomega.
- [4] Rojas R. (1996). Neural Networks A Systematic Introduction, Springer-Verlag, Berlin.
- [5] Malmgren H., Borga M. and Niklasson L. (200). Artificial Neural Networks in Medicine and Biology, Proceedings of the ANNIMAB-1 Conference, Goteborg, Sweden.
- [6] <http://www.cob.calpoly.edu/~eli/pdf/neural.pdf> (última fecha de acceso: Marzo de 2013).
- [7] Worasuchep C. (2011). A Harmony Search with Adaptive Pitch Adjustment for Continuous Optimization, International Journal of Hybrid Information Technology, vol 4. No. 4.
- [8] Stanarevic N. (2012). Hybridizing artificial bee colony (ABC) algorithm with differential evolution for large scale optimization problems, International Journal of Mathematics and Computers in Simulation, Issue 1, Vol. 6, pp.194-202.
- [9] Marcial L., Rivera M., Sandoval M. et al. (2011). Algoritmos bio-inspirados para entrenar una red neuronal que reconoce dígitos manuscritos, Memorias en extenso del IV Taller Latino de Investigación de Operaciones. ISBN 978-607-7760-56-6.
- [10] Basturk B., D. Karaboga (2006). An artificial bee colony (ABC) algorithm for numeric function optimization, In IEEE Swarm Intelligence Symposium 2006, Indianapolis, Indiana, USA.
- [11] Z. Geem (2006). Music-inspired harmony search algorithm: theory and applications, Springer.
- [12] [www.gnu.org/software/octave](http://www.gnu.org/software/octave) (última fecha de acceso: Octubre de 2012).

# Diagnóstico de los residuos sólidos urbanos generados en un sector del estado de Morelos.

Domínguez Ortiz Marco G. <sup>1</sup>

Cortez Lugo Marlene<sup>1</sup>

Pacheco Magaña Lilian E. <sup>1</sup>

Arenas Monreal Luz <sup>1</sup>

Línea de investigación: Ciencias Médicas

<sup>1</sup>Instituto Nacional de Salud Pública

*Ave. Univ. No. 655 Col. Sta. M. Ahuacatlán, Cerrada Los Pinos  
y Caminera, Cuernavaca, Morelos, CP 62100, México  
mark0097@hotmail.com*

## Resumen

**Introducción:** El presente artículo es el resultado de un estudio de composición de los residuos sólidos urbanos (RSU) efectuado en la colonia Atlahuayan, municipio de Yautepec, Morelos, México. Durante el periodo: agosto 2011, al mes de abril del año 2012. Los RSU, están conformados por los desechos de los ciudadanos. Los cuales no tienen valor ni utilidad para la mayoría de las personas **Objetivo:** Realizar un diagnóstico del estado actual que guardan los residuos sólidos urbanos (RSU) generados en la colonia Atlahuayan. **Marco teórico:** El estudio se realizó mediante el modelo eco sistémico el cual incluye participación, equidad y genero, transdisciplina, mas la estrategia de separación de residuos en sus componentes orgánico e inorgánico. **Aplicación:** Los resultados de este diagnóstico habrá de contribuir para crear una fuente de datos actuales que sirva de base para la elaboración de un plan de gestión integral de los RSU. **Metodología** Diseño de estudio comparativo entre componentes cualitativos y cuantitativos de los residuos sólidos urbanos (RSU), se efectuó pesado y segregación de RSU en componentes orgánico e inorgánicos generados por 330 habitantes ubicados en 80 viviendas, a fin de obtener la producción per cápita de RSU. **Modelo:**  $Ppcm = \frac{\text{Cantidad de basura}}{\text{Total de la población}}$  = (Kg/hab/día) **Resultados:** Actualmente el municipio Yautepec cubre el 95 % de la recolección de RSU. Estos residuos tienen como disposición final el tiradero a cielo abierto “El ojo seco” que se encuentra ubicado a escasos metros de la carretera federal Yautepec-Tepoztlan. La colonia de Atlahuayan está conformada por 731 familias (2292 hab.). La producción per cápita promedio fue 0.365 Kg/día/hab. En subproductos se encontró: residuos orgánicos 53 % y 47 % inorgánicos. **Conclusiones:** A nivel nacional, falta un sistema de gestión integral de RSU, por lo que no está estandarizado el sistema, de tratamiento, recolección y disposición final, se carece de datos actualizados por lo que la información es limitada y escasa. Referente a la composición de RSU el 53 % corresponde a residuos tratables, es decir

biodegradables que pueden destinarse a composta o lombricomposta y 23 % reciclables. Si se procediera a la separación de subproductos, únicamente se depositarían al relleno sanitario el 24 % de los RSU totales generados, reduciendo los costos, por disposición final.

**Palabras clave:** Residuos sólidos urbanos; composición; Morelos; Yauatepec

## 1. Introducción

El presente artículo es resultado de un estudio de composición de los residuos sólidos urbanos (RSU) de la colonia Atlahuayan, municipio de Yauatepec, Morelos, México. Los RSU, están conformados por los desechos de los ciudadanos. Los cuales no tienen valor ni utilidad para la mayoría de las personas [1].

A nivel mundial los RSU procedentes de las viviendas se han considerado un problema que es necesario tratar. A nivel local la producción de residuos sólidos municipales sigue aumentando tanto en términos absolutos como per cápita en todo el mundo.

El manejo inadecuado de ellos crea riesgos al medio ambiente y a la salud humana [2].

En México en el año 2010 se reportó una generación de RSU totales de 30,783,527 ton/año de los cuales se reportó que el 75 % de los RSU provienen de origen domiciliario, siendo su disposición final los tiraderos clandestinos, tiraderos a cielo abierto y rellenos sanitarios [3].

El municipio de Yauatepec, está integrado por 62 colonias, las cuales producen 62.5 toneladas de RSU por día; La recolección de los RSU de estas colonias, es brindada por el municipio y servicios concesionados.

Cabe señalar que a la fecha de la realización del estudio, Yauatepec cuenta con un tiradero a cielo abierto, el cual no cumple con la normatividad establecida, NOM-083-SEMARNAT-2003[4] la cual establece las especificaciones de selección del sitio, el diseño, construcción, operación, monitoreo, clausura y obras complementarias de un sitio de disposición final de residuos sólidos urbanos y de manejo especial.

La colonia de Atlahuayan pertenece al área geoestadística básica (AGEB) 015-0 y está conformada por 731 familias (2,292 hab.), con un promedio de integrantes de 3.9 (rango: 1 - 9); un tercio de esta población, se encuentra dentro del rango de edad de 0-14 años.

## 2. Objetivo General

El objetivo de este estudio fue realizar un diagnóstico del estado actual que guardan los residuos sólidos urbanos generados en la colonia Atlahuayan, Yauatepec, Morelos.

### 3. Marco Teórico

La presente investigación se efectuó bajo el modelo ecosisemico, el cual integra tres componentes importantes: participación, transdisciplina y equidad-género. Puesto que la participación asegura el involucramiento de la población, se toma en consideración su perspectiva, conocimiento y experiencias, reconociendo la voz de la comunidad, como punto de vista relevante en la identificación de necesidades y alternativas de solución[5].

Equidad-género, plantea el involucramiento de los distintos grupos sociales dentro de la localidad, niños, jóvenes, mujeres, adultos mayores, grupos organizados por intereses comunes, maestros, ejidatarios, líderes, para la identificación de los problemas dentro de la comunidad y de las causas de la inequidad en salud. Además de considerar las diferencias de género en todas las esferas de la vida comunitaria [5].

Transdisciplina se refiere a enfrentar los problemas desde diversas disciplinas y puntos de vista. Así, desde la transdisciplina se propone que se integren miembros de la población, autoridades (grupos de tomadores de decisiones) y un equipo de profesionistas de diversas disciplinas [5].

Como estrategia para la segregación de los RSU se utilizó la separación mediante sus componentes orgánicos e inorgánicos. Para efectos de la presente investigación se considera el concepto de residuos orgánicos como todo desecho de origen biológico que alguna vez fue parte de un ser vivo y el residuo Inorgánico es todo desecho de origen no biológico, es decir, desecho de origen industrial.

### 4. Aplicación

Los resultados de este diagnóstico habrá de contribuir para crear una fuente de datos actuales que sirva de base para la elaboración de un plan de gestión integral de los RSU debido a que actualmente existen limitadas publicaciones y estudios enfocados a la solución integral de la acumulación de la basura en los ámbitos público y privado. Las posibilidades de tomar en cuenta las poblaciones pequeñas y los sitios más apartados dentro de la comunidad donde habita la población más vulnerable, sin duda redundaran en beneficios para los humanos y el planeta. La posición del conocimiento contribuye a la sensibilización en materia de gestión política nacional donde es importante aplicar la normatividad, capacitar a los funcionarios públicos desde el ámbito nacional hasta las localidades más pequeñas del país propiciando mayor efectividad en el control, separación, manejo y disposición final de los RSU.

Dado que la solución sería intersectorial, la labor de educación para la salud puede aprovechar estos datos de los riesgos y daños que ocasionan los RSU pues considerando que la acumulación de los RSU constituye un problema de salud pública, debido a que la acumulación de los RSU es un medio adecuado para la proliferación de vectores, fauna nociva, (como moscas, mosquitos, cucarachas, ratas, etc.) que constituyen reservorios de parásitos (como cerdos, perros, etc.), algunos de estos vectores y fauna nociva han sido asociados con efectos en la salud de las personas, por ejemplo: a) las cucarachas con fiebre tifoidea, diarrea, lepra y gastroenteritis, ratas con peste, tifo, rabia, diarrea y leptospirosis,

b) moscas con cólera, fiebre tifoidea, diarrea, salmonelosis y disentería y c) los mosquitos con malaria, dengue, encefalitis vírica y fiebre amarilla entre otros. [6]-[10].

En cuestión de la salud ocupacional de los pepenadores, el manejo constante de RSU se ha asociado a enfermedades respiratorias, de la piel y disnea provocado por inhalación de gas nitrógeno [11]-[12].

La magnitud y trascendencia de esta problemática ambiental tiene solución con la participación conjunta de la ciudadanía, las organizaciones no gubernamentales y el gobierno.

#### 4.1. Metodología

Se realizó un estudio comparativo entre componentes cualitativos y cuantitativos de los Residuos Sólidos Urbanos (RSU).

El estudio se efectuó en la colonia Atlihuahayan, municipio de Yautepec, Morelos; en el periodo: agosto del 2011 al mes de abril del año 2012. Los participantes fueron habitantes permanentes de la colonia de Atlihuahayan y servidores públicos del departamento de servicios municipales.

Para la evaluación del departamento de servicios públicos del municipio de Yautepec, se aplicó un cuestionario con el fin de conocer aspectos relacionados con la infraestructura, recursos humanos, marco legal, cobertura de recolección y tratamiento de los RSU del municipio, así como también se efectuó un registro etnográfico.

Para la generación y composición de los RSU se acudió a cada domicilio familiar a lo largo de 16 días (englobando 5 casas por día hasta alcanzar 80 viviendas). Durante las tardes se asistió a los domicilios de las familias participantes efectuando un recorrido de calle por calle para el pesado de los residuos sólidos, el procedimiento consistió en preguntar cuántos habitantes había en la residencia y la cantidad de días que la basura se encontraba almacenada, la cual fue pesada (mediante una báscula TANITA modelo 1631, con una capacidad máxima de 150 kg y mínima de 0.500 gramos) en su totalidad. Para la selección y cuantificación de los RSU se segregaron en sus componentes orgánico, inorgánico, tereftalato de polietileno (P.E.T.), vidrio, aluminio, polietileno de alta densidad (H.D.P.E.), papel-cartón, hojalata y otros. Los datos obtenidos se registraron en la bitácora de pesado usando el estándar de medición en kilogramos.

La producción per cápita de los RSU totales (orgánicos + inorgánicos) se calculó mediante la fórmula  $P_{pcm}$ .

Los datos obtenidos durante el proceso de separación y pesado de residuos fueron transcritos en una hoja de cálculo Microsoft Excel 2007 donde se procedió a efectuar el análisis y determinación de la producción per cápita de los RSU generados por los pobladores de la colonia de Atlihuahayan.

Para el registro etnográfico se utilizó un diario de campo donde se registró una combinación de técnicas: la observación efectuada por el investigador y la información proporcionada por los habitantes de la colonia, mas la información obtenida por entrevistas no estructuradas, así como la aplicación de un cuestionario para la investigación dirigida al personal del departamento de servicios generales del municipio.

Los datos obtenidos se transcribieron en el procesador de texto Microsoft Word 2007, este archivo se transformo en el software etnograf versión 5. Después se realizaron cate-

gorías según los objetivos del proyecto, y fueron analizados mediante la técnica de matriz de análisis.

En el registro etnográfico se identificaron 3 categorías que se describen a continuación:

- 1) Caracterización del sistema del servicio de limpia del municipio.
- 2) Disposición de la basura de la colonia.
- 3) Información de parte de los habitantes en relación al servicio proporcionado por el camión recolector.

## 4.2. Modelo

En el presente estudio la producción per cápita de los RSU totales (orgánicos + inorgánicos) se calculó mediante la siguiente fórmula:

$$Ppcm = \frac{\text{Cantidad de basura}}{\text{Total de la población}} = (\text{Kg/hab/día})$$
 Donde: Ppc: producción per cápita de RSU de las familias participantes. Cantidad de basura: peso de los RSU (orgánicos e inorgánicos). Total de la población: Es la cantidad de personas que integran las 80 familias. Con base en la Ppcm de la muestra se estimó la Ppct de todos los habitantes de la colonia, tomando como total de la población a 2292 habitantes.

## 5. Resultados y discusión

### Resultados.

Los resultados son demostrativos en el municipio de Yautepec, Morelos, se cubre a la fecha del estudio el 95% RSU a través de la recolección municipal, disponiendo de una infraestructura física para la disposición final, la cual consiste en un tiradero a cielo abierto denominado “El ojo seco”, el cual se encuentra ubicado en la Colonia el Zarco, a la altura de la carretera Federal Yautepec-Tepoztlan, el cual no cumple con la normatividad establecida NOM-083-SEMARNAT-2003 “Especificaciones de protección ambiental para la selección del sitio, diseño, construcción, operación, monitoreo, clausura y obras complementarias de un sitio de disposición final de residuos sólidos urbanos y de manejo especial”. En cuanto a los recursos humanos; las rutas se cubren con el trabajo de 26 servidores públicos, de los cuales 11 personas se desempeñan como choferes de transporte recolector y 15 empleados recolectores de RSU, los cuales disponen de 12 camiones recolectores para satisfacer las necesidades de la población municipal, una de las unidades está disponible en forma exclusiva para el área del mercado y los 11 restantes son distribuidos para el servicio habitacional.

Por lo que respecta a la colonia Atlihuahayan, está conformada por 731 familias (2,292 hab.) con un promedio de integrantes de 3.9 (rango: 1 - 9). El resultado de la producción per cápita promedio en la colonia fue de 0.365 Kg/día/hab obtenido en una muestra de 330 habitantes equivalente a 80 familias. Si consideramos que la población total de la colonia es de 2,292 habitantes la producción estimada de RSU al día es de 749.48 Kilogramos como se muestra en la Tabla 1.

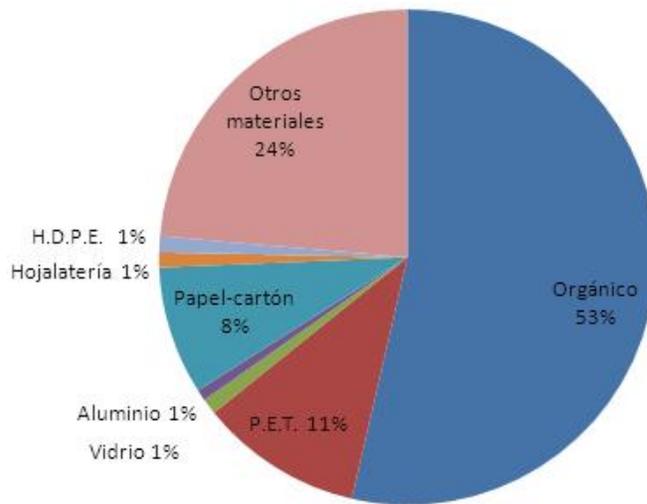


Figura 1: Caracterización de los residuos sólidos estimados de la colonia Atlahuayan, Yautepec, Morelos

**Tabla 1.- Composición de la generación per cápita de residuos en la colonia Atlahuayan, Morelos.**

Subproducto	Ppc. promedio de la muestra (Kg/día/hab) (n=330 hab.)	Estimación de la producción de RSU de la colonia (Kg/día) (N=2,292 hab)
Orgánica	0.191	437.772
P.E.T.	0.038	87.096
Vidrio	0.004	9.168
Aluminio	0.003	6.876
Papel-cartón	0.03	68.78
Hojalatería	0.003	6.876
H.D.P.E	0.004	9.168
Otros (materiales indefinidos)	0.084	192.528
Total	0.356	749.484

En cuanto al tipo de subproductos se encontró: residuos orgánicos en un 53 %, seguido por otros materiales indefinidos no reciclables con un 24 %, el P.E.T con un 11 %, papel-cartón con un 8 %, aluminio con un 1 %, vidrio con un 1 %, H.D.P.E. con un 1 % y hojalatería con un 1 % como se muestra en la Figura 1

Referente al componente cualitativo el presente trabajo reporta 3 categorías de consideración: técnica, social y administrativa del fenómeno en estudio como sigue:

### 1. Caracterización del servicio de limpia del municipio

El sistema de limpia municipal y su personal adscrito en este departamento carece de

herramientas administrativas, para cumplir con su labor y presenta desconocimiento del marco legal e institucional de los residuos sólidos, lo que se ve reflejado en la percepción de los habitantes de la colonia de Atlahuayan, quienes mencionan que es insuficiente el servicio brindado y se ven obligados a sub contratar el servicio de camioneta particular o a efectuar la práctica de quema de basura en su domicilio acto que representa un retroceso cultural.

## **2. Disposición de la basura en la colonia**

Se observa basura acumulada en las calles de la colonia a sí mismo en los patios habitacionales, terrenos baldíos, orillas del río, el único canal de agua que cruza la colonia también presenta RSU, en especial las calles sin pavimentación que se encuentran en los puntos más apartados de la ayudantía municipal. En cuanto a la composición de la basura encontrada en la vía pública se observo residuos de origen orgánico e inorgánico destaca principalmente envolturas de plástico, botellas de P.E.T., vidrio, llantas y animales muertos en descomposición.

## **3. Información de parte de los habitantes en relación al servicio proporcionado por el camión recolector.**

Subjetivamente la información de parte de los habitantes en relación al servicio proporcionado por el municipio, manifiestan que el camión recolector de basura municipal tiene dos días asignados para la colonia de Atlahuayan. Los pobladores exponen su inconformidad debido a que la ruta no cubre las calles no pavimentadas y comunican que las unidades recolectoras se encuentran en condiciones no aptas para desempeñar la importante tarea encomendada. Según la descripción de los puntos anteriores los vecinos se ven obligados a contratar el servicio de recolección a una empresa externa, la cual brinda el servicio cada tercer día a bordo de una camioneta. Sin embargo las personas de menor ingreso no tiene la capacidad de pago para contratar los servicios de recolección privada.

## **Discusión.**

El presente estudio coincide con diferentes estudios efectuados en nuestro país, en la generación per cápita promedio encontrada en nuestro estudio, para un nivel socio-económico medio-bajo, dentro del rango de 0.23 a 0.62 Kg/día/hab. [13]-[15].

De acuerdo a los resultados obtenidos en el estudio RSU efectuado en América Latina y el Caribe (producción per cápita de 0.3 a 0.8 Kg/día/hab.) concuerda con las cifras obtenidas en Atlahuayan y los estudios efectuados dentro de México, coincidiendo con la producción per cápita e índice socioeconómico bajo. [16].

En lo que concierne a la caracterización del material coinciden en la composición respecto a la materia orgánica en el rango de 50.9 % a 63 % [13]-[15].

En el estudio correspondiente a la colonia Atlahuayan, se demostró que el 24 % de la generación de RSU es lo que realmente amerita la disposición final en los tiraderos municipales, de igual manera se observa la similitud con los estudios realizados en las tres investigaciones mencionadas dentro del país[13]-[15].

Las diferencias radican en cuanto a la población de estudio, dado que en dos publicaciones mencionadas [13]-[15]., se investigó a nivel ciudad y la tercera abarcó tres municipios; resaltando que el trabajo realizado en Atlilhuayan a diferencia de otros estudios sólo se indagó la zona habitacional[12]-[15].

## 6. Conclusiones

Actualmente a nivel nacional, no se cuenta con un sistema de gestión integral de los RSU, por lo que existen problemas referentes al tratamiento, recolección y disposición final, además de ello, no se tienen datos actualizados por lo que la información es limitada y escasa. Por otra parte son pocas las ciudades que han realizado diagnósticos (generación y composición) de los RSU, los cuales son muy importantes, ya que cada ciudad, municipio, estado, es diferente debido a los factores culturales, geográficos, socioeconómicos, número de habitantes; por lo que la generación y composición varía entre cada uno de ellos.

Los resultados de este diagnóstico pueden contribuir para crear una fuente de datos actuales que sirva de base para la elaboración de un plan de gestión integral de los RSU. Este estudio solo consideró los RSU de fuentes residenciales, pero también existen los residuos de las fuentes no residenciales (restaurantes, oficinas de gobiernos, mercados o cualquier otro negocio o establecimiento público), los cuales también son grandes generadores de RSU por otra parte dentro de las diferentes percepciones están la de los usuarios que requieren la satisfacción de la necesidad del servicio de recolección de residuos haciéndose manifiesto la urgente necesidad de contar con personal capacitado en la materia a efecto de planear organizar y ejecutar un recorrido integral en la comunidad para cumplir con este cometido, sobretodo contar con un manual de organización, bitácoras de registro de servicios planeados y cumplidos al igual que la adquisición de nuevas unidades con lo cual evitarían incumplir el servicio y manteamiento menos costoso.

La generación per cápita estimada en la colonia resultó por debajo de lo que se reporta a nivel nacional, que es de 1.25 kg /día/hab. [3], con la aclaración que se desconoce si dicho dato incluye la generación de los RSU de las fuentes no residenciales. Relativo a la composición obtenida de la separación de los RSU podemos decir que los residuos sujetos a tratamiento, en este estudio se obtuvo el dato referente al 53 % corresponde a residuos que son tratables, es decir biodegradables los cuales se pueden llegar a someter a composta o lombricomposta para obtener tierra de uso de jardinería. Por otra parte la generación del 23 % de residuos son reciclable, los cuales pueden ser destinados a la venta en empresas recicladoras con los cuales a su vez se obtendría un beneficio económico y ecológico, debido que se aprovecharía al máximo los materiales utilizados como materia prima para producir nuevos productos. Si se llegara a separar estos subproductos, únicamente se depositarían al relleno sanitario el 24 %, lo cual tendría beneficios económicos, debido a que actualmente se tiene un costo de 30 a 640 pesos más impuesto por disponer una tonelada de RSU en el relleno sanitario [17]. Los factores que marcan la diferencia de costo están en función de la densidad poblacional, la cantidad recolectada, la eficiencia en el llenado del vehículo, el estado físico de éstos y el diseño de las rutas para dar un servicio integral y contribuir al mejoramiento del planeta, pues la basura no solo contamina la tierra, sino también el aire y agua repercutiendo en la salud de la comunidad, el deterioro del medio ambiente y la imagen del paisaje.

## Referencias

- [1] Mora J. El problema de la basura en la ciudad de México. [monografía en internet]. México D.F.; 2004 [consultado 2012 junio]. Disponible en: [http : //www.paot.org.mx/paot\\_docs/pdf/basura\\_df.pdf](http://www.paot.org.mx/paot_docs/pdf/basura_df.pdf)
- [2] SEMARNAT. Indicadores Básicos del Desempeño Ambiental de México [monografía en internet]. México 2005 [consultado 2012 junio]. Disponible en: [http : //www.semarnat.gob.mx/informacionambiental/snria/Documents/snria\\_2009/pdf/Indicadores\\_basicos\\_2005.pdf](http://www.semarnat.gob.mx/informacionambiental/snria/Documents/snria_2009/pdf/Indicadores_basicos_2005.pdf).
- [3] Instituto Nacional de Ecología, Estadísticas e indicadores de inversión sobre residuos sólidos municipales en los principales centros urbanos de México. [monografía en internet]. México 1997 [consultado 2012 julio ]. Disponible en: [http : //www2.ine.gob.mx/publicaciones/libros/109/cap8.html](http://www2.ine.gob.mx/publicaciones/libros/109/cap8.html)
- [4] CEAMA. [sitio en internet]. Manejo Integral de los residuos [actualizado 2010; consultado 2012 junio ]. Disponible en: [http : //www.ceamamorelos.gob.mx/secciones/ambiente/residuos\\_solidos.html](http://www.ceamamorelos.gob.mx/secciones/ambiente/residuos_solidos.html).
- [5] Arenas Monreal L, Sosa N, et al. Experiencias de acercamiento comunitario y participación de la población en salud. INSP, Indesol, Sedesol. Cuernavaca. 2008
- [6] Organización Panamericana de la Salud. Salud Ambiental de lo global a lo local. 1ra. Edición: Mc Graw-Hill Interamericana editores, Washington, D.C. 2010.
- [7] Bonfanti, F. A. La incorrecta gestión de los residuos sólidos urbanos y su incidencia en la calidad de vida de la población de Resistencia. UNNE. Comunicaciones Científicas y Tecnológicas. Argentina: 2004.
- [8] Ocampo C, Padilla A, Méndez F. Impacto de un depósito de residuos sólidos en el crecimiento físico infantil. Columbia medica vol. 39 No. 3, 2008.
- [9] Duane J. Gubler. Resurgent Vector-Borne Diseases as a Global Health Problem. Emerging Infectious Diseases, Vol. 4, No. 3, 1998
- [10] Secretaria de salud. Programa de acción específico de 2007-2012 Dengue [monografía en internet]. México 2007 [2012 marzo ]. Disponible en: [http : //www.cenave.gob.mx/progaccion/dengue.pdf](http://www.cenave.gob.mx/progaccion/dengue.pdf)
- [11] Velásquez Monroy O. Los segadores de basura y los factores de riesgo a la salud [monografía de internet]. Secretaria de salud de México, 2000 [consultado 27 agosto 2012]. Disponible en: [http : //www.bvsde.paho.org/bvsacd/acodal/vii.pdf](http://www.bvsde.paho.org/bvsacd/acodal/vii.pdf)
- [12] Organización Panamericana de la Salud. Manual de evaluación de daños y necesidades de salud para sitios de desastres. Washington, D.C.: OPS, 2004
- [13] Buenrostro O, Bocco G, Bernache G. Urban solid wasted generation and disposal in Mexico a case of study. Wastemanagmen and research 2001:19: 169-179.

- [14] Esquinca F, Escobar J, Hernández A, Villalobos J. Caracterización y generación de los residuos sólidos de Tuxtla Gutiérrez, CHIAPAS.[En línea].México 1994 [2011 julio ].Disponible en: *http : //www.bvsde.paho.org/bvsaidis/resisoli/mexico/03364e14.pdf*
- [15] Gobierno del estado de México secretaria de ecología dirección general de normatividad y apoyo técnico. Estudio de generación y caracterización de residuos sólidos municipales. México, 2000.
- [16] Banco interamericano de desarrollo, Organización panamericana. Diagnóstico de la situación del manejo de residuos sólidos municipales en América Latina y el Caribe. Washintong, D.C. 1997
- [17] Instituto Nacional de Ecología. [sitio en internet] Residuos Sólidos Urbanos [actualizado 2007 Noviembre consultado 2012 Noviembre].Disponible en: *http : //www2.ine.gob.mx/publicaciones/libros/495/residuos.html*.

# El modelo de regresión logística: una aplicación al problema de hipertensión arterial, sobrepeso u obesidad de personas que acuden a los Centros de Salud de Chilpancingo, Guerrero.

Moreno Sandoval Ma. Isabel<sup>1</sup>

Díaz González Lucio<sup>2</sup>

Covarrubias Melgar Dante<sup>3</sup>

Sistachs Vega Vivian del Rosario<sup>4</sup>

<sup>1</sup>Unidad Académica de Matemáticas-UAG.

<sup>2</sup>Universidad Autónoma de Guerrero.

<sup>3</sup>Universidad Autónoma de Guerrero.

<sup>4</sup>Universidad de la Habana.

*Ave. Lázaro Cárdenas S/N , colonia la haciendita, C.U,  
Chilpancingo, Guerrero, 39000, México  
isabe.mor.san@gmail.com, luciodiaz@uagro.com  
dcova@uagro.com, vivian@matcom.uh.cu*

## RESUMEN

La regresión logística es una de las herramientas estadísticas con mejor capacidad para el análisis de datos en investigación clínica y epidemiología. El objetivo primordial es modelar cómo influye en la probabilidad de aparición de un suceso, habitualmente dicotómico, la presencia o no de diversos factores y el valor o nivel de los mismos. Uno de los problemas que se presenta en la práctica dentro de las diferentes ramas del quehacer científico y técnico es el de ajustar un modelo a un conjunto de observaciones. El ajuste de un modelo puede tener diferentes objetivos, como puede ser: conocer simplemente la relación existente entre las variables, estimar ciertos parámetros, predecir y comparar grupos de datos. En este trabajo se presentan los resultados de un estudio realizado a los habitantes que acudían a los diferentes Centros de Salud de la Ciudad de Chilpancingo Guerrero. Con el fin de valorar los factores de riesgos asociados a la Hipertensión Arterial, Sobrepeso u Obesidad, esto a partir de una serie de factores que influyen en dichos padecimientos.

## 1. INTRODUCCIÓN

Los niveles de sobrepeso y obesidad han aumentado progresivamente en las últimas décadas tanto en países desarrollados como en muchos otros en vías de desarrollo. Se estima que más de mil millones de adultos sufren de sobrepeso y por lo menos 300 millones son obesos. En América Latina también se está produciendo un rápido y alarmante incremento en la prevalencia de obesidad, debido principalmente a factores tales como el crecimiento económico, la urbanización progresiva y el consiguiente cambio en los estilos de vida que involucra un mayor consumo de alimentos de alta densidad energética, el abandono de dietas tradicionales y la disminución generalizada de los niveles de actividad física.[1]

La epidemia de sobrepeso y obesidad es preocupante debido a que existe abundante evidencia que indica que el peso excesivo aumenta el riesgo de padecer diversas patologías que alteran significativamente la calidad de vida, entre éstas se encuentran las enfermedades cardiovasculares, **hipertensión arterial**, dislipidemia, insuficiencia cardíaca, resistencia a la insulina, diabetes Mellitus del tipo 2, problemas respiratorios, apnea del sueño, enfermedad vesicular, osteoartritis y diversos tipos de cáncer. Un nivel de presión arterial de 140/90 mmHg o más es considerado hipertensión para la mayoría de adultos.[2]

Por esta razón se estudia la regresión logística ya que es útil en los casos en los que se desea predecir la presencia o ausencia de una determinada característica según un conjunto de variables predictoras. De tal manera que la ausencia o presencia de esa característica se denotará mediante el valor uno (presencia) o el cero (ausencia), es decir, la variable dependiente será dicotómica.

Son muchos y variados los fenómenos que se presentan en la vida de los cuales se desconoce su naturaleza, haciéndose imposible utilizar métodos matemáticos determinísticos para su estudio, debido a que en estos están presentes factores aleatorios. Por ello los métodos estadísticos han resultado tan importantes en el desarrollo de muchas ciencias como, la Química, la Agronomía, la Medicina entre otras, pues ellas brindan al especialista una explicación acerca del fenómeno en estudio. El objetivo de la Regresión Logística es encontrar el mejor ajuste del modelo con el menor número de parámetros y describir la relación entre la variable respuesta y un conjunto de variables (covariables) explicativas independientes. En particular, la Regresión Logística es uno de los métodos estadísticos más expresivos y versátiles para el análisis de datos. Muchos especialistas de diferentes ramas han trabajado con la Regresión Logística para predecir o pronosticar el comportamiento de una variable respuesta binaria o dicotómica, en un modelo donde las variables independientes pueden ser de cualquier naturaleza, convirtiéndola en un método estándar para el análisis de regresión cuando los datos son binarios (Hosmer y Lemeshow, 1989).[8]

## 2. Modelo de regresión logística

Este método de regresión aporta en muchos casos una estrategia bastante satisfactoria cuando las distribuciones consideradas no son normales, e incluso cuando algunas de las variables son discretas o cualitativas. Por lo cual es ampliamente utilizado tanto en estudios observacionales como de encuestas y experimentales.

La regresión logística, al igual que otras técnicas estadísticas multivariadas, da la posibilidad de evaluar la influencia de cada una de las variables independientes sobre la variable dependiente y controla el efecto del resto. Estos modelos de regresión logística son modelos de regresión que permiten estudiar si una variable categórica depende, o no, de otra u otras variables.

La Regresión Logística es uno de los métodos estadísticos más expresivos y versátiles para el análisis de datos. Muchos especialistas de diferentes ramas han trabajado con esta técnica para predecir o pronosticar el comportamiento de una variable respuesta binaria o dicotómica, en un modelo donde las variables independientes pueden ser de cualquier naturaleza, convirtiéndola en un método estándar para el análisis de regresión cuando los datos son binarios. Por lo cual el método es efectivo para expresar la probabilidad de que ocurra el hecho en cuestión de una función de ciertas variables independientes que se presumen relevantes o influyentes. Estas variables independientes reciben el nombre de covariables, pueden ser dicotómicas, nominales, ordinales y continuas. Mientras que la variable respuesta o variable dependiente, se codifica en 0 y 1, donde 1 la ocurrencia del evento de interés y 0 significa ausencia.

Su objetivo principal de la regresión logística es el de modelar cómo influye en la probabilidad la presencia de un suceso, habitualmente dicotómico, y la presencia o no de diversos factores y el valor o nivel de los mismos.

El modelo de regresión logística tiene una doble función tanto **Explicativas** como **Predictivas**:

**Explicativas:** Su objetivo es evaluar cómo afectan el cambio en unas características determinadas (variables independientes) sobre otra característica en concreto (variable dependiente).

**Predictivas:** Su objetivo es estimar o aproximar el valor de una característica (variable dependiente) en función de los valores que pueden tomar en conjunto otra serie de características (variables independientes).

El modelo de regresión logística es representado de la siguiente manera:

Supongamos que la probabilidad  $p$  depende de los valores de ciertas variables  $X_1, \dots, X_p$ . Es decir, si  $\mathbf{x} = (x_1, \dots, x_p)'$  son las observaciones de un cierto individuo  $m$  sobre las variables, entonces la probabilidad de que ocurra  $A$  dado  $\mathbf{x}$  es  $p(y=1|\mathbf{x})$ . Indicaremos esta probabilidad por  $p(\mathbf{x})$ . La probabilidad contraria de que  $A$  no suceda dado  $\mathbf{x}$  será  $p(y=0|\mathbf{x}) = 1 - p(\mathbf{x})$ . Donde  $p(\mathbf{x})$  está comprendido entre 0 y 1. [10]

Por diversas razones es muy conveniente suponer un modelo lineal para la llamada transformación logística de la probabilidad.

$$\ln\left[\frac{p(\mathbf{x})}{1-p(\mathbf{x})}\right] = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p = \beta_0 + \beta' \mathbf{x} \quad (1)$$

La expresión anterior es equivalente a esta otra, teniendo como,  $\beta = (\beta_1, \dots, \beta_p)'$  parámetros de regresión:

$$\frac{p(\mathbf{x})}{1-p(\mathbf{x})} = e^{\beta_0 + \beta' \mathbf{x}} \quad (2)$$

despejando  $p(\mathbf{x})$  obtenemos otra forma de escribir el modelo logístico

$$p(\mathbf{x}) = \frac{e^{\beta_0 + \beta' \mathbf{x}}}{1 + e^{\beta_0 + \beta' \mathbf{x}}} \quad (3)$$

### INTERPRETACIÓN DE LOS PARÁMETROS

Para poder interpretar el modelo es necesario conocer algunas medidas de asociación, las cuales permiten dar una interpretación adecuada a los resultados. Para expresar la relación entre dos variables que únicamente pueden optar entre dos opciones de respuesta frecuentemente se recurre a medidas relativas como el odds ratio o el riesgo relativo.

- **Riesgo Relativo (RR):** Expresa el coeficiente de probabilidades del número de sujetos que desarrollan la enfermedad en estudio entre el total de sujetos que presentan un determinado factor dividido por el número de sujetos que desarrollan la enfermedad entre el total de sujetos que no presentan. Toma valores comprendidos entre  $[0, \infty]$ , distribuyéndose asintóticamente. El riesgo relativo de alguna manera nos indica cuántas veces más de riesgo tienen de contraer alguna enfermedad en específico.

El valor de RR igual a 1 indica que no hay asociación positiva, es decir, la presencia del factor se asocia a mayor ocurrencia del evento. Un  $RR > 1$  indica una asociación positiva, lo cual es, la presencia del factor se asocia a mayor ocurrencia del evento. Un  $RR < 1$ , implica que la asociación es negativa.

- **Odds ratio (OR):** Indica cuanto más probable es la ocurrencia del evento que su no ocurrencia.

Un valor de  $OR > 1$  indica una asociación positiva, es decir, la presencia del factor se asocia a mayor ocurrencia del evento. Un valor de  $OR < 1$  implica que la asociación es negativa.[6]

### *ESTIMACIÓN DE LOS PARÁMETROS*

Para la estimación de los coeficientes del modelo y de sus errores estándar se recurre al cálculo de estimaciones de máxima verosimilitud, es decir, estimaciones que hagan máxima la probabilidad de obtener los valores de la variable dependiente  $Y$  proporcionados por los datos de la muestra. Para el cálculo de estimaciones máximo-verosímiles se recurre a métodos iterativos, como el método de Newton-Raphson. Los estimadores de máxima verosimilitud tienen, bajo ciertas condiciones, distribución asintótica normal de media el valor poblacional del parámetro estimado y matriz de covarianzas dada por la inversa de la matriz de información de Fisher. Así, son asintóticamente insesgados y además se puede hacer inferencia sobre ellos basándose en dicha distribución normal cuando el tamaño muestral es suficientemente grande. Estos resultados se deben a Wald. El siguiente paso será comprobar la significación estadística de cada uno de los coeficientes de regresión en el modelo. Para ello se pueden emplear básicamente tres métodos: el estadístico de Wald, el estadístico  $G$  de razón de verosimilitud y la prueba Score.[5]

### *CONTRASTE DE LOS PARÁMETROS*

**El estadístico de Wald:** Contrasta la hipótesis de que un coeficiente aislado es distinto de 0, y sigue una distribución normal de media 0 y varianza 1. Su valor para un coeficiente concreto viene dado por el cociente entre el valor del coeficiente y su correspondiente error estándar. La obtención de significación indica que dicho coeficiente es diferente de 0 y merece la pena su conservación en el modelo. En modelos con errores estándar grandes, el estadístico de Wald puede proporcionar falsas ausencias de significación (es decir, se incrementa el error tipo II). Tampoco es recomendable su uso si se están empleando variables de diseño.

**El estadístico G de razón de verosimilitud:** Se trata de ir contrastando cada modelo que surge de eliminar de forma aislada cada una de las covariables frente al modelo completo. En este caso cada estadístico G sigue una  $X^2$  con un grado de libertad (no se asume normalidad). La ausencia de significación implica que el modelo sin la covariable no empeora respecto al modelo completo (es decir, da igual su presencia o su ausencia), por lo que según la estrategia de obtención del modelo más reducido (principio de parsimonia), dicha covariable debe ser eliminada del modelo ya que no aporta nada al mismo. Esta prueba no asume ninguna distribución concreta, por lo que es la más recomendada para estudiar la significación de los coeficientes.

**El estadístico test de Score:** Está basado en la distribución de las derivadas parciales de la log-verosimilitud, reduce los cálculos con respecto a los anteriores contrastes.

### *BONDAD Y AJUSTE*

Siempre que se construye un modelo de regresión es fundamental, antes de pasar a extraer conclusiones, el corroborar que el modelo calculado se ajusta efectivamente a los datos usados para estimarlo. En el caso de la regresión logística una idea bastante intuitiva es calcular la probabilidad de aparición del suceso. Si el ajuste es bueno, es de esperar que un valor alto de probabilidad se asocie al suceso, y viceversa, si el valor de esa probabilidad calculada es bajo, cabe esperar también ausencia del suceso. Esta idea intuitiva se lleva a cabo formalmente mediante la prueba conocida como de Hosmer–Lemeshow (1989), que básicamente consiste en dividir el recorrido de la probabilidad en deciles de riesgo. Ambas distribuciones, esperada y observada, se contrastan mediante una prueba de  $X^2$ .

**PROBLEMA 2.1** *Se estudió una población de individuos que acudían a los diferentes centros de salud de la ciudad de Chilpancingo, lo cual se quiere determinar los factores de riesgo asociados a la hipertensión arterial.*

*Definimos la variable Y a los individuos que padecen hipertensión arterial:*

*1, Si padece hipertensión arterial.*

*0, No padece hipertensión arterial.*

*Admitamos también que se consideran  $k=10$  variables posiblemente influyentes en el padecimiento de hipertensión arterial. Se muestran en la siguiente tabla:*

Al analizar dichas variables se obtuvieron los siguientes resultados.

Las variables que están seleccionadas son las que salieron significativas para el modelo, las cuales son, la edad y antecedentes de hipertensión.

Nombre de la Variable	Descripción
V1	Edad
V2	Sexo
V3	Estado civil
V4	Ocupación
V5	Peso
V6	Antecedentes de hipertensión
V7	Se estresa en su trabajo
V8	Practica ejercicio
V9	Fuma tabaco
V10	Consume alcohol

Cuadro 1: Variables

Obsevado	Sin HTA	Con HTA	%
<b>Sin HTA</b>	483	8	98.4
<b>Con HTA</b>	126	11	8.0
% Global			78.7

Cuadro 2: Clasificación

V	B	Wald	Sig
V1	.035	26.437	.000
V2	.193	.398	.528
V3	.011	.032	.858
V4	-.005	.016	.900
V5	-.007	.703	.402
V6	-.525	6.178	.013
V7	-.056	.104	.747
V8	.283	2.501	.114
V9	.071	.064	.801
V10	.186	.599	.439
Constante	-2.444	1.068	.022

Cuadro 3: Variables en la ecuación

$$P(Y = 1) = \frac{1}{1 + \exp(2.444 - .035(X_1) + .525(X_6))} \quad (4)$$

Si estamos ante un sujeto de 50 años y tiene antecedentes hereditarios de hipertensión, su probabilidad de padecer hipertension arterial es:

$$P(Y = 1) = \frac{1}{1 + \exp(2.444 - .035(50) + .525(1))} = 0.228 \quad (5)$$

Esto quiere decir que se estima que, aproximadamente, al 23 porciento de los sujetos que posean las características mencionadas se les presentará dicho padecimiento.

*PROBLEMA 2.2 Se estudió una población de individuos que acudían a los diferentes centros de salud de la ciudad de Chilpancingo, lo cual se quiere determinar los factores de riesgo asociados al sobrepeso u obesidad.*

*Definimos la variable Y a los individuos que padecen sobrepeso u obesidad:*

*1, Si padecen sobrepeso u obesidad.*

*0, No lo padecen.*

*Admitamos también que se consideran k= 10 variables posiblemente influyentes en el padecimiento de sobrepeso u obesidad. Se muestran en la siguiente tabla:*

Al analizar las variables se obtuvieron los siguientes resultados:

Las variables que estan seleccionadas son las que salieron significativas para el modelo, las cuales son, la edad y el estado civil.

$$P(Y = 1) = \frac{1}{1 + \exp(.754 - .019(X_1) + .107(X_3))} \quad (6)$$

Si estamos ante un sujeto de 50 años y estado civil casado, su probabilidad de padecer sobrepeso u obesidad es el siguiente:

$$P(Y = 1) = \frac{1}{1 + \exp(.754 - .019(50) + .107(1))} = 0.522 \quad (7)$$

Nombre de la variable	Descripción
V1	Edad
V2	Sexo
V3	Estado civil
V4	Ocupación
V5	Frecuencia en la que se estresa
V6	Frecuencia que realiza ejercicio
V7	Le agraga sal a sus alimentos
V8	Frecuencia con que fuma
V9	Frecuencia con que consume alcohol
V10	Consumo alcohol

Cuadro 4: Variables

Observado	Con	Sin	%
Con	178	133	57.2
Sin	118	199	62.8
% Global			60.0

Cuadro 5: Clasificación

V	B	Wald	Sig
V1	.019	10.192	.001
V2	.402	2.872	.090
V3	-.107	.051	.035
V4	-.042	1.649	.199
V5	-.020	.116	.734
V6	.077	2.356	.125
V7	-.164	1.410	.235
V8	-.027	.092	.761
V9	.023	.032	.858
Constante	-.754	.765	.382

Cuadro 6: Variables en la ecuación

Esto quiere decir que se estima que, aproximadamente, al 52.2 por ciento de los sujetos que posean las características mencionadas se les presentará dicho padecimiento.

### 3. CONCLUSIONES

De acuerdo a los resultados obtenidos se concluye que los factores de riesgo para padecer hipertensión arterial son, la edad y antecedentes hereditarios. Lo que significa que la edad es un factor de riesgo para el problema de hipertensión arterial y los antecedentes hereditarios son un factor de protección. Para el padecimiento de sobrepeso y obesidad los factores de riesgo son la edad y el estado civil donde, la variable edad es un factor de riesgo para padecer sobrepeso u obesidad y la variable estado civil es factor de protección.

### Referencias

1. World Health Organization, Obesity: Preventing and Managing the Global Epidemic. Report of a WHO consultation of Obesity. Geneva, 2000.
2. World Health Organization. The World Health Report. Reducing Risks, Promoting Healthy Life. Geneva: WHO; 2002.
3. D. Wayne, Bioestadística: Base para el análisis de las ciencias de la salud. Cuarta edición: Limusa, 1997.
4. D. Montgomery, Introducción al Análisis de Regresión Lineal. Tercera edición. México: Continental, 2006.
5. A. Salazar, Modelos de respuesta discreta en R y su aplicación con datos reales. Primera edición. Universidad de Granada.
6. E. Sanchez, Regresión Logística en Salud Pública. Primera edición. España: EASP, 2000.
7. F. BARÓN, Apuntes de Bioestadística. Primera edición. Universidad de Málaga, 2004.
8. D. Hosmer, Logistic Regression. Segunda edición. Canadá: John Wiley & Sons, 2000.
9. J. Bernardo, Bioestadística. Primera edición. España: Vicens-Vives, 1981.
10. J. Montero, Análisis Multivariante. Primera edición. España: Universidad de Extremadura, 2008.

# Evaluación del riesgo por contaminantes tóxicos en la Ciudad de México

García Reynoso, J. A.<sup>1</sup> y Martínez Catana, N.

<sup>1</sup> UNAM, Centro de Ciencias de la Atmósfera.

## Resumen

En este trabajo se presenta una evaluación preliminar de riesgos a la salud humana por efectos cancerígenos y por efectos no cancerígenos tanto a agudos como crónicos por inhalación de tóxicos atmosféricos presentes en la atmósfera de la Ciudad de México. Los compuestos estudiados son los conocidos como BTXE que conforman el Benceno, Tolueno, Xilenos y Etil Benceno. La evaluación abarca los meses de junio a diciembre de 2011. Estos contaminantes son medidos por la red automática de monitoreo atmosférico del gobierno de la Ciudad de México y se emplea la metodología de la agencia de Protección Ambiental de América del Norte.

De los resultado se encontró que para los efectos no cancerogénicos los niveles de concentración ambiental de los BTXE no requieren de un estudio más detallado, caso contrario para los efectos cancerogénicos, donde si se tienen niveles de riesgo no aceptables.

**Palabras clave:** Riesgo, BTEX o (benceno, tolueno, xileno, etilbenceno), Ciudad de México.

## 1. Introducción

Cuando se habla de compuestos tóxicos del aire se refiere a un grupo de compuestos químicos nocivos para la salud. En esta categoría de contaminantes se agrupan a muchas sustancias con efectos diversos, los contaminantes tóxicos que representan un problema a la salud en la Ciudad de México son: Benceno, Tolueno, Xilenos y Etilbenceno o mejor conocidos como BTXE los cuales son reportados por el Sistema de Monitoreo Atmosférico de la Ciudad de México, estos pueden causar cáncer y otros efectos a la salud como irritación de las vías respiratoria, problemas en el sistema nervioso, defectos congénitos, daños al sistema inmune y también respiratorio. Las personas se pueden exponer por diferentes rutas en al aire por inhalación, absorción de piel, al comer y beber agua, este trabajo se enfoca a la exposición vía inhalación

El problema de la calidad del aire en la Ciudad de México aún no se ha resuelto a causa del crecimiento demográfico, de la concentración industrial y el incremento de vehículos ya sea del transporte público y privado, esto ha llevado a provocar el deterioro ambiental en la región.

Los efectos agudos a la salud ocurren poco después que una persona inhala un contaminante tóxico del aire. Dentro de los contaminantes ambientales podemos encontrar los contaminantes criterio como el ozono, dióxido de azufre, monóxido de carbono, dióxido

de nitrógeno, partículas y plomo. Sin embargo también en la atmósfera se encuentran otros compuestos conocidos como tóxicos atmosféricos que son los que han contribuido al aumento de morbilidad y mortalidad.

El presente trabajo, tiene como objetivo realizar una evaluación preliminar de riesgo calculando cuantitativamente el riesgo a presentar cáncer, así como a efectos no carcinogénicos y un análisis comparativo de los mismos.

## 2. Antecedentes

Los contaminantes tóxicos del aire pueden darse de manera natural o provocados por el hombre de forma natural como incendios de bosques, erupciones volcánicas, mientras las provocadas por el hombre son por fabricación de productos químicos, fabricación de pesticidas, incineradores, hornos etc. Se considera riesgo la probabilidad de ocurrencia de consecuencias adversas, indeseables a la vida humana, salud, propiedad o al ambiente. [1]

Los efectos agudos pueden manifestarse en cuestión de minutos, horas o días, mientras que los crónicos aparecen solamente en semanas, meses o años después del contacto con la sustancia tóxica [2]. Riesgos agudos/ Riesgos crónicos de los contaminantes tóxicos atmosféricos para el Benceno, Etilbenceno, Tolueno y Xilenos son:

### 2.1. Benceno

Riesgos agudos /Riesgos crónicos: El Benceno puede ser muy tóxico por ingestión, inhalación y por el contacto de la piel, es irritante en la piel, ojos, nariz y la garganta. Cuando se calienta hasta que se evapora emite humos tóxicos de monóxido de carbono y dióxido de carbono, también es un precursor de cáncer. Síntomas: Los síntomas de exposición pueden incluir mareo, vómito, pérdida de conocimiento, depresión en el sistema nervioso central, mareo, cefalea, constricción de pecho y euforia. Puede causar somnolencia, pulso rápido, delirio, neumonía química, falta de respiración, trastornos visuales, fatiga, vértigo, disnea, la exposición por inhalación provoca disfunción en el sistema nervioso central.

### 2.2. Etilbenceno

Riesgos agudos / Riesgos crónicos: Etilbenceno es un irritante, puede ser tóxico por ingestión, inhalación y absorción de la piel. En altas concentraciones puede ser narcótico y precursor del cáncer. Síntomas: Los síntomas de exposición pueden incluir irritación, enrojecimiento e inflamación de la piel, irritación de la nariz, garganta y ojos, puede provocar conjuntivitis, puede provocar lesiones en la cornea, dermatitis, mareos, narcosis y sensación de obstrucción de pecho.

### 2.3. Tolueno

Riesgos agudos / Riesgos crónicos: El Tolueno puede ser tóxico e irritante. Puede ser nocivo por inhalación, ingestión o absorción de piel. Es fácilmente absorbido por la piel. Cuando se calienta hasta su descomposición emite gases tóxicos de dióxido de carbono y monóxido de carbono. Síntomas: Los síntomas de exposición incluyen irritación de ojos, dilatación de las pupilas, fatiga ligera, trastornos oculares, enrojecimiento de

la visión. También causa mareo, cefalea, náuseas, confusión mental, enfermedad en la sangre, dermatitis. Otros síntomas como alucinaciones o distorsiones, narcosis en altas concentraciones, cambios de actividad motora.

## 2.4. Xilenos (mezclas)

Riesgos agudos / Riesgos crónicos: Es un irritante y puede ser tóxico si se ingiere. Cuando se calienta hasta la descomposición, esta mezcla emite vapores tóxicos de monóxido de carbono, dióxido de carbono y puede ser un narcótico en altas concentraciones. Síntomas: Los síntomas de exposición a los Xilenos pueden incluir irritación en los ojos, nariz y garganta, sequedad y pérdida de grasa de la piel que puede conducir a la dermatitis, neumonitis química, edema pulmonar, hemorragia, depresión del sistema nervioso central, mareo, inconsciencia, anorexia, náuseas, vómitos y dolor abdominal, puede ser narcótico en altas concentraciones. También causa cefalea, fatiga, cansancio, irritabilidad y trastornos gastrointestinales, puede causar daño ocular reversible, sensación en las membranas mucosas salivación, vómito con sangre, trastornos de la coordinación motora, trastornos del habla, temblores, respiración superficial, irregularidades ventriculares, parálisis y convulsiones.

## 3. Desarrollo del trabajo.

Se obtuvieron las concentraciones medidas en la atmósfera de la Ciudad de México de los tóxicos a estudiar, las mediciones comprenden el año 2011, éstas son realizadas por la Dirección de Monitoreo Atmosférico del Distrito Federal, los datos se obtuvieron del portal del Sistema de Monitoreo Atmosférico de la Ciudad de México [3]. Los contaminantes tóxicos obtenidos fueron: Benceno, Tolueno, Xileno y Etilbenceno, son cinco las estaciones donde se realiza el monitoreo que son Merced, FES Acatlán, San Agustín, UAM Iztapalapa y Pedregal. La ubicación de las estaciones se muestra en la figura 1

### 3.1. Métricas de Riesgo

Las métricas utilizadas para la evaluación de riesgo comprenden el índice de riesgo (IR) tanto para exposición aguda como crónica y el incremento en la probabilidad de cáncer en el tiempo de vida [4].

El índice de riesgo se calcula como la relación de la concentración medida con la concentración de referencia como se muestra en la ecuación 1

$$IR = \frac{C_m}{C_{ref}} \quad (1)$$

donde: IR es el índice de riesgo,  $C_m$  es la concentración medida y  $C_{ref}$  es la concentración de referencia, si se emplea el valor de referencia agudo se obtiene el IR agudo y si es el crónico se obtiene el índice crónico.

El incremento en la probabilidad de tener cáncer en el tiempo en el tiempo de vida (LCP) se estima mediante el uso del “factor de riesgo unitario”, (URF), que representa un factor de conversión de la concentración exterior a LCP y representa el número de casos probables a contraer cáncer en 1,000,000 de habitantes. La forma como se calcula este valor se presenta en la ecuación 2

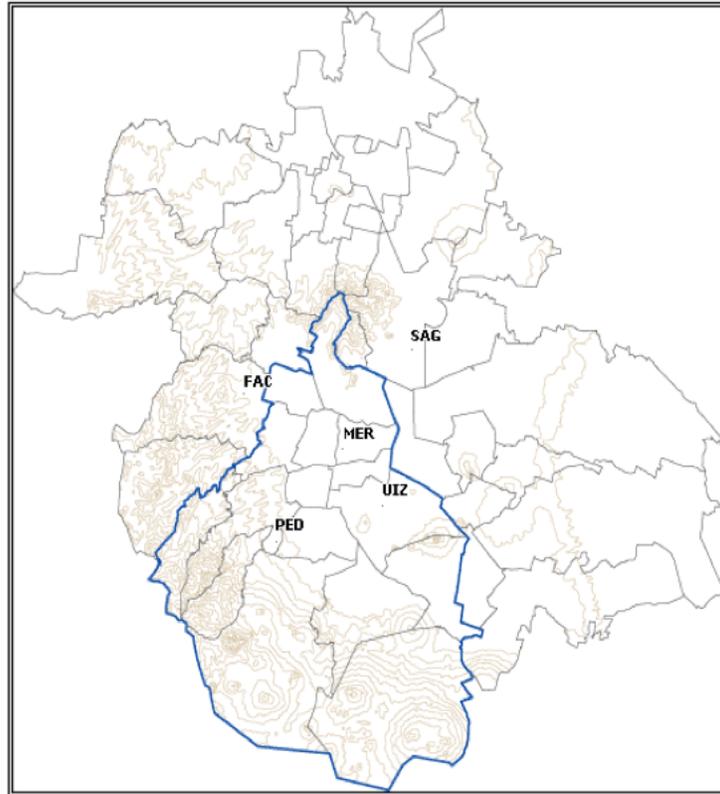


Figura 1: Estaciones donde se realiza el muestreo de BTEX.

$$LCP = URF \cdot C_m \quad (2)$$

El incremento anual teórico en los casos de cáncer (IA) es la incidencia durante el tiempo de vida dividida por el tiempo promedio de vida de un adulto, es decir el número de casos que se incrementan por año y se puede estimar mediante la ecuación 3

$$IA = \frac{\text{Incidencia en el Tiempo de Vida}}{\text{Esperanza de Vida}} \quad (3)$$

La incidencia refleja en número de casos nuevos en un período de tiempo.

Una forma de expresar lo resultado y que sean más fácil de interpretar es mediante la pérdida de esperanza de vida (PEV) que representa el numero de días que se pierden por estar expuesto a un riesgo. Así tenemos que la PEV se puede calcular a partir del riesgo con la ecuación 4.

$$PEV = 1.1 \cdot 10^6 \cdot LCP \quad (4)$$

para  $LCP < 10^{-3}/\text{año}$ . Este se puede calcular para diversas causas de muerte que varían desde enfermedades hasta accidentes.

### 3.2. Datos de referencia

En el cuadro 1 se presentan los valores de concentración de referencia para evaluar los efectos agudos y crónicos, así como también el factor de riesgo unitario.

Cuadro 1: Concentraciones de referencia para niveles agudos y crónicos, y el factor de riesgo unitario

Tóxico	Concentración aguda ( $\mu\text{g}/\text{m}^3$ )	Concentración crónica ( $\mu\text{g}/\text{m}^3$ )	URF ( $(\mu\text{g}/\text{m}^3)^{-1}$ )
Benceno	1.3E+03	6.00E+01	2.9E-05
Tolueno	3.70E+04	3.00E+02	
Xilenos	2.20E+04	7.00E+02	
Etilbenceno		2.00E+02	2.50E-06

Los valores obtenidos de concentración de referencia y el factor de riesgo unitario son los reportados por la oficina de evaluación de riesgos ambientales a la Salud del estado de California de los Estados Unidos [5].

De acuerdo a la información del Instituto Nacional de Estadística y Geografía en el año 2011 la esperanza de vida se refiere al número de años que en promedio se espera que viva una persona después de nacer [6], en el Distrito Federal esperanza de vida de hombres y mujeres es de 76.4 años.

## 4. Resultados

Se presentan las concentraciones de los contaminantes tóxicos emitidos a la atmósfera en la Ciudad México del Sistema de Monitoreo Atmosférico, durante los meses de junio a diciembre de 2011, la caracterización de los riesgos en cancerígenos y no cancerígenos, y incremento anual teórico en los casos de cáncer.

Las concentraciones promedio durante el segundo semestre del 2011 se presentan en el cuadro 2

Cuadro 2: Concentraciones de BTEX en  $\mu\text{g}/\text{m}^3$

	Benceno	Etil Benceno	Tolueno	Xilenos
Merced	4.7	3.0	30.5	9.2
FES Acatán	2.4	1.7	10.8	4.5
San Agustín	2.9	3.6	16.1	7.1
UAM Iztapalapa	3.6	5.4	21.6	12.9
Pedregal	2.8	1.8	16.9	4.7

Para el caso del riesgo no cancerogénico (IR crónico y agudo) se obtuvieron valores menores a la unidad para todos los compuestos., por lo que no se presentan. Sin embargo

para el caso de los efectos cancerogénicos se obtuvieron valores mayores a 1 caso por millón. En el cuadro 3 se muestran los valores de riesgo en casos por millón, la pérdida de esperanza de vida y el incremento anual teórico en los casos de cáncer (IA).

Cuadro 3: Riesgo en casos por millón (R), la pérdida de esperanza de vida (PEV) días y el incremento anual teórico en los casos de cáncer (IA)

Zona	Casos	PEV	IA
Merced	144	158	2
FES Acatlán	74	81	1
San Agustín	92	101	1
UAM Iztapalapa	118	130	2
Pedregal	85	94	1

Se observa que los sitios con riesgos mayores se encuentran en la Merced y la UAM Iztapalapa y el lugar con menor riesgo es en la FES Acatlán. Para toda la zona metropolitana se tendría un incremento anual teórico en los casos de cáncer de hasta 7 casos al año.

Si bien los resultados preliminares incluyen solamente el segundo semestre de 2011, se puede determinar que el riesgo no es aceptable y se requiere de realizar un estudio más detallado para evaluar mejor los niveles de exposición de la población y las magnitudes de riesgo.

## 5. Conclusiones

Se puede estimar de manera preliminar el riesgo por la exposición a tóxicos atmosféricos en la Ciudad de México. De los resultados preliminares se determinó que el riesgo no es aceptable ya que se rebasa el nivel recomendado de 1 caso por millón, por lo cual se recomienda realizar un estudio más detallado para determinar los niveles de exposición de la población y de las magnitudes de riesgo, con el objeto de desarrollar medidas y políticas para conocer el nivel de reducción de las emisiones y con ello tener una calidad del aire que no afecte la salud.

## Referencias

- [1] H. Keith Lawrence and M. Walter Mary *Handbook of Air Toxics. Sampling Analysis and Properties*. Published by Lewis Publishers C.R.C. Press, Inc. 1995. Printed in the United States of America. Pages 184-186, 349-351, 567-570, 601-608.
- [2] Devotta Sukumar and Chalapati Rao C.V. *Environmental status of India*. Published by Atlantic Publishers and Distributors (P) Ltd., 2008. Printed in India. Page 38
- [3] Secretaria del Medio Ambiente del Distrito Federal *Bases de Datos del Sistema Automático de monitoreo Ambiental*  
<http://www.calidadaire.df.gob.mx/calidadaire/index.php?opcion=2&opcioninfoproductos=2>,  
 Última visita 13 febrero 2013

- [4] Gratt, Lawrence B., *Air Toxic Risk Assessment and Management: Public Health Risk from Normal Operations* John Wiley & Sons Inc. 1996. p 388.
- [5] Oficina de Evaluación de Riesgos Ambientales a la Salud *Consolidated Table of OEH-HA/ARB approved Risk Assessment Health Values*,  
<http://www.arb.ca.gov/toxics/healthval/contable.pdf>, Última visita 13 febrero 2013.
- [6] Instituto Nacional de Estadística y Geografía <http://www.inegi.org.mx/> última vez revisada 11-abril 2012.

*Inga jinicuil* (jinicuil) Como árbol frutal nativo: valor comercial, calidad de frutos y valor cultural como determinantes para permanecer o ser eliminados.

(Ma. Teresa Patricia Pulido-Salas<sup>1</sup>), (Miguel Equihua Zamora<sup>2</sup>),  
(Arturo Pérez Vázquez<sup>3</sup>)

<sup>1</sup>Centro de Investigación Científica de Yucatán, Calle 43 núm. 130  
Col. Chuburná de Hidalgo, C.P.97200. Mérida, Yuc. México

<sup>2</sup>Instituto de Ecología, A.C. Carr. antigua a Coatepec No. 351  
Congregación El Haya 91070 Xalapa, Veracruz México

<sup>3</sup>Colegio de Postgraduados Campus Veracruz Km 26.5  
Carr. Veracruz-Xalapa, rancho Tepetates 91690

Tepetates, Mpio. Manlio Fabio Altamirano, Veracruz. México  
(pulidosalas@gmail.com), (equihuam@gmail.com), (parturo@colpos.mx)

**Resumen.** En México, uno de los problemas que enfrenta el ambiente y en particular la biodiversidad, es la escasa o nula valoración de las especies nativas en términos monetarios y culturales. Ante la falta de datos, ambos aspectos se minimizan frente a los retos propios de las economías emergentes como son la pobreza, la producción de alimentos y la creación de empleos. Los árboles nativos con frutos comestibles han aportado beneficios que no han sido suficientemente cuantificados. Este estudio se realizó en Coatepec, Veracruz, que es una importante región cafetalera inserta en el Bosque Mesófilo de Montaña. *Inga jinicuil* (Leguminosae) es un árbol cuyas vainas se comercializan cada año por sus semillas dulces. Mediante un muestreo polietápico se encontró que los consumidores locales están dispuestos a pagar de 15 % a 100 % más sobre el precio actual existe, lo que constituye un mercado potencial aún no aprovechado. Localmente le confieren también un valor cultural (valor de tradición). Por otra parte, mediante un análisis de Componentes Principales se obtuvo un diagnóstico de la población de árboles usando variables morfológicas del árbol y de los frutos (tamaño de fruto, contenido de azúcar). Los mejores árboles se encontraron en un pequeño rango altitudinal lo que permite inferir que es una especie útil en riesgo. La estadística fue crucial para obtener un diagnóstico multifacético de la situación que enfrenta esta especie nativa de uso tradicional, en una región megabiodiversa que ha sido mermada en años recientes.

**Palabras clave:** *Jinicuil, Árbol, Fruto, Tradicional, Comercio.*

## 1. Introducción

*Inga*, es un género de leguminosas arbóreas exclusivamente americanas (Pennington, 1997). *Inga jinicuil* se encuentra desde la zona montañosa central de Veracruz hasta las montañas de Centroamérica. En Coatepec, Ver., es conocido comúnmente como “jinicuil” y sus frutos (vainas) son objeto de comercio cada verano.

Las semillas de *Inga jinicuil* tienen una cubierta blanca y dulce por lo que se consumen localmente como fruto fresco. Se ha determinado que existe un mercado actual y potencial cuya aportación económica no ha sido suficientemente reconocida (Pulido-Salas, 2009). El consumo de semillas es una tradición prehispánica (León, 1992), iniciada probablemente en las tierras altas de Veracruz (Pennington, 1997). Se reportan también otros usos: como medicinal (Argueta y Cano, 1994), vainas secas para combustible, la madera para construcciones rústicas entre otras. Es una especie con potencial melífero y como forraje por su contenido de proteína bruta (17%) en las semillas (Geilfus, 1989). Esta especie, por su rápido crecimiento, esta recomendada para reforestar áreas naturales en el BMM y vegetación riparia (Castillo-Campos, 1991). Ortíz-Ceballos (2004) en un estudio comparativo de varias especies arbóreas, determinó que tiene una alta capacidad de almacenamiento de carbono.

## 2. Marco teórico

En Veracruz la distribución de *Inga jinicuil* coincide con plantaciones de café (Pennington y Sarukhán, 1998; Challenger, 1998). En Coatepec, coincide principalmente con cafetales insertos en el Bosque Mesófilo de Montaña (BMM) (Luna-Monterrojo, 1997). Se ha calculado que en México, el BMM, con endemismo cercano al 30%, ocupa únicamente el 1% del territorio nacional, a manera de archipiélagos vegetacionales (Rzedowski, 1991). El BMM de Veracruz, por afectaciones derivadas del desarrollo regional, es actualmente un ecosistema amenazado, a pesar de su gran riqueza biológica (Williams-Linera, 2007).

En Coatepec y sus alrededores, uno de los sistemas agrícolas más evidentes es el cafetal, el cual tiene especies frutales y otros árboles de sombra incluyendo a *Inga jinicuil*. Son espacios productivos con una cubierta vegetal que contribuye a la conservación biológica (Perfecto et al., 1996), lo cual es relevante en países megadiversos, tal como es el caso de México (Mittermeier y Goettsch, 1992). Los cafetales arbolados son piezas importantes en para la conservación de biodiversidad local, ya que fungen como resguardo de una gran riqueza de especies, incluyendo a algunas endémicas o amenazadas (Moguel y Toledo, 1996; Gallina et al., 1996; Manson et al., 2008). En años recientes, el número de cafetales con árboles de jinicuil, se ha reducido de manera notable.

El manejo sustentable de especies útiles, debe tener inicio en la autoecología de la especie y en selección de germoplasma, especialmente en el caso de los árboles, que frecuentemente son recursos amenazados, (Wickens, 2001). Una visión desde diversas perspectivas, es esencial para lograr decisiones bien informadas para una producción agrícola. La hipótesis planteada en este trabajo fue que existen factores tanto biológico-ecológicos como factores socio-económicos que determinan la permanencia o eliminación de *Inga jinicuil* en la región.

### 3. Objetivo general

Realizar un diagnóstico del estado de la población de árboles de *Inga jinicuil* y de la percepción de los consumidores, que permitan explicar su permanencia o eliminación como árbol frutal en Coatepec, Ver.

### 4. Aplicación

La sustentabilidad, como una meta deseable para los agroecosistemas, incluye aspectos biológicos, ecológicos, económicos y sociales (Tommasino, 2000). *Inga jinicuil* es considerada especie multiusos de rápido crecimiento (Pennington, 1997) con germinación de 100 % (Benítez et al., 2004). Al implementar adecuadamente un sistema sustentable, se asegura su capacidad de permanencia (Villa-Issa, 2008). Un diagnóstico biológico-ecológico de la población de árboles de jinicuil, conjuntamente con un diagnóstico de la dinámica de aceptación del fruto y su demanda, permite un mejor entendimiento de la problemática de la especie y por lo tanto, permite hacer planteamientos viables para su conservación.

### 5. Materiales y Métodos

Área de estudio. El municipio de Coatepec se ubica en la zona montañosa central del estado de Veracruz, en las faldas del Cofre de Perote ( $19^{\circ} 27' N$  y  $96^{\circ} 58' O$ ). Con una superficie de  $155.81 km^2$  y altitudes que van de los 500 a los 3000m *snm* y la cabecera municipal Coatepec, se encuentra a 1256m de altitud. En 2005 (INEGI) se registró una población de 79,787 habitantes. El clima es templado húmedo con temperatura media anual de  $19.2^{\circ} C$  y precipitación pluvial media anual de 1926mm (Gómez y Soto, 1990). En los alrededores de la cabecera municipal, existen encinares y bosque mesófilo de montaña, que se combina con plantaciones de café de sombra. En las partes bajas existen cultivos de café de sol y caña de azúcar, además algunos elementos florísticos de selva baja caducifolia (Luna-Monterrojo, 1997; Ortiz-Ceballos, 1995; Muñoz et al., 2008). La cuenca cafetalera Xalapa-Coatepec es una de las más antiguas y prósperas de México. Esta actividad define su perfil actual ya que mantiene una importante producción de café (Báez-Landa, 2004). Este trabajo se realizó de 2005 al 2009.

Con base en mapas municipales de las principales poblaciones donde se cultiva café con sombra diversificada (Ortiz-Ceballos, 1995; Luna, 1997; Muñoz-Villers y López-Blanco, 2008), se eligió una finca por localidad. Seleccionamos a 18 dueños pero solo 14 cafetales bajo los siguientes criterios: disponibilidad del dueño para colaborar en este estudio, presencia de árboles de *Inga jinicuil* y que, en conjunto, los predios permitieran representar el gradiente altitudinal donde se encontró la especie.

Encuesta a consumidores. Mediante cuestionarios semi-estructurados, y un muestreo polietápico en la cabecera municipal, se expusieron escenarios hipotéticos como indica el método de Valoración Contingente (Garrod y Willis, 1999) y se indagaron datos acerca de aceptación por gusto y precio del jinicuil así como la calidad deseada como fruto de temporada. Para ello se marcaron 400 cuadros de iguales dimensiones sobre un mapa de la cabecera municipal. Posteriormente con una tabla de números aleatorios se obtuvo el número de entrevistas a realizar (214) correspondiente con el número de habitantes y que

se obtuvo mediante la fórmula que es usada para evaluaciones sociales de interés forestal (Arana-Ovalle, 2003), a saber.

$$n = \frac{1}{\left(\frac{d^2}{Z^2 * S^2}\right) + \frac{1}{n}} \quad (1)$$

En cada cuadro elegido se eligieron dos casas y se entrevistó en cada casa a una persona adulta. Encuesta a propietarios. Se elaboró un cuestionario semi-estructurado que incluyó preguntas relacionadas con datos de los propietarios, de la finca, de la historia de los árboles y de los factores que intervienen en la toma de decisiones sobre su permanencia o eliminación. Concertando previamente una cita, se aplicaron los cuestionarios en entrevista cara a cara. Las 14 fincas seleccionadas se encuentran entre los 850 y los 1530 *snm*, abarcando un gradiente altitudinal de 680*m*. El 74% de las fincas está ubicada entre 1000 y 1200*m snm*. Muestreo de árboles. En cada finca se marcaron cuatro árboles adultos un total de 54 árboles (mayores de siete años o *d.a.p.*  $\geq 15$ *cm*), preferentemente no cercanos entre sí, para evitar un posible parentesco y la consiguiente similitud de datos morfológicos. Se registraron datos del sitio y en cada árbol se midió altura total (mediante el método del leñador) y altura de la primera rama; también el diámetro a la altura del pecho (*d.a.p.*) con cinta diamétrica. Muestreo de vainas y contenido de azúcar. Durante la época de maduración (agosto-octubre) se colectaron vainas frescas para obtener datos morfológicos y de contenido de azúcar. La copa fue dividida en cuadrantes imaginarios (Martínez-Moreno et al., 2006) (Figura 1). Se colectaron de cuatro vainas por cuadrante

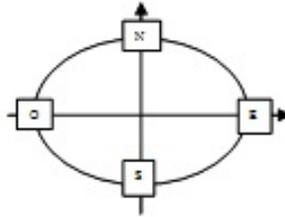


Figura 1: Cuadrante imaginario delimitado en la copa de los árboles de *I. jinicuil* para el muestreo de frutos (vainas). Modificado de Martínez-Moreno et al., 2006.

(16 vainas por árbol). De cada vaina se tomaron datos de peso fresco, longitud, ancho y número de semillas. Se tomó lectura del contenido de azúcar en la cubierta blanca de las semillas con un refractómetro (ATAGO ATC-1, Brix 0-32) y se calculó el promedio de los cuatro valores máximos obtenidos en Grados Brix, para cada árbol. Análisis estadísticos. Los datos recabados con los cuestionarios y los datos de los árboles se capturaron en tablas *Excel* y se utilizaron herramientas estadísticas (Ojeda, 1998). Se realizó análisis de varianza y análisis de regresión ( $P \geq 0.05$ ), así como análisis univariado y multivariado por Componentes Principales con el programa *Statistica 6*.

## 6. Resultados y Discusión

Consumidores. De los entrevistados, el 36% fueron hombres y 64% mujeres con edad promedio de 38 años (18 a 80 años). El 98% de los entrevistados lo consumen regular-

mente; el 100 % esta de acuerdo en el precio actual y el 32 % estaría dispuesto a pagar un precio mayor. La calidad de fruto fue definida así: vainas sin manchas, grandes, llenitas, con semillas blancas, esponjosas y dulces. También le confieren valor de tradición ya que durante la época de cosecha de jinicuiles algunos realizan reuniones para comerlos en familia a lo que denominan "jinicuizada". Propietarios. De los 18 entrevistados el 38 % fueron mujeres y el 62 % hombres con edad promedio de 60 años ( $\pm 15.5$ ). El 55 % de las fincas tiene más de 10 ha ( $\pm 12.7$ ). El 72 % manifestó tener otros ingresos, además del ingreso por café. Refirieron que los árboles de jinicuil dan buena sombra al café y se tienen que eliminar epífitas de sus ramas ("destenchar") una vez al año o cada dos años; pero el consenso es que no requieren cuidados especiales. No propagan estos árboles para sustituir a los ejemplares viejos. Las fincas, el 100 % son propiedad privada con una extensión que va de 1 a 30 ha. En promedio se tienen 2.9 árboles de jinicuil/ha. Los árboles. El intervalo de edad de árboles fue entre 40 y 80 años (56 años promedio). Algunos árboles tienen más de 100 años, según el cálculo que hizo cada propietario. Se pudo definir que, los árboles que tienen una producción regular de vainas, son adultos mayores a siete años de edad o con un d.a.p. mayor a 15 cm. No se realiza propagación dirigida y tampoco se sustituyen los árboles viejos o enfermos. Solo el 44 % de propietarios entrevistados ha sembrado, tolerado o favorecido árboles de jinicuil en los últimos 10 años. El análisis de correlación mostró significancia ( $p > 0.05$ ) entre altitud y *d.a.p.*, altitud y cobertura, altitud y longitud de vaina, altitud y número de semillas por vaina y, longitud de vainas y número de semillas por vaina. Estos resultados se muestran en el cuadro 1. La correlación

<b>Variable morfológica</b>	<b>Valor de correlación</b>
Altitud (m snm) vs d.a.p	0.58
Altitud vs Cobertura	0.62
Altitud vs Long de vaina	0.75
Altitud vs Número de semillas por vaina	0.59
Longitud de vainas vs Número de semillas por vaina	0.57

Cuadro 1: Análisis de correlación entre variables morfológicas

encontrada entre altitud, tamaño de vainas y cobertura, es ser un factor a considerar para mejorar el manejo, con el fin de obtener mejores árboles o frutos y árboles más saludables. Al parecer, existe una pequeña franja altitudinal que favorece un mejor desarrollo de estos árboles. Se encontró que el contenido de azúcar (grados Brix) es alto con respecto a frutas de alta demanda como el mango (19° *Brix*). También se encontró que el contenido de azúcar varía en las semillas de una misma vaina, razón por la cual se tomó el valor más alto para calcular los promedios por árbol. En el Cuadro 2 se resumen los valores encontrados para las variables analizadas. El análisis de regresión y correlación mostró que las variables que explican en un mayor porcentaje la variabilidad encontrada son: Altura de la 1ª. rama, Altitud y Longitud de vainas. Para este último fue tomado el promedio de los cuatro valores máximos: uno/ cuadrante/árbol. La altura de la primera rama muestra una alta dependencia de la variable altura total del árbol, lo que, posiblemente se explica como una consecuencia del crecimiento. La altitud muestra el valor más alto como variable explicativa para la longitud de vainas (cuadro3), lo que nos conduce hacia una mejor calidad de fruto en términos comerciales. La Figura 2 representa el resultado del análisis por Componentes Principales, donde los ejes cortos en torno al origen corresponden a las variables menos representadas, mientras que el ángulo entre los ejes

Variablen	Intervalo obtenido (mín.-máx)	Promedio (DSt)	Moda
Altura del árbol (m)	(4 -17)	12 ( $\pm 2.7$ )	12
Altura donde aparece la 1 <sup>a</sup> rama (m)	(0.6 - 6)	2.7 ( $\pm 1.08$ )	2.5
Diám. a la alt. del pecho <i>d.a.p.</i> (cm)	(23 - 109)	55.4 ( $\pm 20.3$ )	47
Cobertura (m <sup>2</sup> )	(20 -320)	125.5 ( $\pm 57.9$ )	120
Prom. Ancho Máx de vainas/árbol(cm)	(4.5 - 7.6)	6 ( $\pm 0.78$ )	5.37
Prom. Long. Máx de vainas/árbol(cm)	(23 - 46)	34.2 ( $\pm 5.8$ )	35
Prom. Long. Máx de semillas/árbol(cm)	(2.9 - 4.0)	3.4 ( $\pm 0.30$ )	3.5
Número de semillas /vaina	(2 - 20)	11( $\pm 3.7$ )	12
Promedio Azúcar /árbol ( $^{\circ}$ Brix)	(7 - 16)	11.9 ( $\pm 2.2$ )	12

Cuadro 2: Caracterización de variables morfométricas de *I. jinicuil* con promedios de los 54 árboles muestreados en 14 fincas.

Variable independiente	Variable dependiente	Valor de r	Valor de R <sup>2</sup> (%)
Altitud	Altura del árbol	0.19	3
Altitud	Longitud de vainas	0.46	21
Altitud	<i>d.a.p.</i>	0.43	18
Altura del Árbol	1 <sup>a</sup> Rama	0.53	28
Altura del Árbol	Cobertura	0.43	18
<i>d.a.p.</i> ,	Prom Long Máx vainas	0.42	17
<i>d.a.p.</i> ,	Cobertura	0.43	19
<i>d.a.p.</i> ,	Prom Máx $^{\circ}$ Brix	0.24	06
Cobertura	Prom Máx $^{\circ}$ Brix	0.15	02
Promedio de Longitud Máxima de vainas	Prom Máx $^{\circ}$ Brix	0.13	01
Altitud	Prom Máx $^{\circ}$ Brix	0.5	0.3

Cuadro 3: Valores obtenidos en el análisis de Regresión y de Correlación para los 54 árboles

biplot aproxima el grado de correlación (Jongman et al., 1987; Cárdenas et al., 2006). Se infiere que la longitud de vainas, el *d.a.p* y la altitud son variables explicativas, en un mayor porcentaje, para la diversidad morfológica encontrada, siendo las dos últimas las que muestran mayor relación, lo que se confirma con los valores de  $R^2$ . El análisis de regresión mostró un bajo coeficiente de determinación ( $R^2$ ) para explicar la relación entre las variables morfométricas. Esto se debe a la alta variabilidad que se reporta para el género *Inga*, incluyendo la posible hibridación (Pennington, 1997) por lo que se considera un género en plena diversificación (Richardson et al., 2001). La hibridación es favorecida en cafetales de Coatepec ya que coexisten varias especies. El análisis por factor de correlación aporta eigen valores de 2.5 para el Factor 1, compuesto por variables de altitud y morfometría del árbol (altura total, altura de 1<sup>a</sup> rama, *d.a.p.* y cobertura) lo que explica el 32 % de la variación. El valor para el Factor 2 compuesto por variables relacionadas con el fruto

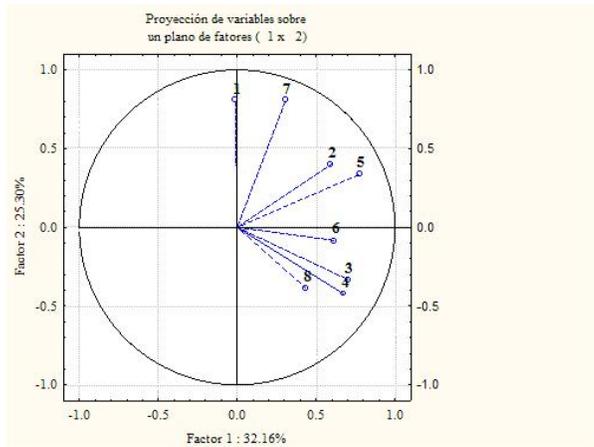


Figura 2: Proyección de las variables sobre un plano de factores de correlación (1 x 2). Las variables representadas son: 1) Árbol; 2) Altitud; 3) Altura de árbol; 4) Altura de la 1ª. rama; 5) Diámetro a la altura del pecho (*d.a.p.*); 6) Cobertura; 7) Promedio de valores máximos de longitud de vaina; 8) Promedio de valores máximos de °Brix.

(promedio de los valores máximos de longitud de vainas y promedio del valor máximo de °Brix/ árbol), mismos que aportan un valor de 1.55, lo que explica el 22% de la variación (Figura 3). El análisis por factores de correlación permite señalar al menos cuatro árboles

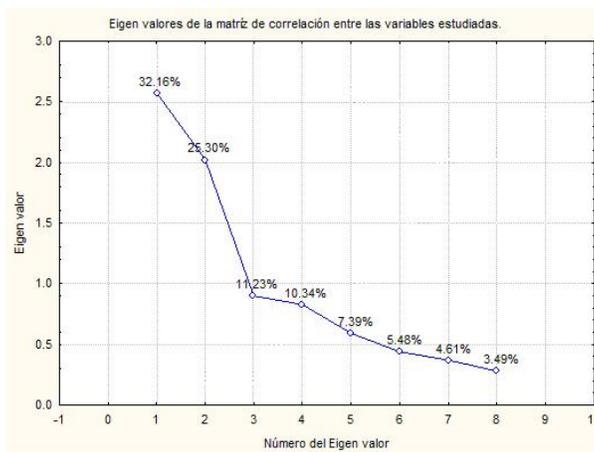


Figura 3: Eigenvalores obtenidos en el análisis por Factores de Correlación.

(ED III, GP II, GP IV y TS III) que tienen mejor calidad, es decir, con mejor relación entre características deseadas como tamaño de vaina y contenido de azúcar (Figura4). En cuanto a la relación entre altitud y los promedios de la longitud máxima de vainas por

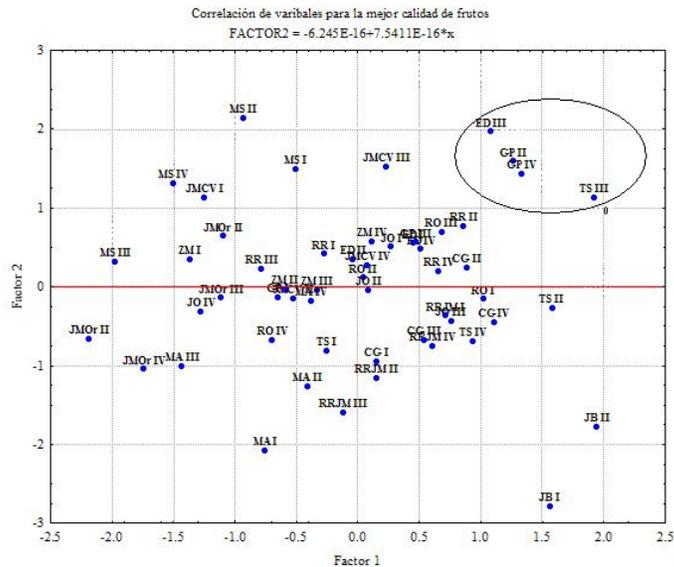


Figura 4: Gráfica que permite inferir sobre cuales podrían ser considerados con mejores características, quedando éstos en el cuadrante superior derecho.

finca, el análisis de regresión muestra una  $r = 0.92$  y una  $R^2$  de 85% (Figura 5). Esto indica que las vainas más grandes se encontraron en altitudes entre 1050 y 1150 msnm, por lo que se infiere que la altitud es un factor importante para el tamaño de las vainas. Este resultado es relevante para una eventual planeación del manejo como árbol frutal, si recordamos que los consumidores expresaron preferencia por vainas grandes. Existe entonces una zona de limitado rango altitudinal que es favorable para la permanencia de *Inga jinicuil* como árbol frutal que sería redituable en un mercadeo regular. En el análisis de correlación de variables por finca se obtienen valores de significancia ( $p > 0.05$ ) para la altitud, lo que confirma que los árboles con mejores características para la producción de frutos, por lo menos por longitud de vainas, podrían estar restringidos a cierto intervalo de altitud. Esta correlación a su vez podría estar relacionada con la zonificación climática del municipio (Gómez y Soto, 1990), ya que las fincas muestreadas se encuentran en la zona de temperatura media anual no mayor de  $20^{\circ}C$  ni menor de  $14^{\circ}C$  y una precipitación pluvial anual no mayor de  $2000mm$  ni menor a  $1500mm$ . De la relación entre  $^{\circ}Brix$  y altitud, no se encontró una relación estrecha  $r = 0.1$  y  $R^2 = 1\%$ , debido a la gran variación encontrada en el contenido de azúcar dentro de una misma vaina. De los valores de contenido de azúcar, los valores más altos de grados Brix promediados por árbol, se encontraron en altitudes cercanas a los 1100 msnm (Figura 6), lo que corresponde con el dato promedio de longitud máxima de vainas. Cabe señalar que de enero de 2006 a diciembre de 2008 se perdieron 4 de los 54 árboles marcados para este estudio. Uno por urbanización (TS-I), y dos por eventos climáticos o pudrición (ED-I y RR-IV); uno más por su cercanía a una construcción (JB-II) con buena producción de frutos dulces (hasta  $17^{\circ}Brix$ ). Esto representa una pérdida del 2.4% anual. Entre las razones para eliminar o

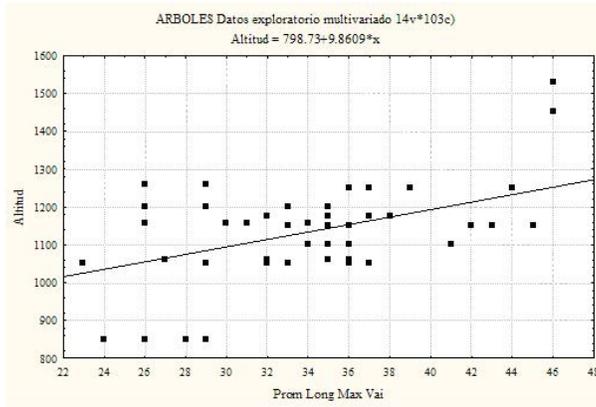


Figura 5: Correlación entre altitud (m snm) y promedio de longitud de vainas/finca.

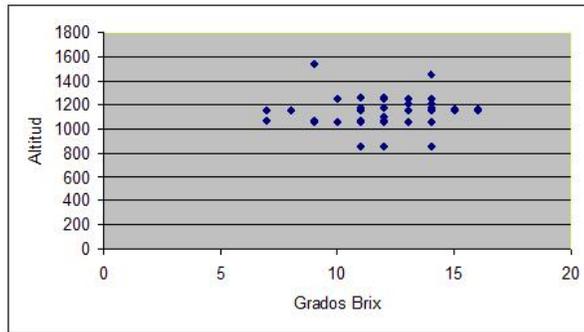


Figura 6: Dispersión de los datos obtenidos de contenido de azúcar en los 54 árboles muestreados (promedio de valores máximos por árbol), con respecto a la altitud.

sustituir un árbol de jinicuil los propietarios coinciden en tres: 1) La aparición de alguna otra oportunidad para mejorar los ingresos 2) La edad avanzada de los árboles por lo que se debilitan convirtiéndose en un riesgo latente por la caída eventual y 3) El robo de vainas que da como consecuencia daños a los cafetos. Se identificaron en total diez factores que contribuyen a determinar la permanencia o eliminación de la especie en la región estudiada: cuatro socio-económicos y seis ecológico-biológicos. Tal diversidad de factores, vistos como causas que afectan a la especie estudiada, nos conducen a aportar una visión transdisciplinaria (Ruíz-Rosado, 2006), con la finalidad de buscar un manejo adecuado ya que siendo un árbol multibenéfico, su permanencia es deseable en la región de Coatepec, favoreciendo la conservación de la biodiversidad local y de un fruto nativo con valor de tradición.

## 7. Conclusiones

La permanencia de árboles de *Inga jinicuil* está relacionada a tres tipos de factores: la aparición de oportunidades económicamente atractivas para el propietario, el envejecimiento de los árboles sin ser replantados y el riesgo de robo de vainas, con el eventual daño para los cafetos. Sin embargo, el mercado potencial encontrado no está siendo aprovechado para mantener estos árboles para producción de frutos. Dada la relación encontrada entre la altitud y las características de fruto, las mejores vainas se encuentran en un reducido gradiente altitudinal (entre 1000 y 1200msnm = 200m). La altitud parece ser un factor determinante para obtener una mejor producción y calidad de vainas para satisfacer el mercado potencial encontrado. Sin embargo, esta franja altitudinal corresponde con los alrededores del poblado Coatepec, que es donde se concentra el crecimiento urbano. Esta misma relación de variables señala también la vulnerabilidad de la especie. La pérdida de cuatro árboles marcados, permitió calcular una tasa de pérdida de 2.4% anual. Esto aunado a que la edad promedio de los árboles actuales es de 56 años, y que el reemplazo comúnmente no está contemplado en el manejo de las fincas, podría significar que de continuar las condiciones actuales, la especie *Inga jinicuil* podría desaparecer en la región en 40 años. Para un manejo sustentable de *Inga jinicuil* se debe incluir la propagación sistemática para mantener una población saludable con edades intermedias. La identificación de diversos factores que intervienen en la problemática de la especie, permite sugerir que con canales eficientes para la información dirigida a los propietarios, se puede lograr mayor objetividad para las decisiones que afectan a esta especie.

## 8. Agradecimientos

Al CONACYT por el apoyo a través de la beca doctoral 218825. A la M.C. Griselda Benítez del Depto. de Ecología aplicada del INECOL, por conceder el apoyo para el trabajo de campo.

## Referencias

- [1] Arana-Ovalle R.I. Métodos de muestreo. Tesis profesional. División de Ciencias Forestales, Universidad Autónoma Chapingo, Chapingo, México. 2003.
- [2] Argueta, A. y L. Cano (Coord.). Atlas de las plantas de la medicina tradicional mexicana, vol. I, II y III. Instituto Nacional Indigenista. México, D.F. 1786 p. 1994.
- [3] Báez Landa, M. Los señores, la tierra y los indios, la formación de una región cafetalera de Veracruz. Editora de Gobierno. Xalapa, Veracruz. 281 p. 2004.
- [4] Benítez, G., M.T. Pulido-Salas y M. Equihua. Árboles multiusos nativos de Veracruz para reforestación restauración y plantaciones. Comisión Nacional Forestal, Sistema SIGOLFO, e Instituto de Ecología, A.C. Xalapa, Ver. México. 420 p. 2004.
- [5] Cárdenas, O., C. Noguera, P. Galindo y J.L. Vicente. Alternativa a la regresión con componentes principales basada en biplot de regresión. *Interciencia* 31 (3): 160-167. 2006.

- [6] Castillo-Campos, G. Vegetación y Flora del Municipio de Xalapa, Veracruz. Instituto de Ecología. Xalapa, Ver., México. 148 p. 1991.
- [7] Challenger, A. Utilización y Conservación de los ecosistemas terrestres de México. CONABIO, Instituto de Biología, UNAM y Asoc. Sierra Madre. México, D.F. pp: 443-518. 1998.
- [8] Gallina, S., S. Mandujano y A. González-Romero. Conservation of mammalian biodiversity in coffee plantations of Central Veracruz, México. *Agroforestry Systems* 33: 13-27. 1996.
- [9] Geilfus, F. El árbol al servicio del agricultor, manual de agroforestería para el desarrollo rural. Vol. 2. ENDA-Caribe-CATIE. Turrialba, Costa Rica. 625 p. 1989.
- [10] 10) Garrod G. y K.G. Willis. 1999. Economic valuation of the environment. Edgard Elgar. Northampton, MA. USA. 384 p.
- [11] Gómez, M. y M. Soto. Atlas climático del municipio de Coatepec, Ver. Instituto de Ecología, A.C. Xalapa, Ver. México. 47 p. 1990.
- [12] INEGI. Censo Nacional de Población. Instituto Nacional de Estadística, Geografía e Informática. 22 enero 2007. <http://www.inegi.gob.mx/inegi/default.aspx>. 2005.
- [13] Jongman, R.H.G., C.J.F. ter Braak y O.F.R. van Tongeren (Eds.). Data analysis in community and landscape ecology. Pudoc. Wageningen. 299 p. 1987.
- [14] Luna-Monterrojo, V. Flora y vegetación del municipio de Coatepec, Ver. Tesis Profesional. Facultad de Biología, Universidad Veracruzana. 192 p. 1997.
- [15] Manson, R., V. Hernández-Ortíz, S. Gallina y K. Mehtleter (Eds.). Agroecosistemas cafetaleros de Veracruz. Instituto de Ecología, A.C. Xalapa, Ver. 248 p. 2008.
- [16] Martínez-Moreno, E., T. Corona-Torres, E. Avitia-García, A.M. Castillo-González, T. Terrazas-Salgado y M.T. Colinas-León. Caracterización morfométrica de frutos y semillas de nanche (*Byrsonima crassifolia* (L.) H.B.K.). *Revista Chapingo Serie Horticultura* 12 (1): 11-17. 2006.
- [17] Moguel, P. y V.M. Toledo. Café, luchas indígenas y sostenibilidad: el caso de México. *Ecología Política* 18:23-36. 1996.
- [18] Ojeda, M.M. Análisis exploratorio de datos: con énfasis multivariado en el contexto de aplicaciones ecológicas. Fac. Estadística e Informática, Universidad Veracruzana. Xalapa, Ver. México. 92 p. 1998.
- [19] Ortiz-Ceballos, G. Los ambientes para la producción agropecuaria y forestal en el municipio de Coatepec, Ver. Tesis de Maestría en Ciencias. Colegio de Postgraduados, Campus Veracruz. 145 p. 1995.
- [20] Pennington, T.D. The genus *Inga*, botany. The Royal Botanic Gardens. Kew, Reino Unido. 844 p. 1997.
- [21] Pennington, T.D. y J.Sarukhán. Árboles Tropicales de México. 2ª. Ed. Fondo de Cultura Económica y Universidad Nacional Aut. de México (UNAM). 521 p. 1998.

- [22] Perfecto I., Rice, R.A., Greenberg R. y Van der Voort M. Shade coffee: a disappearing refuge for biodiversity. *Bioscience* 46 (8): 598-608. 1996.
- [23] Pulido-Salas, M.T. *Inga jinicuil* ("jinicuil") en Coatepec, Ver., México, como árbol frutal nativo: potencial económico y factores que afectan su permanencia. Tesis doctoral. Colegio de Postgraduados Campus Veracruz. Tepetates, Ver. México. 178 p. 2009.
- [24] Richardson, J.E., R.T. Pennington, T.D. Pennington y P.M. Hollingsworth. Rapid diversification of a species-rich genus of neotropical rain forest trees. 2001.
- [25] Ruíz-Rosado, O. Agroecología: una disciplina que tiende a la transdisciplina. *Inter-ciencia* 31(2): 140-245. 2006.
- [26] Rzedowski, J. El endemismo en la flora fanerogámica mexicana: una apreciación analítica preliminar. *Acta Botánica Mexicana* 15: 47-64. 1991.
- [27] Sousa, S., M. El género *Inga* (Leguminosae: Mimosoideae) del sur de México y Centroamérica, estudio previo para la flora Mesoamericana. *Annals of the Missouri Botanical Garden* 80: 223-269. 1993.
- [28] Tommasino, H. Sustentabilidad rural: desacuerdos y controversias. En: G. Foladori y H. Tommasino. El concepto de desarrollo sustentable 30 años después. *Desenvolvimento e Meio Ambiente* Cap. 4. 1: 412-56. 2000. <http://sprausg.googlepages.com/DESARROLLOSUSTENTABLE30AOSDESPUS.pdf> . 16 junio 2008.
- [29] Villa-Issa, M. Qué hacemos con el campo mexicano? Colegio de Postgraduados. Colegio de Puebla y Mundi Prensa. México 230 p. 2008.
- [30] Wickens, G.E. *Economic botany, principles and practices*. Kluwer. Academia Publ. Amsterdam. 441 p. 2001.
- [31] Williams-Linera, G. El bosque de niebla del centro de Veracruz: ecología, historia y destino en tiempos de fragmentación y cambio climático. Comisión para el conocimiento y uso de la biodiversidad (CONABIO) e Instituto de Ecología, A.C. (INECOL) Xalapa, Ver. México. 208 p. 2007.

# Modelación del Ozono en la Estación Agua Santa de la Ciudad de Puebla usando L-moments

Juan Antonio Cruz Juárez<sup>1</sup>  
Hortensia Reyes Cervantes<sup>1</sup>  
Sara Rodríguez Rodríguez<sup>1</sup>  
Humberto Vaquera Huerta <sup>2</sup>.

Fac. de Ciencias Físico-Matemáticas; Universidad Autónoma de Puebla<sup>1</sup>,  
Colegio de Posgraduados; Campus Texcoco<sup>2</sup>.  
juanantonio\_63@hotmail.com

## Resumen

En este presente trabajo se utiliza la técnica descrita por Hosking en 1989. El método se basa primordialmente en el cálculo de los estimadores de L-Moments, cuentan con la propiedad de robustez, son útiles en bases de datos pequeñas y pueden contener datos atípicos. Los estimadores de L-Moments caracterizan a los estimadores de momentos ponderados, dado que son combinaciones lineales de los estadísticos de orden. Se presenta la aplicación de esta metodología a un problema físico-ambiental, correspondiente al contaminante ozono en una estación de monitoreo que forma parte de la Red de Monitoreo Atmosférico del Estado de Puebla (REMA). También se supone que los datos siguen una distribución de Valores Extremos Generalizada (VEG). Algunos resultados que nos arrojan, es poder encontrar los estimadores de los parámetros de la distribución, más aún, con el fin de realizar la comparación entre los estimadores de máxima verosimilitud y los de L-Moments se hallan los estimadores por el método de máxima verosimilitud, resultando que son muy similares en cuanto a las cantidades obtenidas. Esto nos da confianza en usar esta metodología para estaciones meteorológicas que por alguna causa todavía no tienen un buen manejo de la información y se pierden.

**Palabras clave:** Ozono, estación de monitoreo, L-moments.

## 1. Introducción.

Se supondrá como ley válida en las observaciones una distribución de Valores Extremos Generalizada (VEG). Muchas estaciones de monitoreo al comenzar al operar sus sistemas de toma de datos sufren de pérdida de información, debido a que los aparatos no están bien calibrados o todavía no se cuenta con todo el equipo funcionando adecuadamente. Por lo cual, se empieza a sufrir con pérdida de datos por algunos años y es aquí donde se plantean alternativas para poder analizar la información cuando hay ausencia de datos o existen valores atípicos.

La metodología presentada se usa en un problema real del área físico-ambiental, es la contaminación por ozono en la estación Agua Santa que se encuentra al Sur de la Ciudad de Puebla.

## 2. Marco teórico

El análisis de datos en la vida cotidiana es un trabajo arduo que requiere información tecnológica para manipular grandes cantidades de información y desgraciadamente los fenómenos ambientales son inestables debido al calentamiento global. Actualmente existen diversas técnicas variadas que ayudan a los investigadores para interpretar o predecir los fenómenos que están estudiando.

Los fenómenos ambientales son de gran estudio actualmente dado que proveen al investigador de herramientas para toma de decisiones ante los gobiernos y la sociedad. La Distribución de Valores Extremos Generalizada (VEG), es la función de densidad idónea que mejor representa la información en este tipo de problemas. Esta función es usada en diversos campos del conocimiento, como son:

1. Finanzas, Riesgos.
2. Problemas ambientales como: Nivel del mar, Velocidad del viento, Nivel de un río o presa, Concentración de contaminantes, Lluvias, Oleaje, Etc.
3. Geografía.

La teoría de Valores extremos inicia fundamentalmente en un artículo propuesto por Fisher & Tippet (1928), en donde se enfocan tres familias de distribución, Gumbel (I), Fréchet (II) y Weibull (III).

TEOREMA 2.1 *Supongamos que existen  $a_n > 0$ ,  $b_n \in \mathbb{R}$  para  $n \geq 1$  tales que*

$$P(M_n - b_n/a_n \leq x) = F^n(a_n + b_n) \rightarrow G(x) \quad (1)$$

*débilmente cuando  $n \rightarrow \infty$  y donde  $G$  es una función no degenerada, entonces  $G$  pertenece a alguna de las siguientes tres familias de distribuciones:*

$$I : G_0(x) = \exp(-e^{-x}), -\infty < x < \infty. \quad (2)$$

$$II : G_{1,\alpha}(x) = \begin{cases} \exp(-x^{-\alpha}) & \text{si } x \geq 0 \\ 0, & \text{si } x < 0 \end{cases} \quad (3)$$

$$III : G_{2,\alpha}(x) = \begin{cases} \exp(-(-x)^{-\alpha}) & \text{si } x \leq 0 \\ 1, & \text{si } x > 0. \end{cases} \quad (4)$$

En los años 50 estas tres familias se unifican para formalizarlas en una sola, a esta distribución se le conoce como Distribución de Valores Extremos Generalizada:

$$G_\xi(x) = \exp\{-(1 + \xi x)^{-1/\xi}\} \quad (5)$$

o incluyendo los parámetros de localización y escala:

$$G_{\xi,\mu,\sigma}(x) = exp\left\{-\left(1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right)^{-1/\xi}\right\} \quad (6)$$

definida sobre el conjunto

$$\{x : 1 + \xi\left(\frac{x - \mu}{\sigma}\right) > 0\} \quad (7)$$

dado que de lo contrario,  $G$  adoptaría el valor de 0 o 1.

La teoría de valores extremos encuentra diversas aplicaciones en muchos fenómenos que se estudian.

## 2.1. Estimadores de L-Moments

Los estimadores más usados en la modelación estadística son conocidos como estimadores de máximo-verosimilitud que tienen propiedades que los hacen ser los mejores cuando el tamaño de la muestra es grande, mientras que en tamaños muestrales finitos (pequeños) se hace uso de otro tipo de estimadores, como momentos de probabilidad ponderada, L-Moments, etc. Los estimadores L-Moments caracterizan a los estimadores de momentos ponderados, dado que son combinaciones lineales de los estadísticos de orden.

Los primeros L-Moments son [7]:

$$\lambda_1 = E(X_{(1:1)}) \quad (8)$$

donde  $X_{(1:1)}$  es estadístico de orden 1.

$$\lambda_2 = 1/2E(X_{(2:2)} - X_{(1:2)}) \quad (9)$$

donde  $X_{(2:2)}$  (máximo) y  $X_{(1:2)}$  (mínimo)

$$L - CV = \tau = \lambda_2/\lambda_1 \quad (10)$$

$$\tau_3 = \lambda_3/\lambda_2 \quad (11)$$

$$\tau_4 = \lambda_4/\lambda_2 \quad (12)$$

Estos estimadores tienen las siguientes propiedades:

1. Existencia. Si la media de la distribución existe, entonces todos los L-Moments existen.
2. Unicidad. Si la media de la distribución existe, entonces los L-Moments son únicos para definir una distribución, esto es, no dos distribuciones tienen los mismos L-Moments.

3. Términos:

$\lambda_1$  es el L-localización o media de la distribución.

$\lambda_2$  es el L-escala.

$\tau$  es el L-CV.

$\tau_3$  es la L-asimetría y  $\tau_4$  es el L-curtosis.

## 2.2. Estimadores de L-Moments y la Distribución de VEG

La Distribución de Valores Extremos Generalizada tiene tres parámetros: escala ( $\sigma$ ), localización ( $\mu$ ), y forma ( $\xi$ ).

En esta sección nos encargaremos de mostrar la relación de L-Moments con la distribución de valores extremos, se describen la expresiones para los primeros cuatro L-Moments y los parámetros estimados.

Como se ha comentado la distribución VEG es de utilidad para describir fenómenos en donde haya observaciones extremas, que comúnmente se presenta en fenómenos naturales.

De igual manera se describen tres familias de distribuciones que son: Weibul ( $\xi < 0$ ), Frechet ( $\xi > 0$ ), Gumbel ( $\xi = 0$ ).

La función de distribución de Valores Extremos Generalizada se define como [4]:

$$f(x) = \alpha^{-1} \exp[-(1 - \xi)Y - \exp(Y)] \quad (13)$$

$$F(x) = \exp[-\exp(Y)] \quad (14)$$

donde

$$Y = \begin{cases} -\xi^{-1} \log[1 - \xi(x - \mu)/\sigma] & \text{si } \xi \neq 0 \\ (x - \mu)/\sigma & \text{si } \xi = 0 \end{cases} \quad (15)$$

y

$$x(F) = \begin{cases} \mu + \sigma(1 - [-\log(F)]^\xi)/\xi & \text{si } \xi \neq 0 \\ \mu - \sigma \log[-\log(F)] & \text{si } \xi = 0. \end{cases} \quad (16)$$

Los rangos de la variable aleatoria (v.a.)  $x$  son:

$$-\infty < x < \mu + \sigma/\xi \quad \text{si } \xi > 0 \quad (17)$$

$$-\infty < x < \infty \quad \text{si } \xi = 0 \quad (18)$$

$$\mu + \sigma/\xi \leq x < \infty < 0 \quad \text{si } \xi < 0. \quad (19)$$

Los L-moments son [7]:

$$\lambda_1 = \mu + \sigma[1 - \Gamma(1 + \xi)]/\xi \quad (20)$$

$$\lambda_2 = \sigma(1 - 2^{-\xi})\Gamma(1 - \xi)/\xi \quad (21)$$

$$\tau_3 = 2(1 - 3^{-\xi})/(1 - 2^{-\xi}) - 3 \quad (22)$$

$$\tau_4 = \frac{5(1 - 4^{-\xi}) - 10(1 - 3^{-\xi}) + 6(1 - 2^{-\xi})}{(1 - 2^{-\xi})}, \quad (23)$$

donde,  $\Gamma(\cdot)$  denota la función gamma completa:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt, \quad (24)$$

con,  $x \in \mathbb{R}$ .

De las ecuaciones anteriores se obtienen los parámetros de escala, localización y forma.

$$\sigma = \frac{\lambda_2 \xi}{(1 - 2^{-\xi}) \Gamma(1 + \xi)} \quad (25)$$

$$\mu = \lambda_1 - \sigma[1 - \Gamma(1 + \xi)]/\xi. \quad (26)$$

No existe una solución explícita para el parámetro  $\xi$  en términos de los L-Moments y sólo se realizan métodos numéricos para su hallazgo. Aunque Hosking [9] da una aproximación:

$$\xi \approx 7.8590c + 2.9554c^2,$$

donde,

$$c = \frac{2}{3 + \tau_3} - \frac{\log 2}{\log 3}.$$

### 2.3. Diagramas de radio de L-Moments

Una herramienta fuerte con la que cuenta el análisis de datos a través de los L-Moments se presenta en el estudio de los diagramas de radio de los L-Moments.

Los diagramas de radio basan su estudio principalmente en dos de los cua-tro estimadores de L-Moments ( $\tau_3, \tau_4$ ), dado que es común analizar los dos primeros ( $\lambda_1$  y  $\lambda_2$ ) en cualquier distribución, ya sea de un modelo lineal o lineal generalizado.

Tal vez una de las formas más adecuadas de visualizar cómo los L-Moments contribuyen a definir un tipo de distribución es a través del denominado Diagrama de radio de L-Moments.

### 2.4. Construcción de los diagramas de radio de los L-Moments

Hosking and Wallis (1997, p.208) reportan aproximaciones polinómicas para la caracterización de  $\{\tau_3, \tau_4\}$  relacionando estas aproximaciones mediante la construcción de los diagramas de radio de L-Moments.

La aproximación polinomial es de la siguiente forma:

$$\tau_4 = \sum_{j=0}^8 A_j \tau_3^j \quad j > 0 \quad (27)$$

de donde los coeficientes  $A_j$  son seleccionados para tres parámetros de las distribución. En el Cuadro 1, se presentan las siguientes distribuciones: GEV, distribución de valores extremos generalizada; GLO, distribución logística generalizada; GNO, distribución normal generalizada; PE3, distribución Pearson tipo III; – implica que los coeficientes son cero.

	GEV	GLO	GNO	GPA	PE3
$A_0$	0.10701	0.16667	0.12282	0.	0.12240
$A_1$	.11090	–	–	.20196	–
$A_2$	.848338	.83333	.77518	.95924	.30115
$A_3$	-.06669	–	–	-.20096	–
$A_4$	.00567	–	.12279	.04061	.95812
$A_5$	-.04208	–	–	–	–
$A_6$	.03763	–	-.13638	–	-.57488
$A_7$	–	–	–	–	–
$A_8$	–	–	.11368	–	.19383

Cuadro 1: Valores de los coeficientes del polinomio de aproximación por Hosking y Wallis (1997) del estimador L-kurtosis como función de L-asimetría

### 3. Ozono en Puebla

Las múltiples actividades que se desarrollan en las zonas urbanas tienen la finalidad de proveer bienes y servicios a la sociedad. Sin embargo, esto conlleva un costo ecológico, el cual se ve reflejado en la degradación de los recursos naturales y en la generación de residuos y contaminantes.

La calidad de aire en las zonas urbanas se mide con base a la concentración de contaminantes atmosféricos presentes, de esta manera, mientras más grande sea la generación, emisión y residencia de los mismos, menor será la calidad del aire, y en consecuencia, mayor será el impacto sobre los diferentes escenarios naturales y humanos.

Por lo que respecta a la Ciudad de Puebla las emisiones de contaminantes atmosféricos empiezan a tener presencia significativa a partir de los años 60 y 70, reflejándose en paralelo con su incremento demográfico, industrial y comercial, que en consecuencia demandan un mayor suministro de bienes y servicios, así como de transporte [17].

#### 3.1. Aplicación

Se supone que las observaciones de Ozono pueden aproximarse con la distribución de Valores Extremos Generalizada. En la teoría de VEG se contemplan observaciones aleatorias e independientes, puesto que las observaciones que se analizan no cumplen con este requisito debido a que los datos son consecutivos y están bajo las mismas condiciones ambientales, se procede a realizar bloques de los datos obtenidos y de esta manera disminuir la correlación entre las observaciones de  $O_3$ , [4].

La estación de la cual se toman las base de datos a trabajar surgen de la estación Agua santa que forma parte de la Red de Monitore atmosférico del Estado de Puebla, de las cuales se inicia con la limpieza de los datos. Terminada la limpieza de nuestra base de datos iniciamos con la el análisis, tanto descriptivo, como inferencial de los datos, las variables a considerar son: Presión Barométrica (BPR), Monóxido de carbono(CO), Humedad Relativa (HR), Dióxido de Nitrógeno ( $NO_2$ ), Ozono ( $O_3$ ), Partículas menores a 10 micrometros (PM10), Dióxido de Azufre ( $SO_2$ ), Temperatura (TEMP), Radiación ultravioleta tipo A y tipo B (UV-A y UV-B, respectivamente), Dirección del viento (WD), Velocidad del viento (WS).

La estación en cuanto a  $O_3$ , aproximadamente el 25% de los datos se encuentran por arriba de 0.7 ppm.



**1.- Estación Agua Santa:**  
 Ubicación: Zona Sur de la Cd. de Puebla.  
 Dirección: Prol. 11 sur y periférico.  
 Localidad: Agua Santa.

Figura 1: Ubicación geográfica Agua Santa.

Contaminante	Media	s.d.	25 %	50 %	75 %	n
BPR	596.39	1.54	595.615	596.09	597.00	487
CO	0.61	0.38	0.370	0.57	0.79	487
HR	40.49	15.93	30.050	40.10	50.90	487
NO <sub>2</sub>	10.42	6.46	5.950	9.10	13.10	487
O <sub>3</sub>	60.19	23.57	42.850	57.10	72.95	487
PM10	42.57	31.09	19.700	36.50	57.50	487
SO <sub>2</sub>	3.49	1.85	2.100	3.10	4.60	487
TEMP	21.84	2.52	620.400	21.70	23.15	487
UV.A	435.35	174.09	341.500	471.00	571.00	487
UV.B	52.64	38.21	32.500	47.00	63.00	487
WD	172.64	80.41	120.800	172.50	216.70	487
WS	5.04	3.71	2.800	3.70	6.10	487

Cuadro 2: Resumen estadístico de los contaminantes de la estación Agua Santa

En el gráfico de dispersión (ver Figura 2) podemos observar la dispersión del contaminante ozono en la estación Agua Santa, las frecuencias máximas de ozono parecen tener una tendencia a la baja durante un cierto periodo, sin embargo, estas frecuencias tienen una tendencia a la alza después del periodo a la baja.

Presentamos los L-Moments para la distribución VEG de nuestros datos (Cuadro 3):

Estación	$\lambda_1$	$\lambda_2$	$\tau_3$	$\tau_4$
Agua Santa	56.90	13.615	0.115	0.101

Cuadro 3: L-Moments para los datos de O<sub>3</sub> en la estación Agua Santa.

A partir de los estimadores de L-Moments podemos construir los parámetros para la distribución deseada, que en nuestro caso es la Distribución De Valores Extremos Generalizada, (ver cuadro 4).

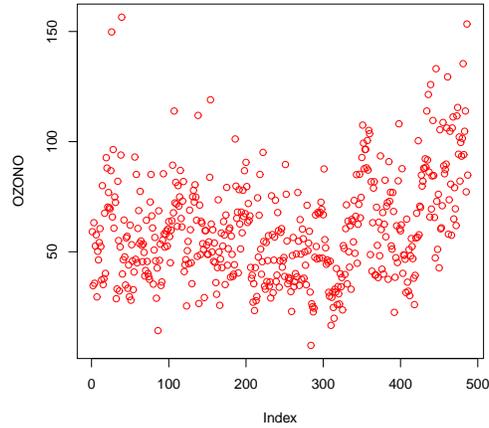


Figura 2: Dispersión de  $O_3$  en la estación Agua Santa.

Estación	$\mu$	$\sigma$	$\xi$
Agua Santa	49.71	19.41	0.039

Cuadro 4: Parámetros de la distribución VEG por el método de L-Moments para los datos de  $O_3$  en la estación Agua Santa.

En la sección 2.2 se habló de la importancia de los diagramas de radio de L-Moments, aplicando esta técnica, obtenemos los siguientes resultados.

En la Figura 3 se muestran los diagramas de radio para la estación Agua Santa que es parte de la REMA en la cual se observa que la distribución que mejor aproxima las observaciones son: la Distribución De Valores Extremo Generalizada, Pearson tipo III y Normal Generalizada, que tienen tres parámetros sin embargo existe otra distribución que solo contiene dos parámetros, Distribución Gumbel.

## 4. Conclusiones

De acuerdo a lo observado, podemos concluir que en efecto la disminución de algunos contaminantes, ha estado presente en la estación Agua Santa que forma parte de la REMA, sin embargo el haber considerado la distribución VEG como válida, no es suficiente para realizar un buen diagnóstico de nuestros datos.

Lo anterior es constatado por el diagrama de radio, los datos tienen cuatro posibles distribuciones, donde, tres distribuciones posibles contienen tres parámetros y la restante solo dos parámetros.

Se obtienen los estimadores por máxima verosimilitud y estos son los resultados:

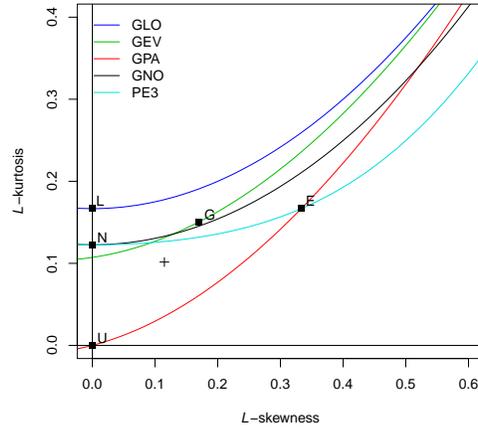


Figura 3: Diagrama de radio de  $O_3$  en la estación Agua Santa.

Método	$\mu$	$\sigma$	$\xi$
Máxima verosimilitud	49.730	19.237	-0.0368
L-Moments	49.71	19.41	0.039

Cuadro 5: Parámetros estimados por el método de MV y el de L-Moments en la estación Agua Santa.

## Referencias

- [1] Asquith, William H. ,*Distributional Analysis with L-moments Statistics using the R Environment fo Statistical Computing*, primera edición, 2011.
- [2] Cánavos, George C., *Probabilidad y Estadística, Aplicaciones y métodos*, 1988.
- [3] Chan, L. K.,*On a characterization of distributions by expected values of extreme order statistics*, Am. Math. Mthly, 74,950-951, 1967.
- [4] Coles, A. *An introduction to statistical modeling of extreme values*, tercera edición, editorial Springer Verlag, ISBN 1852334592, 2004.
- [5] Gail, D. A. S. and Gastwirth, J. L., *A scale-free goodness-of-fit test for the exponential distribution based on the Gini statistic*, J. R. Statist. Soc. B, 40, 350-257, 1957.
- [6] Greenwood, J.A., *Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form*, Wat. Res. Res., 15 (5), pp 1049-1054, 1979.
- [7] Hosking, J. R. M. ,*Some theoretical results concerning L-moments*, IBM Research Division,T. J. Watson Research Center Yorktown Heights, NY 10598, 1989.

- [8] Hosking, J. R. M., *L-moments: analysis and estimation of distributions using linear combinations of order statistics*, Journal of the Royal Statistical Society, Series B, 52, 105-124, 1990.
- [9] Hosking, J. R. M., *Moments or L-moments? An example comparing two measures of distributional shape*, The American Statistician, 46, 186-189, 1992.
- [10] Hosking, J.R.M. and Wallis, J.R., *Regional Frequency Analysis*, 1997.
- [11] Peel, Murray C. Q., et. al., *The utility of L-moment ratio diagrams for selecting a regional probability distribution*.
- [12] Reiss R. y Thomas M., *Statistical Analysis of extreme values*, Birkauer Verlag, Germany, segunda edición, 2001.
- [13] Secretaría de Sustentabilidad Ambiental y Ordenamiento Territorial (SSAOT), *Red de Monitoreo Atmosférico de Puebla (REMA)*, [www.remapuebla.gob.mx](http://www.remapuebla.gob.mx)
- [14] Reyes, Hortensia J., Vaquera, Humberto y Villaseñor, José A., *Estimation of trends in high urban ozone levels using the quantiles of (GEV)*, 2010.
- [15] R, *Paquete lmom en R*, <http://cran.r-project.org/web/packages/lmom/lmom.pdf>
- [16] R, *Paquete lmomco en R*, <http://cran.r-project.org/web/packages/lmomco/lmomco.pdf>
- [17] Semarnat, SSAOT, Puebla, *Programa de Gestión de calidad de la zona Metropolitana del Valle de Puebla 2006-2011*, 2006.
- [18] Semarnat, SSAOT, Benemérita Universidad Autónoma de Puebla, *Programa de Destiación de Calidad del aire del estado de Puebla 2012-2020*, 2012.
- [19] Yee, T. W., *VGAM, Family Functions for Extreme Value Data*, November 20, 2007.

# Predicción de la precipitación media anual para el estado de Puebla.

Marcial Castillo Luis René<sup>1</sup>

Rivera Martínez Marcela<sup>1</sup>

Sandoval Solís María de Lourdes<sup>1</sup>

Línea de investigación:Atmosférico.

<sup>1</sup>Benemérita Universidad Autónoma de Puebla.

*14 Sur y Av. San Claudio, Col. San Manuel, C.U,*

*Puebla, Puebla, CP 74570, México.*

*e-mail:lmc@cs.buap.mx*

## Resumen

En este trabajo, se implementan las técnicas media móvil, ajuste polinomial, interpolación usando transformada de Fourier y suavizado exponencial, como métodos para realizar la predicción de la precipitación media anual en el estado de Puebla. Se desarrollan las pruebas usando la base de datos histórica de la SEMARNAT y se presenta la comparación entre los métodos tomando como medida de error, el error absoluto entre el valor real y el que predice la técnica.

**Palabras clave:** Predicción, precipitación media, SEMARNAT.

## 1. Introducción

La predicción de series temporales constituye un área de gran aplicación práctica en diversos ámbitos del conocimiento, como son: ingeniería, economía, finanzas, demografía, meteorología, hidrología, medio ambiente, medicina, etc. En este trabajo, se investigan distintas metodologías que se utilizan para realizar la predicción de series temporales, se implementan los métodos y se obtiene un producto de software que es capaz de realizar la predicción de la precipitación media anual para el estado de Puebla. Este trabajo, está organizado de la manera siguiente: en la sección 2, se detallan los objetivos de este trabajo; se explican en la sección 3, los métodos de predicción como son: media móvil [1], ajuste polinomial [2], interpolación usando transformada de Fourier[3] y suavizado exponencial [4]; en la sección 4, se explica el problema de aplicación; en la sección 5, se muestran los resultados, donde se comparan los métodos implementados en base al error absoluto de predicción en datos conocidos y se da la predicción de cada uno de los métodos para el año 2012; se reportan las conclusiones y trabajo futuro en la sección 6.

## 2. Objetivos

Los objetivos de este trabajo son:

1. Comparar las técnicas de predicción de series temporales como son: media móvil, ajuste polinomial, interpolación usando transformada de Fourier y suavizado exponencial.
2. Realizar la implementación computacional de las técnicas de predicción usando el lenguaje de programación OCTAVE [5].
3. Probar los métodos en su funcionamiento usando como medida el error absoluto, para las pruebas se usa la base de datos de la SEMARNAT [6], la cual proporciona los datos históricos de la precipitación media anual para el estado de Puebla desde el año 1941 al año 2011.
4. Realizar las predicciones y observar qué técnicas funcionan mejor para este caso particular.

### 3. Marco Teórico

Una serie temporal es una sucesión de observaciones consecutivas de una variable en distintos momentos del tiempo [1], estas observaciones provienen a veces de una función de distribución conjunta que puede o no ser diferente en cada instante de tiempo y otras veces como generalmente ocurre en los problemas reales, ni siquiera se conoce la distribución de donde provienen dichas observaciones [2]. El interés principal en las series temporales, es estudiar cómo se producen los cambios de esa variable con respecto al tiempo, todo el análisis se desarrolla con el fin de predecir el posible futuro de la variable temporal. Existen varias tareas que se realizan en el análisis de las series de tiempo [7], estas son:

1. Descripción. La finalidad es obtener estadísticas y grafos que expliquen la serie temporal.
2. Análisis e interpretación. El objetivo es encontrar un modelo que describa acertadamente los datos en dependencia del tiempo.
3. Predicción. En esta tarea, conocidas las muestras u observaciones de la serie temporal, se desea predecir el valor siguiente o los valores siguientes de la serie temporal.
4. Control. La meta es ajustar varios parámetros de control, de modo que la serie se ajuste a lo que ocurre en la realidad.
5. Ajuste. Para un modelo lineal, los errores formarán parte de la serie temporal de las observaciones que se encuentran correlacionadas y se desea ajustar la varianza.
6. Filtrado. Se busca en el filtrado realizar la separación del ruido de la serie temporal original ya sea mediante filtrados lineales o no lineales. El filtro de Kalman es uno de los métodos más usados, es un método para actualización de las estimaciones de parámetros instantáneamente cuando una nueva observación se introduce, basado en la verosimilitud de los datos actuales, no necesita reestimar una gran cantidad de parámetros.

Las series temporales, pueden ser continuas (se conoce la variable de estudio en cualquier instante de tiempo) o discretas (los valores de la variable sólo se conocen en determinados

instantes de tiempo).

Se dice que una serie temporal es estacionaria, si la función de distribución de probabilidad que define la serie temporal se mantiene inalterable al desplazar la variable temporal un cierto periodo de tiempo, generalmente los datos de la vida real son no estacionarios. La base de datos de la serie temporal de SEMARNAT [6], cumple con ser una serie temporal discreta, no estacionaria, y no se conoce su función de distribución de probabilidad. A continuación se explican cada una de las técnicas de predicción implementadas.

### 3.1. Media móvil

La media móvil, no toma en cuenta las observaciones lejanas en el tiempo, es decir, se calcula la predicción  $P_{t+1}$  como la media de una cantidad fija de observaciones anteriores  $(t), (t-1), \dots, (t-k)$  con  $k > 1$ . El cálculo se realiza de la manera siguiente [1]:

$$P_{t+1} = \frac{1}{k} \sum_{i=t-k+1}^t Y_i, \quad (1)$$

donde  $Y_i$  son las observaciones.

Si utilizamos un valor de  $k$  muy pequeño, entonces, la media móvil tiende a repetir la serie temporal en su pasado más inmediato, pero si utilizamos un valor muy grande de  $k$ , entonces la media móvil suavizará demasiado a la función de predicción. Por lo que se puede concluir que el valor de  $k$  no debe ser muy pequeño ni muy grande con la finalidad de cometer menos error en la predicción.

### 3.2. Ajuste polinomial

El método de ajuste polinomial consiste en usar un polinomio de grado  $n$  de la forma [2]:

$$f_n(x) = \sum_{i=0}^n a_i x^i, \quad (2)$$

donde los parámetros a ajustar son:  $a_0, a_1, \dots, a_n$ . El objetivo es encontrar los valores de los parámetros que mejor ajusten a las observaciones  $Y_0, \dots, Y_m$  en el sentido de los mínimos cuadrados lineales, es decir, se resuelve el problema:

$$\text{minimizar}_{a_0, \dots, a_n} \frac{1}{2} \|f_n(x) - Y\|_2^2 = \frac{1}{2} \sum_{i=0}^n (f_n(x_i) - Y_i)^2, \quad (3)$$

donde  $\|\cdot\|_2$  denota la norma 2, definida para un vector  $x = [x_1, \dots, x_n]$  como la raíz cuadrada de  $x_1^2 + \dots + x_n^2$ . Observe que la función objetivo depende de los parámetros  $a_0, a_1, \dots, a_n$ . Solicitando que el gradiente de la función objetivo sea cero, se obtiene el sistema de ecuaciones lineales conocidas como ecuaciones normales. Resolviendo este sistema de ecuaciones lineales se obtienen los parámetros  $a_0, a_1, \dots, a_n$ . Una vez encontrados los parámetros, se utiliza el polinomio  $f_n(x)$  para realizar las predicciones. Es importante mencionar que si el grado del polinomio de ajuste es grande, entonces, se produce gran oscilación en la función de predicción, por lo que, se concluye que el grado del polinomio de ajuste no debe ser muy grande. La característica principal del ajuste polinomial es que no necesariamente pasa sobre las observaciones.

### 3.3. Interpolación usando transformada de Fourier

El método de interpolación con transformada de Fourier consiste en expresar las observaciones  $Y_i, i = 1, \dots, N$ , como una combinación lineal de exponenciales complejas de periodo  $N$  [3]:

$$Y_n = \sum_{k=\langle N \rangle} a_k e^{jk \frac{2\phi}{N} n}, \quad (4)$$

donde

$$a_k = \frac{1}{N} \sum_{k=\langle N \rangle} Y_k e^{-jk \frac{2\phi}{N} n}, \quad (5)$$

son las constantes de la combinación lineal, conocidas como los coeficientes de Fourier y la sumatoria es tomada sobre un periodo  $N$  de la exponencial compleja. A la ecuación (4) se le conoce como la serie inversa de Fourier y a la ecuación (5) se le conoce como la transformada de Fourier. Para predecir cualquier valor se utiliza la ecuación (4). La transformada de Fourier es un método de interpolación, por lo que la característica principal de tal método es que pasa exactamente sobre las observaciones.

### 3.4. Suavizado exponencial

El método de suavizado exponencial más simple, únicamente necesita estimar un parámetro, es decir, se calcula la predicción  $P_{t+1}$  como [4]:

$$P_{t+1} = \alpha Y_t + (1 - \alpha)P_t, \quad (6)$$

donde  $\alpha$  toma valores entre 0 y 1. De forma que la predicción se obtiene asignando a la última observación un peso  $\alpha$  y un peso de  $(1 - \alpha)$  a la predicción anterior. La ecuación anterior se puede extender de modo que se obtiene lo que se conoce como el alisado exponencial:

$$P_{t+1} = \sum_{i=0}^{t-1} \alpha(1 - \alpha)^i Y_{t-i} + (1 - \alpha)P_t. \quad (7)$$

El suavizado exponencial, asigna a las observaciones distintos pesos, estos pesos decrecen de modo exponencial, de modo que, las observaciones más recientes tienen más peso en la predicción que las observaciones más alejadas. Esta técnica trata de ajustarse a la evolución de la serie temporal. Para el caso del parámetro  $\alpha$ , se tiene que, si se asigna un valor de  $\alpha$  muy cercano a cero, entonces los datos de inicialización tienen un papel determinante en la predicción, mientras que, si  $\alpha$  es cercano a uno, entonces la predicción utiliza la observación más reciente como predicción. El valor de  $\alpha$  se encuentra generalmente a prueba y error, y cambia de una aplicación a otra. En general, el suavizado exponencial es un buen método de predicción cuando se aplica a series temporales que no muestran alguna tendencia o bien que son series temporales no estacionarias [2].

## 4. Aplicación

El objetivo de este trabajo, es predecir la precipitación media anual para el estado de Puebla, usando para ello la base de datos de la SEMARNAT que contiene datos históricos desde 1941 al 2011 [6]. Los métodos media móvil, ajuste polinomial, interpolación por transformada de Fourier y suavizado exponencial son usados para predecir el año 2012.

## 5. Resultados y discusión

Las implementaciones de los métodos de predicción explicados en la sección 3, fueron realizadas bajo el lenguaje de programación OCTAVE [5]. Se utilizó para realizar los experimentos una computadora con procesador INTEL Core 2 QUAD a 3.0Ghz con 2GB de memoria RAM.

En este trabajo, se realizaron dos experimentos, el primer experimento tiene el objetivo de encontrar los valores óptimos de todos los parámetros de ajuste que usan cada uno de los métodos de predicción implementados, en donde, se usa como medida de error el valor absoluto entre el valor que obtiene el método de predicción y el valor real conocido. La finalidad del segundo experimento, es realizar pruebas que nos ayuden a comprobar qué método de predicción funciona mejor cuando predicen la precipitación media anual para el estado de Puebla.

### 5.1. Experimento 1

El método media móvil, sólo depende de un parámetro  $k$ . Para obtener el valor óptimo de  $k$ , se buscó el mejor valor entero de  $k$  sobre el rango de valores de 1 a 50, que diera menor error en la predicción, se encontró que el valor óptimo fue para  $k$  igual a 4.

El método ajuste polinomial, depende sólo del parámetro  $k$ , por lo que se buscó el grado del polinomio adecuado sobre el rango de valores de 1 a 50, el mejor grado encontrado del polinomio fue de 4, que es cuando se comete menos error en la predicción.

El mejor valor encontrado del parámetro  $\alpha$ , para el método suavizado exponencial fue de 0.22, mismo que se eligió de los valores entre 0 y 1, iniciando con 0 y tomando incrementos de .01 hasta llegar a 1.0 y que diera menor error en la predicción.

Es importante mencionar que, el método de interpolación por transformada de Fourier no necesita ajustar parámetros y eso lo hace un método muy importante en la predicción de series temporales.

### 5.2. Experimento 2

En este experimento, se realizaron varias pruebas, las cuales se explican a continuación. La figura 1, muestra los resultados obtenidos por el método media móvil, cuando se predice la precipitación media anual para el estado de Puebla del 2006 al 2011, se mide el error absoluto entre el valor real y el obtenido por el método, finalmente se predice la precipitación media anual para el año 2012 cuyo valor real es desconocido. La Figura 2, muestra los resultados obtenidos por el método ajuste polinomial, cuando se predice del 2006 al 2011 y la precipitación media para el año 2012 cuyo valor real se desconoce.

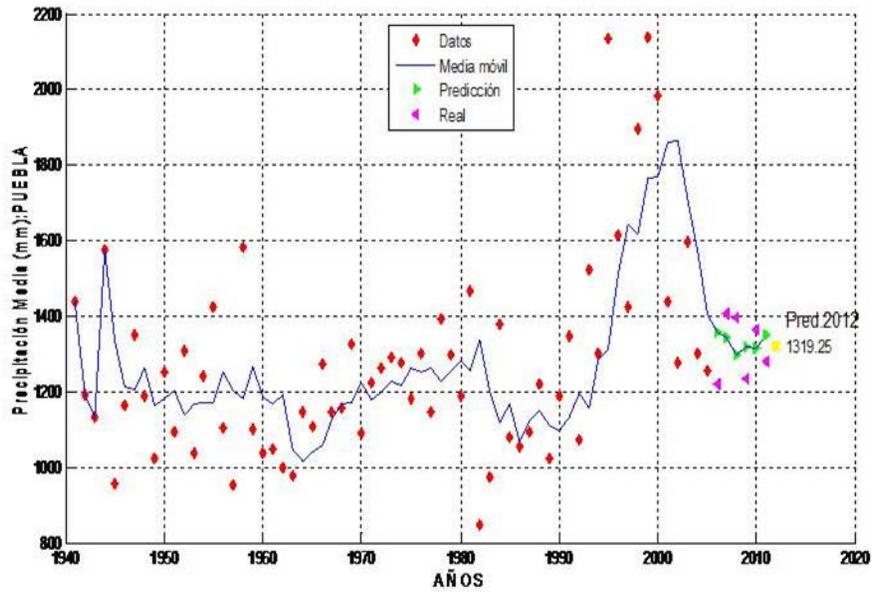


Figura 1. Predicción por media móvil. Ambientales 2012. Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla (FCFM-BUAP).

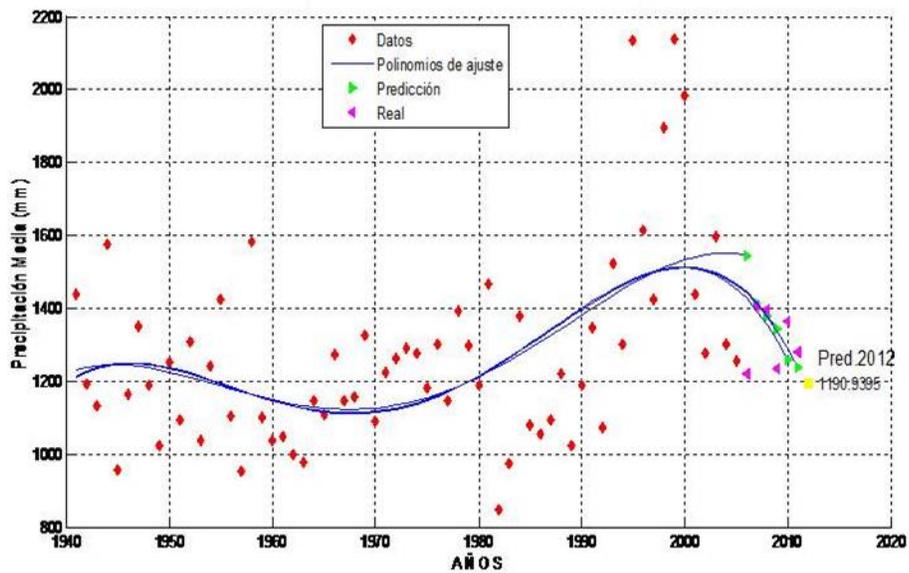


Figura 2. Predicción por polinomios. Ambientales 2012. FCFM-BUAP.

Las Figuras 3 y 4, muestran respectivamente los resultados de los métodos interpolación usando la transformada de Fourier y suavizado exponencial, cuando se predice del 2006 al 2011 y la precipitación media para el año 2012 cuyo valor real se desconoce.

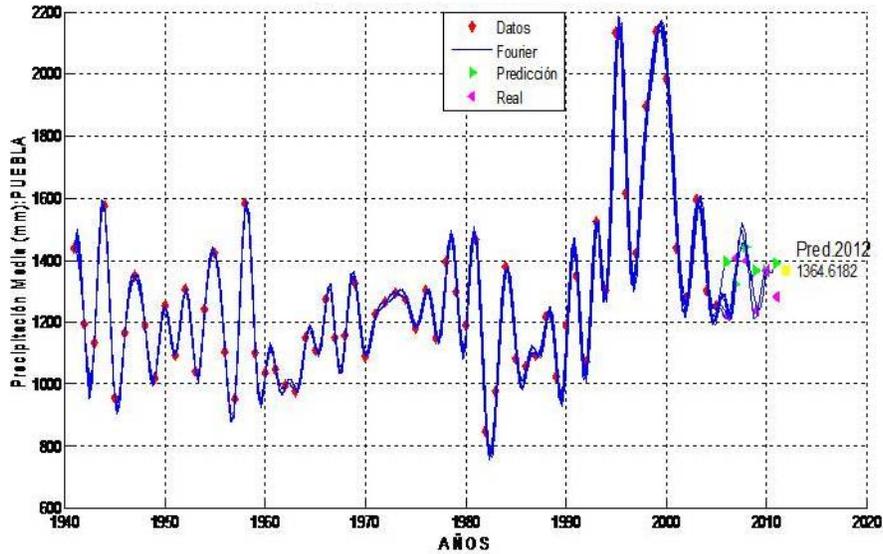


Figura 3. Predicción por transformada de Fourier. Ambientales 2012 FCFM-BUAP

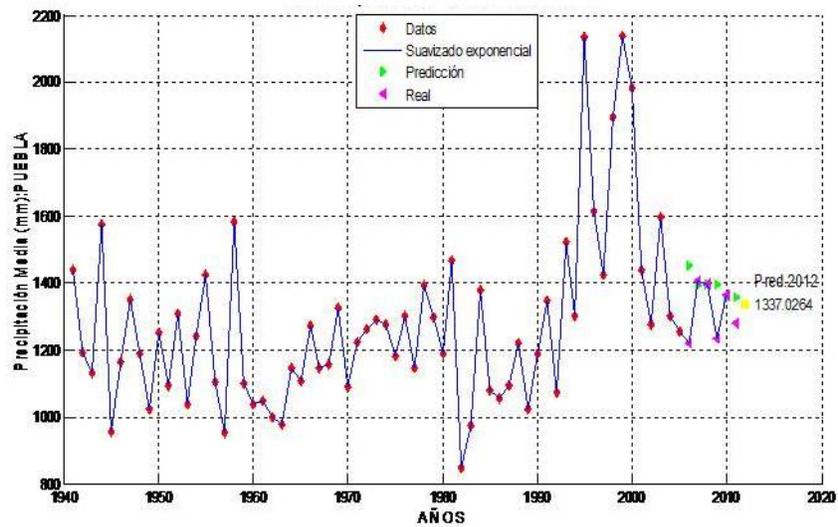


Figura 4. Predicción por suavizado exponencial. Ambientales 2012. FCFM-BUAP.

El Cuadro 1, muestra el error absoluto (entre el valor conocido y el predecido) cometido en la predicción que realiza cada uno de los métodos implementados cuando se les utiliza para predecir la precipitación media anual en el estado de Puebla del 2006 al 2011, es decir, una predicción de 6 años.

Cuadro 1. Predicción de la precipitación media de 2006 a 2011.

<b>Método de predicción</b>	<b>Error absoluto</b>
Media móvil	503.75
Ajuste polinomial	604.70
Interpolación por transformada de Fourier	542.68
Suavizado exponencial	493.39

El Cuadro 2, muestra el error absoluto cometido en la predicción que realiza cada uno de los métodos implementados cuando se les utiliza para predecir la precipitación media anual en el estado de Puebla del 2001 al 2011, es decir, una predicción de 11 años.

Cuadro 2. Predicción de la precipitación media de 2001 a 2011.

<b>Método de predicción</b>	<b>Error absoluto</b>
Media móvil	2045.8
Ajuste polinomial	3492.8
Interpolación por transformada de Fourier	1223.8
Suavizado exponencial	1851.1

El Cuadro 3, muestra el error absoluto cometido en la predicción que realiza cada uno de los métodos implementados cuando se les utiliza para predecir la precipitación media anual en el estado de Puebla del 1991 al 2011, es decir, una predicción de 21 años.

Cuadro 3. Predicción de la precipitación media de 1991 a 2011.

<b>Método de predicción</b>	<b>Error absoluto</b>
Media móvil	4783.5
Ajuste polinomial	6102.1
Interpolación por transformada de Fourier	4332.7
Suavizado exponencial	4861.3

El Cuadro 4, muestra el error absoluto cometido en la predicción que realiza cada uno de los métodos implementados cuando se les utiliza para predecir la precipitación media anual en el estado de Puebla del 1981 al 2011, es decir, una predicción de 31 años.

Cuadro 4. Predicción de la precipitación media de 1981 a 2011.

<b>Método de predicción</b>	<b>Error absoluto</b>
Media móvil	6343.0
Ajuste polinomial	7953.3
Interpolación por transformada de Fourier	6566.6
Suavizado exponencial	6556.8

Se observa en los Cuadros 1 - 4, que conforme aumenta el número de años a predecir difiere la técnica de predicción a utilizar. Para los últimos 5 años se recomienda usar el método de suavizado exponencial, para predecir desde 10 hasta 20 años se recomienda usar el método de interpolación por transformada de Fourier y finalmente para predecir 30 años o más se recomienda usar el método media móvil.

Finalmente en la Cuadro 5, se muestra la predicción que hacen cada uno de los métodos implementados de la precipitación media anual en el estado de Puebla para el año 2012.

Cuadro 5. Predicción de la precipitación media para el año 2012.

<b>Método de predicción</b>	<b>Estimación para el año 2012</b>
Media móvil	1319.25
Ajuste polinomial	1190.98
Interpolación por transformada de Fourier	1364.61
Suavizado exponencial	1337.02

## 6. Conclusiones y trabajo futuro

1. El método de predicción, de la precipitación media anual para el estado de Puebla que comete menor error, tomando el error como el valor absoluto entre el valor obtenido por el método y el valor real conocido, es el método de interpolación mediante transformada de Fourier, seguido del método de suavizado exponencial y del método de media móvil.
2. El método de predicción que comete mayor error en todas las pruebas del experimento 2, es el de ajuste polinomial, por lo que se recomienda no usar polinomios para predecir la precipitación media anual en el estado de Puebla.
3. El método de predicción que no requiere de ajuste de parámetros y es muy bueno como método de predicción para la precipitación media anual para el estado Puebla es el de interpolación mediante transformada de Fourier, por lo que se recomienda su uso.
4. Como trabajo futuro, se propone identificar una cierta cantidad de cúmulos, usando estos en lugar de los datos observados e identificar el periodo de la niña y el niño con el fin de agregar esta información a los métodos de predicción implementados.

## Referencias

- [1] Peter J. Brockwell, Richard A. Davis (2010), Introduction to Time Series and Forecasting, Edition:2nd, Springer.
- [2] Jonathan D. Cryer, Kung-Sik Chan (2010), Time Series Analysis: With Applications in R (Springer Texts in Statistics), Springer.
- [3] Robert M. Gray, Joseph W. Goodman (1995), Fourier Transforms An Introduction for Engineers, Kluwer Academic Publishers.
- [4] Robert H. Shumway, David S. Stoffer (2010), Time Series Analysis and Its Applications: With R Examples (Springer Texts in Statistics), Springer.
- [5] <http://www.gnu.org/software/octave> (último acceso Agosto de 2012)
- [6] [http:// www.semarnat.gob.mx](http://www.semarnat.gob.mx) (último acceso Mayo de 2012)
- [7] Gesine Reinert, Time Series,2010: [http://www.stats.ox.ac.uk/ reinert](http://www.stats.ox.ac.uk/reinert) (último acceso Noviembre de 2012)

# Recuperación de aceite quemado de motor aplicando dos métodos: fisicoquímico y microbiológico, como medio preventivo de la contaminación ambiental.

Pozos Gutiérrez Arturo<sup>1</sup> Castro Lino Alejandra<sup>1</sup>

Meléndez Balbuena Lidia<sup>1</sup>

Castro Caballero Leopoldo<sup>1</sup>

González Coronel Marco Antonio<sup>1</sup>

Guadalupe López Olivares<sup>1</sup>

Línea de investigación: Química: Ambiental.

<sup>1</sup>Benemérita Universidad Autónoma de Puebla,

Fac. Ciencias Químicas, Departamento de Química Inorgánica.

*Ave. San Claudio y Rio Verde, Col. San Manuel, C.U,*

*Puebla, Pue., 72000, México.*

*alcastro1228@yahoo.com.mx*

## Resumen

Para el aprovechamiento energético de los aceites quemados de motor se pueden seguir dos caminos diferentes en función de sus aplicaciones. Primero.- El aceite quemado se destina como combustible en instalaciones con alta potencia térmica a altas temperaturas y de gran consumo de combustible. Segundo.- Puede ser usado en la aplicación de tratamientos fisicoquímicos y biológicos más complejos con el fin de fabricar un combustible que pueda tener un espectro de utilización más amplio en instalaciones con menos potencia térmica o en motores de combustión y calderas. En la realización del presente trabajo se observa que el aceite quemado es vertido indiscriminadamente a la naturaleza, provocándole daños irremediables al ambiente por lo que esta investigación concentra sus esfuerzos en obtener aceite de motor limpio, haciendo uso de los siguientes métodos: el fisicoquímico o llamado ácido arcilla y biológico o mejor conocido como microbiológico utilizando una cepa de *Pseudomonas putida* sp por sus capacidades degradativas en hidrocarburos para la recuperación del mismo y evitar que este siga contaminando.

**Palabras clave:** Recuperación, aceite, métodos, ambiental.

## 1. Introducción

Los aceites lubricantes se distinguen entre sí, según sus propiedades o según el comportamiento que tengan dentro de las máquinas. Un buen aceite lubricante, a lo largo de su utilización no debe formar excesivos depósitos de carbón, ni tener tendencia a la formación de lodos, tampoco debe congelarse a bajas temperaturas. Sus propiedades más importantes que deben tener son: color, que nos da el grado de pureza o refinamiento, densidad, viscosidad que será la resistencia que este presenta a fluir, índice de viscosidad que representa la variación de este con la temperatura, punto de inflamación el cual es la temperatura cuando se forman vapores donde se inflaman en presencia de una llama, punto de congelación que es la temperatura a la cual pierde sus propiedades de fluido y ahora se convierte en sólido, acidez o mejor dicho acidez mineral que es el grado de refinamiento del aceite lubricante que puede presentar en una reacción ácida o alcalina, y por último índice de basicidad el cual es la propiedad que presenta de neutralizar los ácidos formados por la combustión dentro de los motores [2].

Los aceites lubricantes están formados por compuestos tóxicos que al quemarse, tirarse al suelo o a la basura y en un mayor perjuicio entren en contacto con el agua, contaminará cualquiera de las formas antes mencionadas [2].

En nuestro país se usan aproximadamente setecientos cincuenta y siete millones de litros de aceite lubricante, que después de ser usados son indebidamente desechados al ambiente llámese, aire, agua o tierra. En 3.785 Litros de aceite quemado contamina 3,785,000 litros de agua que es la cantidad que usan cincuenta personas en un año. Si se tira en el suelo por la penetración al mismo destruye al humus vegetal, acabando con la fertilidad de la tierra. Cuando cinco litros de aceite lubricante es quemado contamina cerca de 2204 litros de cuerpos de aire que es la cantidad que respira una persona en tres años. Si todo el aceite que se utiliza en nuestro país de forma indebida se le diera un trato correcto, no dependeríamos tanto de los recursos no renovables [1].

Los procesos de regeneración: fisicoquímicos y biológicos que mediante distintos tratamientos del residuo, permiten recobrar el material de las bases lubricantes presentes en el aceite original, de manera que resulten aptas para su reformulación y utilización.

Todo residuo o desecho que pueda causar daño a la salud o al ambiente es considerado como un residuo peligroso, fundamento por el cual los gobiernos tienen la responsabilidad de promover la adopción de medidas para reducir al máximo la generación de estos desechos, así como establecer políticas y estrategias para su manejo.

## 2. Objetivo General

Recuperar aceite lubricante limpio a partir de aceite quemado de motor aplicando los métodos: fisicoquímico ácido arcilla y biológico por *pseudomonas putida* sp (aureoginosa) en la obtención de aceite limpio.

## 3. Marco Teórico

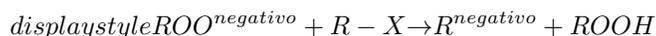
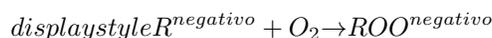
Un aceite lubricante es un líquido usado para disminuir la fricción entre dos superficies, estos son usados en el interior de motores donde las condiciones de operación degradan

sus compuestos cuyas características no permiten su utilización como lubricante [7].

El aceite lubricante está compuesto en general (excepto aceites sintéticos) por una base orgánica y aditivos, estos últimos utilizados para aumentar su rendimiento, eficiencia y vida útil. La base orgánica contiene en su mayoría compuestos aromáticos polinucleares (PNA) por sus siglas en inglés, algunos son estructuras de 4, 5, y 6 anillos aromáticos, los cuales son considerados cancerígenos. Los compuestos polinucleares aromáticos tienen altos coeficientes de fluidez, buena conductividad calorífica, capacidad de aislamiento eléctrica y flamabilidad [8].

El contacto con este tipo de compuestos origina riesgos para la salud, principalmente atacan el sistema hepático y enzimático. Por mencionar, en la industria eléctrica a través del material utilizado se producen reacciones de acné, así mismo en la industria incineradora por los vapores generados provocan intoxicaciones. En las industrias de la comunidad económica europea, en Estados Unidos y del Japón es un delito ecológico derramarlos a la naturaleza, es muy costosa su transportación así como su eliminación de las plantas que los incineran a altas temperaturas [1].

La reacción que a continuación se describe representa la oxidación que sufre el aceite quemado de motor. Donde  $R^{negativo}$ , representa al aceite nafténico, parafínico o aromático.



La reacción no se inicia hasta pasado un cierto periodo de inducción el cual corresponde al intervalo necesario para la formación de los peróxidos, que actúan como catalizadores, durante este periodo la oxidación del aceite es muy débil. Los hidrocarburos parafínicos se oxidan por los extremos de la cadena formando ácidos o cetoácidos corrosivos (pasando por los correspondientes productos intermedios). Con los hidrocarburos nafténicos se rompe la cadena y ocurre un proceso análogo al de los hidrocarburos parafínicos [11].

### 3.1. Antecedentes biológicos

El término biodegradación nos indica la extensión en la que se ha alterado una molécula, o en su caso señala si en el proceso muestra energía. Quizá resulte más apropiado hablar de bio transformación para referirse al proceso a través del cual un ser vivo modifica un compuesto sin llegar a mineralizarlo. En algunos casos el producto resultante puede ser incluso más tóxico y perjudicial que la sustancia de partida [19]. Los compuestos mineralizables se convierten en bióxido de carbono ( $CO_2$ ), agua y formas inorgánicas por la acción de los seres vivos, predominantemente por los microorganismos y como parte del constante reciclaje de los átomos de carbono. La mineralización de un compuesto implica su alteración estructural y la formación de intermediarios metabólicos que pueden servir de elementos estructurales de la célula o del combustible al oxidarse. Los intermediarios pueden convertirse en distintos compuestos orgánicos antes de su combustión final. En el proceso de transformación de un compuesto en otro puede sufrir la pérdida de uno o más de sus elementos estructurales o de sólo una reordenación de sus átomos [20]. El principal problema en el método biológico es la baja en la tasa de degradación cuando disminuye la concentración de los contaminantes (compuestos a degradar), en la mayoría de los casos

esto es ocasionado por la eliminación de co-sustratos necesarios para la degradación de otros contaminantes por medio del co-metabolismo. Esto lleva a diseñar procesos basados en reactores controlados, para una correcta biodegradación del aceite y no simplemente agregar el inoculo a los aceites para degradarlos [21]. El segundo tratamiento que se le ha dado es a través de microorganismos aerobios en este caso la pseudomona putida sp (aureoginosa) que puede biodegradar y después de este proceso recuperar la base lubricante del aceite así como desechos acuosos para darles un uso adecuado. La llamada biodegradación o mejor conocida como biotransformación microbiana es una técnica que permite utilizar un amplio espectro de bacterias. Por medio de pruebas bioquímicas se ratifica la identificación de la misma cepa contando con una pureza del noventa y nueve por ciento la cual garantiza el proceso de biodegradación [22].

#### 4. Metodología

En el diagrama de flujo No. 1 se presenta el mecanismo realizado para la preparación del medio filtrante.

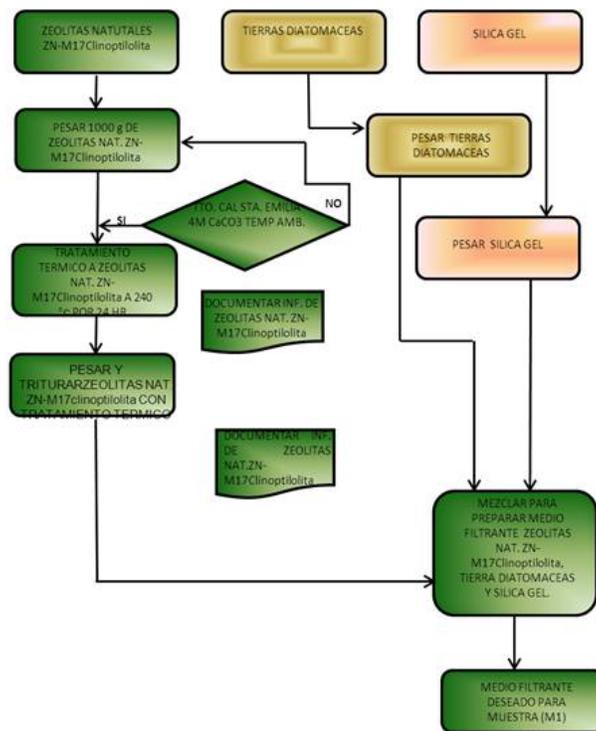


Figura 1: Preparación del medio filtrante

#### 4.1. Medios filtrantes

La filtración es una de las técnicas empleadas para separar partículas sólidas de un líquido. Haciendo pasar el líquido a través de un medio filtrante, sobre el cual se depositan los sólidos. Las características del medio filtrante debe mantener los sólidos a filtrar dando lugar a un filtrado claro, no obstruirse, ser químicamente resistente y tener resistencia física para soportar las condiciones del proceso, que como medio filtrante se debe desprender fácilmente del recipiente que lo contiene y finalmente debe de ser de costo moderado. Para este trabajo se emplearon zeolitas naturales ZN M-17(clinoptilolita), desecantes de silica gel y tierras diatomáceas [16]. En la Figura 1 se presenta el tipo de zeolitas utilizadas en este trabajo con sus características físicas.



Figura 2: Preparación de Zeolita ZN M-17 (Clinoptilolita) de granulometría menor a 0.9 mm. sin tratamiento térmico.

**Zeolitas.** La estructura altamente porosa de las zeolitas puede capturar partículas contaminantes de hasta 4 micras. Las zeolitas están cargadas negativamente de forma natural, por lo que pueden adsorber cationes, como metales pesados, hidrocarburos (alcanos y alquenos) [16]. Las zeolitas naturales ZN M-17(clinoptilolita) con una granulometría menor a 0.9 mm., densidad de 1.7684 g/mL. se trató con tratamiento térmico a 240 A° C con óxido de calcio 4 M. (Cal Santa Emilia).

En la Figura 3 a, b y c se aprecia la preparación del medio filtrante que dio los resultados para obtener la el aceite lubricante limpio, se agrega primero 5 gramos de zeolitas ZN M-17 clinoptilolita, se mezclan 5 gramos silica gel y 5 gramos de tierras diatomáceas para terminar de presentar el medio filtrante.



Figura 3: Preparación del medio filtrante a, b y c.

El diagrama de flujo No. 2 presenta el mecanismo realizado por medio del método fisicoquímico. El cual nos muestra como se prepararon las diferentes combinaciones de zeolitas naturales ZN M-17, silica gel y tierras diatomáceas.

Se toma un volumen de 25 mL de aceite quemado y se lleva a evaporación calentado

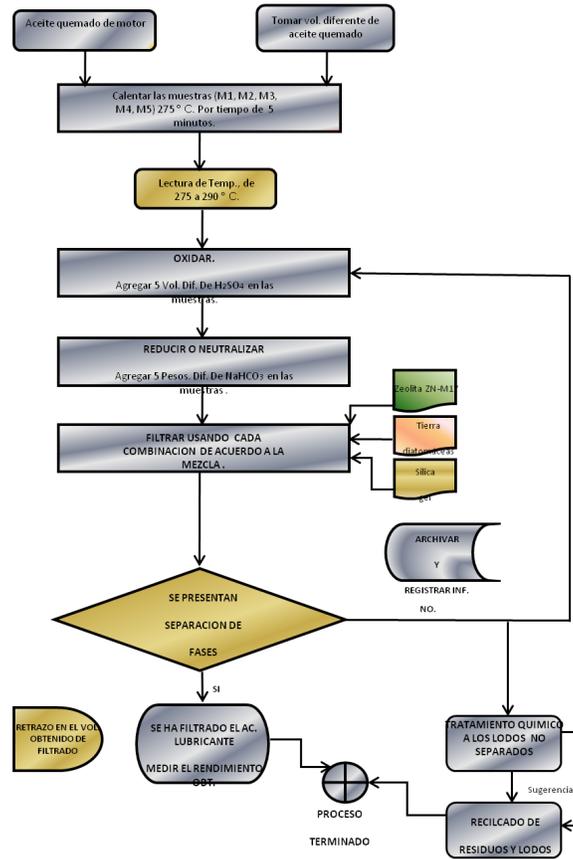


Figura 4: Método fisicoquímico

a una temperatura de 270 celsius durante 5 minutos para eliminar agua e hidrocarburos que se encuentran presentes en la muestra, posteriormente a la muestra se le agrega 50 mL de  $H_2SO_4$  (ácido sulfúrico) concentrado al 99.8% con agitación lenta obteniéndose así una mezcla muy reactiva de acción vigorosa por el contacto entre el aceite y el ácido. En seguida se le agrega 10 g de  $NaHCO_3$  (carbonato de sodio) para neutralizar y pasar a la siguiente fase de filtrado. De este filtrado se obtiene un volumen de 45 mL de aceite acido limpio el cual representa un rendimiento del 60% en volumen. El medio filtrante usado fue una mezcla de 5 gramos de zeolita (clinoptilolita) ZN M17, 5 gramos de silica gel y 5 gramos tierras diatomáceas los cuales fueron vertidos en un embudo de vidrio ver figura 4.

La zeolita que se utilizó en esta parte fue de tipo natural ZN M17 (Clinoptilolita) con una granulometría menor de 0.9 mm, de una densidad de 1.7684 g/mL a la cual se le dio un tratamiento térmico con oxido de calcio (Cal Santa Emilia) a 4 Molar. Obteniéndose una separación de fases como se muestra a continuación. En la figura 5 se observa el filtrado de aceite limpio el cual se ha obtenido usando un medio filtrante de mezcla de



Figura 5: Toma de muestras de aceite usado en volumen de 25 mL con calentamiento a 270 C adicionándole 50 mL de H<sub>2</sub>SO<sub>4</sub> concentrado al 99.8 %

tierras diatomáceas, zeolita ZN-M17, y silica gel. Obteniéndose una separación de fases



Figura 6: Separacion de fases

como se muestra a continuación.

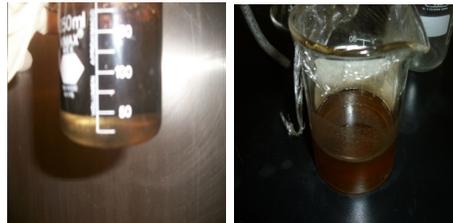


Figura 7: Se observa el filtrado de aceite limpio el cual se ha obtenido usando un medio filtrante de mezcla de zeolitas ZN M-17, silica gel y tierras diatomáceas.

El proceso tiene un rendimiento global de 60 % en volumen. A continuación se muestran en la Tabla 1 las pruebas que se realizan al aceite quemado con diferentes mezclas de medio filtrante.

En la Figura 7. Se representan estos datos donde se aprecia el volumen de ácido sulfúrico como de aceite usado y su posterior neutralización con bicarbonato de sodio. De los valores obtenidos y graficados a quince muestras se observa que las condiciones ideales donde existe las mejores condiciones de filtrado de muestra se obtiene en los puntos siete y doce que se muestran en el gráfico.

Numerodemuestra	Vol. DeaceitemL	Vol. H <sub>2</sub> SO <sub>4</sub> mL	PesoengNaHCO <sub>3</sub>
1	150	150	8.9725
2	60	60	5
3	60	25	5
4	50	10	1
5	50	10	4.035
6	25	15	2.725
7	25	50	8.9621
8	20	10	1
9	20	15	4.035
10	15	10	2.725
11	15	10	3.0025
12	25	50	8.9621
13	50	10	1
14	50	10	4.035
15	25	10	8.0025

Cuadro 1: Método fisicoquímico

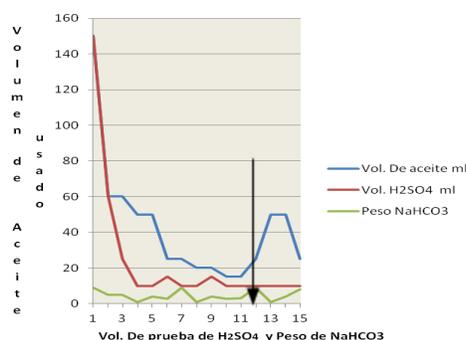


Figura 8: Método biológico o de degradación microbiana.

Para la obtención de la cepa se preparan las condiciones adecuadas para su desarrollo y crecimiento de las colonias bacterianas tal como lo es la *Pseudomonas putida* sp, para la identificación de esta bacteria es por medio de una Galería API 20 NE (No entero bacterias). Esta bacteria es de la clasificación no fermentadora la cual debe permanecer en refrigeración a temperatura de 4 a 5 °C. Las condiciones para examinar las colonias son bajo luz ultravioleta a 260nm. Se considera un resultado positivo la observación de fluorescencia, que es un pigmento de color amarillo, amarillo-verdoso fluorescente que rodea la colonia o que se extiende por el medio de cultivo debido a fenómenos de difusión. Esta cepa para madurar requiere de cultivarse por 24 horas. Para tal efecto, se hizo uso de un densímetro con escala de nefelómetro de Mac Farland. Así mismo se llevó a cabo la prueba positiva en la obtención de la bacteria de oxidasa, a partir del uso de sensibilizadores, los cuales contienen oxalato de dimetil-para-fenilendiamina, el cual es el sustrato de la enzima oxidasa. Los mos que producen la enzima oxidasa se evidencian porque en presencia de oxígeno atmosférico y citocromo c, se oxida el sustrato presente en los discos a un compuesto de color rojo fucsia. Después de pasadas las 24 horas la bacteria fue cultivada con éxito así como se entregó la cepa pura la cual fue ratificada por pruebas bioquímicas y respaldada por el sistema de identificación bacteriológica semi-automatizada mini API. Esta bacteria producida en caldo biológico se preparó en 5 tubos de ensayo en volumen de 1 ml. quemado para la prueba de biodegradación colocando

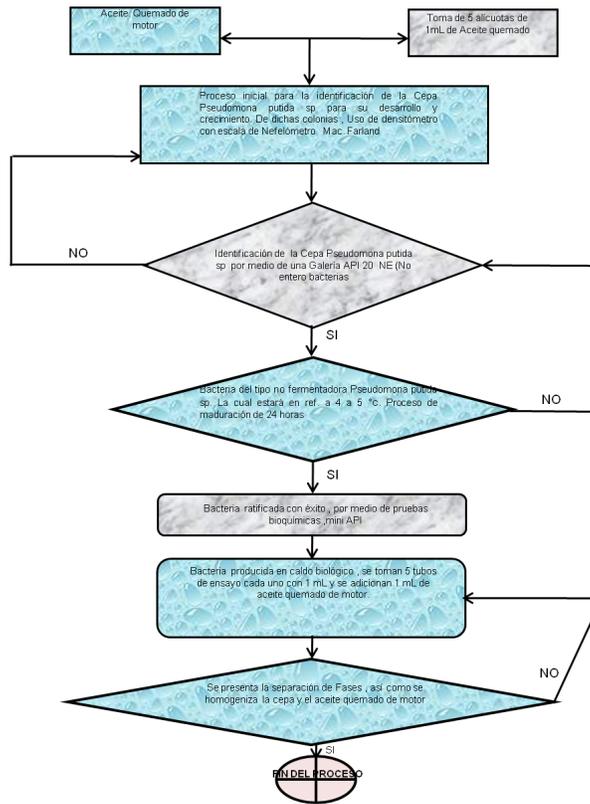


Figura 9: Método biológico o degradación.

estos tubos dentro de una incubadora para mantener las condiciones de crecimiento de la cepa e iniciar el proceso de biotransformación del aceite quemado con agitación diaria a temperatura de 36 °C para evitar que la cepa muera, este proceso se prolonga por lapso de treinta días para hacer notar los cambios biológicos que transforman el aceite y observar que finalmente la cepa cumple con el proceso de biotransformación notándose así separación del aceite limpio.

## 5. Resultados y discusión

A continuación se representan las tablas comparativas de rendimiento lo cual permite observar que el método fisicoquímico es más eficiente que el método biológico. Método fisicoquímico: Tabla de resultados del método fisicoquímico.

Método biológico: Resultados.



Figura 10: Método biológico para la toma de 1 ml de cepa de pseudomona putida sp dentro de un campo de esterilidad.

Vol. de aceite mL	Vol. $H_2SO_4$ mL	Peso en g $NaHCO_3$
150	150	9
60	60	5
60	25	5
50	10	1
50	10	4.035
25	15	2.725
25	50	9
20	10	1
20	15	4.035
15	10	2.725
15	10	3.0025
25	50	8.9621
50	10	1
50	10	4.035
25	10	8.0025

Cuadro 2: Tabla de Resultados del método fisicoquímico

Prueba A	Vol.Ac	Vol. Cepa "P"	T, C	Num. días.	Cambios 1a. etapa	Cambios 2da. etapa
1	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada
2	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada
3	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada
4	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada
5	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada
Prueba	Vol.Ac*	Vol. Cepa "P"	T,C	Num. días.	Cambios 1. etapa	Cambios 2da. etapa.
6	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada
7	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada
8	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada
9	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada
10	1 mL	1 mL	36	30	Se mezclan	Se Biodegrada

Cuadro 3: Resultados de la Prueba "A" y "B"

## 6. Conclusiones

1. Se concluye que a partir de la técnica fisicoquímica, se obtuvo un rendimiento del 60% en volumen para muestra de 75 ml por lo cual se establece que es un método

efectivo y de la misma forma los costos involucrados van conforme a los reactivos usados.

2. El aceite lubricante recuperado en este trabajo sirve como energético aunque si hay pequeñas modificaciones en sus propiedades físicas una de ellas es el color. La desventaja que presenta este método biológico es que solo se tomó un tipo de bacteria para la biodegradación del aceite quemado.
3. El método biológico fue un proceso bueno pero de bajo rendimiento por lo que se propone realizar más pruebas a las muestras de aceite en presencia de mínimo diez bacterias diferentes lo cual podría mejorar el rendimiento de las mismas ya que solo la presencia de una clasificación de bacteria *pseudomona putida* sp (aureoginosa) hizo que el rendimiento se viera afectado.
4. Con base a los resultados obtenidos en este trabajo se concluye que los métodos aquí utilizados son el comienzo para el desarrollo de una nueva tecnología, en la reutilización de los desechos automotrices como el aceite lubricante, no solo para la mejoría de la ecología local sino en todo nuestro país.
5. Es de vital importancia hacer conciencia a todos los consumidores de aceites lubricantes tanto de industria como de sectores gubernamentales y de nuestra sociedad en general de darle un lugar preponderante a la reutilización de este recurso, para a su vez disminuir la contaminación que provoca al aire, suelo y agua y al mismo tiempo no sobre explotar nuestros recursos no renovables.

## Referencias

- [1] Benavente, R.G; *Aceite lubricante usado*. Bravo energy. Chile, S. A. Boletín No. 2. Junio 1999.
- [2] [html.rincondelvago.com/ lubricantes/lubricante; Html](http://html.rincondelvago.com/lubricantes/lubricante.html). (31 Marzo 2011).
- [3] Baugton, H. A.; *Environmental assessment of used management methods*. In: *Environmental science technology*. Vol. 38, No. 2. 2004.
- [4] Dávila, M. M. y Figueruelo, J. (2002); *Química física del medio ambiente y de los procesos medio ambientales*. Editorial. Reverté, S. A.
- [5] López, R. A. y Díaz, H. I. (1993); *Algunas técnicas de purificación de aceites lubricantes usados*. Tesis de licenciatura no publicada, Benemérita Universidad Autónoma de Puebla.
- [6] Diario oficial de la federación.; (*Título II, capítulo VI, título III, capítulo III*) (1997, 21 de Enero). *Reglamento federal de seguridad, higiene y medio ambiente*. México; Méx.
- [7] Ruíz, E.; *Aceites lubricantes para motores a gasolina. Curso de educación continua, fundamentos básicos de lubricación*. Bogotá, Colombia. Mayo 1991.

- [8] Venkateswaran, Y. H.; *Sequential enrichment of microbial populations exhibiting enhanced biodegradation of crude oil*. Canadian journal of microbiology. Vol. 41 (1995),p. 767-775.
- [9] [www.monografias.com](http://www.monografias.com);
- [10] Albarracin, P.; *Lubricación industrial y automotriz*. Editorial Omega. 21 Febrero 2012.
- [11] <http://translate.google.com//theallnotes.com/branches> (1 de Abril 2011).;
- [12] [www.monografias.com](http://www.monografias.com).; *Química, propiedades físicas de los lubricantes*.(31 de marzo del 2011).
- [13] Quarry.P.2000;*Biodegradation of oils used. Equipment & Tech Handbook*. Vol. 92 Issue 6, p.126.
- [14] Florez Piedrahita, C. A.; *Transformación de los aceites usados para su utilización como energéticos en procesos de combustión. En: Escenarios Y Estrategias*. Bogotá. Diciembre de 2001 No 8.,p. 28 – 32.
- [15] <http://theallnotes.com/branches/CS/2ndsem/Chemistry/Unit3/9/>( 1 de Abril del 2011).:
- [16] <http://www.emyucatan.com/zeolita.html>.Zeolitas característica y propiedades. (Dic/08/2011).;
- [16] <http://www.auxifilt.com.ar/que-son-auxiliares-filtrantes.php>; *Que son los auxiliares filtrantes Tierras de Diatomeas Filtrantes - Auxiliares Filtrantes* . (28 Mayo del 2010).
- [18] Madigan, M. T., Martinako, M. J., Parker, J.; *Biología de los Microorganismos*. (10ma. Ed.). Southern. Illinois University Carbondale. Ed. Prentice Hall. 2006.
- [19] Lansing, M., Prescott, J. P. Harley Donald, A. K.; *Microbiología. (5a Edición)* .Mac Graw Hill. 2002.
- [20] Ward, O.; Singh, A. y Van Hamme. J. Accelerated.: *Biodegradation of petroleum hydrocarbon waste. En: Journal Industrial Microbiology Biotechnology*. (2003) Vol. 30:p 260-270.
- [21] Venkateswaran Y. H.; *Sequential enrichment of microbial populations exhibiting enhanced biodegradation of crude oil. En: Canadian Journal Of Microbiology*. Vol. 41 (1995),p.767-775.
- [22] Pantoja, J. L. M. Moreno, P. M.; *Qué se hace en España con los aceites usados. En : Ingeniería Química*. Enero 1995, p. 113-117.

# Teledetección del Carbono Orgánico en Suelos en la zona de Tetelilla, Teziutlán, Puebla.

Torre T. Edgardo<sup>1</sup>, Linares Fleites Gladys<sup>2</sup>, Sandoval S. María de Lourdes<sup>5</sup>  
Tenorio A. M. Guadalupe<sup>2</sup>, Peña M. Ricardo<sup>3</sup>  
Castelán V. Rosalía<sup>2</sup>, Rodríguez H. América<sup>4</sup>

<sup>1</sup> Estudiante del Doctorado en Ciencias Ambientales, BUAP.

<sup>2</sup> Prof-Inv. Posgrado Cs. Ambientales y Depto. en Cs. Agrícolas, BUAP.

<sup>3</sup> Centro de Química. BUAP.

<sup>4</sup> Universidad Autónoma de Guerrero.

<sup>5</sup> Facultad de Computación. BUAP

Fax 2295500. Ext 7351

14 Sur 6301. Ciudad Universitaria. Jardines de San Manuel.

Puebla, Pue. México. Ext 7355, CP 72470.

e-mail: etorres23@hotmail.com

## Resumen

La concentración de carbono orgánico del suelo (COS) es una propiedad fundamental del suelo para guiar aplicaciones agrícolas. En el presente trabajo se han utilizado datos de teledetección para estimar el contenido del COS a partir de índices de vegetación en la región RTP 105: Cuetzalan, definida por la CONABIO como prioritaria para la conservación. La metodología se basa, primero, en el procesamiento digital e interpretación de las imágenes que proceden del satélite Landsat-5 sensor TM, que proporciona imágenes de 8 bits en 7 y 8 bandas espectrales, tres de las cuales son visibles (banda 1, banda 2 y banda 3), y se usan para la determinación de índices de vegetación, tales como el NDVI, el SAVI y el RS. A partir de estos índices, se han comparado diferentes modelos de regresión, regresión por mínimos cuadrados ordinarios (OLS), regresión con componentes principales (RCP) y regresión paso a paso (RPP), desarrollándose modelos entre la variable dependiente COS y las variables independientes índices de vegetación. La técnica que brindó las predicciones más precisas fue la RPP.

*Palabras clave:* Cambio Climático, Análisis de Componentes Principales, Índices de Vegetación, Reflectividad.

## 1. Introducción

Uno de los problemas ambientales de mayor envergadura, a los que se está enfrentando la humanidad en la actualidad, es el cambio climático.

El *cambio climático* es un problema con características únicas, ya que es de naturaleza global, y su impacto es a largo plazo e involucra interacciones complejas a escala mundial entre procesos tanto naturales (fenómenos ecológicos y climáticos) como sociales, económicos y políticos (Martinez 2004) . Acuerdos internacionales, particularmente el Protocolo de Kyoto, ha incrementado la atención sobre las concentraciones del dióxido de carbono atmosférico y el balance global del carbono por ser tópicos de gran interés científico y de importancia política.

En este contexto se están ideando nuevas metodologías de investigación y se están utilizando nuevas herramientas de búsqueda de información (Chuivenco 2007) . Los instrumentos de teledetección son útiles en estudios de ecosistemas a gran escala, ya que permiten obtener información de las condiciones biofísicas de manera más eficaz que las mediciones tradicionales *in situ*.(Moslem *et al*, 2010).

Los estudios relacionados con los ecosistemas forestales en México y su relación con el impacto y la vulnerabilidad del cambio climático, han sido, hasta ahora muy generales. (Villers y Trejo, 2004). Es necesario, desarrollar estudios a nivel local, que permitan esclarecer el papel del suelo en el secuestro de carbono, ya que el carbono es un elemento fundamental de los compuestos orgánicos. Por otra parte, la concentración de carbono orgánico del suelo (COS) es una propiedad fundamental del suelo para guiar aplicaciones agrícolas (Sandoval *et al*, 2003).

Los modelos de regresión son de extraordinaria importancia para relacionar informaciones obtenidas tradicionalmente en el campo y en el laboratorio con información obtenida a través de imágenes satelitales. Dado que existen diversas técnicas, se ha considerado conveniente comparar algunos modelos de regresión, a saber, la regresión por mínimos cuadrados ordinarios (OLS), la regresión con componentes principales (RCP) y la regresión paso a paso (RPP), desarrollando modelos entre la variable dependiente COS y las variables independientes índices de vegetación, que permitan alcanzar buenas predicciones del COS en la zona de estudio y contribuir a la toma de decisiones de los especialistas en agronomía y en ciencias ambientales.

En el presente trabajo se desarrollan ideas fundamentales sobre teledetección, índices de vegetación y modelos de regresión y se aplican en la selección de un modelo adecuado a la predicción del COS en la zona mencionado. Finalmente, se brindan conclusiones y referencias

## 2. Objetivo General

El objetivo de este estudio es determinar algunos índices de vegetación de banda ancha a partir de teledetección ambiental en la región de Tetelilla, Teziutlán, Puebla, y relacionarlos con el Carbono Orgánico del Suelo.

### 3. Marco Teórico

Una breve explicación sobre los temas de teledetección e índices de vegetación, que pueden obtenerse a partir de imágenes satelitales, son necesarios para la comprensión de los resultados.

#### 3.1 Teledetección.

La Teledetección no es una ciencia que se le podría denominar pura. Es un compendio de ciencias y tecnologías que se han desarrollado a lo largo de los dos últimos siglos y que han dado lugar a una materia multidisciplinaria basada en el conocimiento de la Física y las Matemáticas y que se ha desarrollado, principalmente, gracias a los avances tecnológicos en Telecomunicación y Computación. Sólo la adecuada conjunción de todas ellas da lugar a la Teledetección. Ésta se podría definir como la ciencia y/o el arte de adquirir información sin contacto directo entre el captador y el *objetivo*.

El término Teledetección es una traducción del inglés «*Remote Sensing*», y se refiere, no sólo, a la captación de datos desde el aire o desde el espacio sino también a su posterior tratamiento. Una definición más formal la describe como la técnica de adquisición y posterior tratamiento de datos de la superficie terrestre desde sensores instalados en plataformas espaciales, en virtud de la interacción electromagnética existente entre la Tierra y el sensor, siendo la fuente de radiación bien proveniente del sol (teledetección pasiva) o del propio sensor (teledetección activa)..

Los términos plataformas, sensores y canales deben ser bien comprendidos. Un sensor es el aparato que reúne la tecnología necesaria para captar imágenes a distancia y que es transportado en una plataforma. Puede captar información para diferentes regiones del espectro y cada una de estas regiones se denomina canal o banda. Por ejemplo, Landsat es una plataforma que contiene dos sensores Landsat-TM (Thematic Mapper), y Landsat-ETM+, el primero de los cuales permite captar radiación en 7 bandas (azul, verde, rojo, 3 en el infrarrojo cercano y 1 en el infrarrojo térmico) y el segundo en 4 bandas (verde, rojo y 2 en el infrarrojo cercano). En este trabajo se utiliza información de esta plataforma y del sensor TM, en sus bandas visibles (banda 1 banda 2 y banda 3).

#### 3.2 Índices de vegetación

El empleo de índices, calculados a partir de la reflectividad en diferentes bandas, que nos indiquen la abundancia y estado de la vegetación, se basan en el comportamiento reflectivo peculiar de la vegetación. La signatura espectral característica de la vegetación sana muestra un fuerte contraste entre la baja reflectividad en el rojo y la alta reflectividad en el infrarrojo. Esta diferencia es tanto mayor cuanto mayor es la densidad de la vegetación y mejor su estado fitosanitario. En esta idea se basan la mayor parte de los índices de vegetación. El más conocido es el Índice Normalizado de Vegetación (NDVI) expresado como

$$NDVI = \frac{B_4 - B_3}{B_4 + B_3}$$

donde:

B4 : Banda 4 de la imagen de Landsat TM (reflectividad en el infrarrojo cercano)

B3 : Banda 3 de la imagen de Landsat TM (reflectividad en el rojo)

Este índice varía entre  $-1$  y  $1$ , lo que facilita su interpretación. Puede señalarse como umbral crítico para cubiertas vegetales un valor de NDVI de  $0.1$  y para la vegetación densa de  $0.5$ . Cualquier valor negativo implica ausencia de vegetación por lo que debieran reclasificarse a cero. (Rouse,1974)

Otros índices de vegetación que pueden obtenerse a partir de las bandas antes citadas son: el de Razón Simple (SR) y el Índice de Vegetación con Ajustes al Suelo (SAVI). La expresión del primero (Birth y Mcvey, 1968) es:

$$SR = \frac{B_3}{B_4}$$

El Índice de Vegetación con Ajustes al Suelo (Huete, 1988) incorpora a la ecuación un parámetro  $L$ , que generalmente se toma como  $L = 0.5$ . Su expresión es.

$$SAVI = (1 + l) \frac{B_4 - B_3}{B_4 + B_3}$$

## 4. Métodos de regresión: OLS, RPP y RCP

A continuación se explican tres métodos de regresión de amplia utilización en la práctica (Linares, 2006) El modelo de regresión para estos métodos puede escribirse como:

$$y = 1\beta_0 + X\beta + \epsilon$$

donde,

$y$  es un vector  $n \times 1$  de observaciones de la variable dependiente, es una constante desconocida

$X$  es una matriz  $n \times p$  que consiste de  $n$  observaciones de las  $p$  variables

$\beta_0$  es un vector  $p \times 1$  de coeficientes de la regresión (parámetros desconocidos)

$\beta$  es un vector  $p \times 1$  de coeficientes de la regresión (parámetros desconocidos), y

$\epsilon$  es un vector  $n \times 1$  de errores independientes e idénticamente distribuidos con media cero y varianza  $\sigma^2$ .

### 4.1 Regresión por mínimos cuadrados ordinarios. (OLS)

Cuando la matriz  $X$  tiene rango  $p$ , el estimador OLS se obtiene minimizando la suma de cuadrados de los residuos. Este estimador es un vector de dimensión  $p$  por  $1$  cuya

expresión es

$$b_{OLS} = (X'X)^{-1}X'y$$

son estimadores insesgados lineales con varianza mínima. Cuando las variables independientes están altamente correlacionadas,  $X'X$  es mal condicionada y la varianza de los estimadores OLS se hace muy grande, y por tanto, son estimadores muy inestables.

#### 4.2 Regresión paso a paso (RPP)

La RPP permite la re-especificación del modelo. Este procedimiento comienza con la hipótesis de que no hay regresores en el modelo además de la ordenada al origen. El primer regresor que se selecciona para entrar al modelo es el que tenga la máxima correlación simple con la variable respuesta, luego en cada paso se reevalúan todos los regresores que habían entrado antes al modelo, mediante sus estadísticas parciales F. Un regresor agregado en una etapa anterior puede volverse redundante, debido a las relaciones entre él y los regresores que ya están en la ecuación. Si la estadística F de una variable es menor que la F de salida, esa variable se elimina del modelo. En este método de regresión se requieren dos valores de corte, la F de entrada y la F de salida.

#### 4.3 Regresión con componentes principales (RCP)

La RCP es una de las maneras de tratar los problemas de mal condicionamiento de matrices. Básicamente lo que hace es obtener el número de componentes principales que brinda la variación máxima de X. Realmente es un método de regresión lineal en que la respuesta es regresada sobre los componentes principales de la matriz X.

La expresión del estimador en RCP es

$$b_{RCP} = V_m \alpha_m$$

donde  $V_m$  es una matriz que consiste de los primeros m vectores de norma unidad ,

$$\alpha_m = (Z'_m Z_m)^{-1} Z'_m y$$

$Z$  son las componentes principales y  $m$  es el número de estas componentes principales retenidas en el modelo. RCP da estimadores sesgados de los parámetros.

#### 4.4 Criterios de Comparación de modelos de regresión

Antes de comparar la capacidad predictiva de los modelos es útil introducir varias medidas del ajuste del modelo a los datos y de la potencia de predicción de esos modelos (Montgomery *et al*, 2004).

La Suma de Cuadrados de Error de Predicción, conocida como la estadística *PREES*, se considera una medida de lo bien que funciona un modelo de regresión *para predecir nuevos datos*. Se define como la suma de los residuales *PRESS* al cuadrado que son los

residuos que se obtiene entre el valor observado y el valor predicho de la  $i$ -ésima respuesta observada, basado en un ajuste de modelo con los  $n - 1$  puntos restantes de la muestra.

Con la estadística *PREES* se puede calcular otro estadístico conocido como  $R^2$  para la predicción, que se define como:

$$R_{prediccion}^2 = 1 - (PRESS/Sumadecuadradostotal)$$

Una aplicación muy importante de estos estadísticos es comparar modelos de regresión. En general un modelo con pequeño valor de *PREES* es preferible a uno con *PREES* grande. El  $R_{prediccion}^2$  se interpreta de manera similar al estadístico  $R^2$  utilizado usualmente para medir la bondad del ajuste del modelo a los datos: a mayores valores de éstos estadísticos mayor es la bondad del ajuste y mayor es la capacidad predictiva del modelo.

## 5. Aplicación

### 5.1 Localización de la zona de estudio

La zona estudiada corresponde a Tetelilla, Teziutlán, estado de Puebla, México, que se encuentra inmersa en la RTP 105: Cuetzalan, definida por la CONABIO como prioritaria para la conservación, y se ubica en las coordenadas extremas:  $19^{\text{grados}} 46^{\text{min}} 23^{\text{seg}}$  y  $20^{\text{grados}} 11^{\text{min}} 55^{\text{seg}}$  de latitud norte y  $97^{\text{grados}} 09^{\text{min}} 17^{\text{seg}}$  a  $97^{\text{grados}} 38^{\text{min}} 36^{\text{seg}}$  de longitud oeste. Está conformada por 28 municipios, compartidos entre los estados de Veracruz y Puebla y en ella predominan el bosque Mesófilo de Montaña, la selva alta perennifolia y las asociaciones de bosques de encino y pino.

### 5.2 Metodología

La metodología consiste en, primero, en el procesamiento digital e interpretación de las imágenes que proceden del satélite Lansat-5 sensor TM, (Thematic Mapper) que proporcionan imágenes de 8 bits en 7 y 8 bandas espectrales, tres de las cuales son visibles (banda 1, banda 2 y banda 3). Estas imágenes son especialmente adecuadas para la estimación de índices de vegetación, entre los que se encuentran el NDVI, el SAVI y el RS.

En base a la imagen satelital captada desde el espacio por el sensor TM se generó un mapa de la RTP 105. y se realizó la corrección geométrica para posteriormente analizarlas obteniendo la estimación de los índices de vegetación mencionados utilizando sistemas de información geográfica, Esta información se validó en campo con puntos georeferenciados con GPS.

A partir de estas determinaciones, y tomando en consideración la información sobre propiedades del suelo obtenidas en el laboratorio, fue posible llevar cabo diferentes técnicas de regresión. Dado que el uso principal de los modelos de regresión es la predicción de futuras observaciones utilizaremos la estadística *PREES* y el  $R_{prediccion}^2$  para seleccionar el mejor modelo (Linares *et al*, 2011). Los resultados que se presentan en cada método fueron obtenidos en MINITAB 15.

N	Media	Des. Est.	Mínimo	Máximo	Coef. Asimetría
24	39.13	21.44	12.48	87.00	0.68

Cuadro 1: Estadísticos descriptivos de COS

Valor propio	2.0873	0.8406	0.0721
Proporción	0.696	0.280	0.024
Proporción Acum.	0.696	0.976	1.000
VARIABLE	CP1	CP2	CP3
NDVI	0.384	0.906	0.177
SAVI	0.633	-0.398	0.605
SR	-0.673	0.143	0.726

Cuadro 2: Análisis de Componentes Principales de los índices de vegetación

### 5.3 Resultado y discusión

El Cuadro 1 muestra los estadísticos descriptivos más importantes de la variable COS tomada como respuesta en los modelos considerados. Puede observarse que el coeficiente de asimetría no sobrepasa el valor 1, luego puede considerarse una variable simétrica.

El Análisis de Componentes Principales de los índices de vegetación muestra que una sola dimensión es capaz de explicar casi el 70 por ciento de la variabilidad y que en esa dimensión (CP1) se destacan los índices SAVI y SR, oponiéndose ambos. El Cuadro 2 muestra los valores y vectores propios de la matriz de correlaciones, a partir de la cual se desarrolló el análisis.

A continuación se muestran los resultados obtenidos por cada método de regresión en el problema bajo estudio.

#### **Regresión por mínimos cuadrados ordinarios (OLS)**

El Cuadro 3 muestra los coeficientes del modelo de regresión OLS, los errores estándar de los coeficientes, el estadístico t de Student con sus correspondientes valores p y los factores de inflación de la varianza (VIF). Estos últimos indican problemas de multicolinealidad ya que todos son mayores que 1.

Predictores	Coeficientes	SE Coef.	Estadístico t	Valor de p	VIF
Constante	17.06	53.76	0.32	0.754	
No. Pixel NDVI	52.33	28.54	1.83	0.082	1.5
No. Pixel SAVI	23.11	66.88	0.35	0.733	6.5
SR	-3.43	32.94	-0.11	0.916	7.5

Cuadro 3: Regresión OLS

Paso	1
Constante	18.49
NDVI	60
Valor de t	2.59
Valor de p	0.017

Cuadro 4: Regresión RPP

Predictores	Coefficientes	SE Coef.	Estadístico t	Valor de p
Constante	39.133	4.013	9.75	0.000
CP1	6.569	2.31	2.31	0.030

Cuadro 5: Regresión RCP

### Regresión paso a paso (RPP)

En este método de regresión se requieren dos valores de corte, la F de entrada y la F de salida y hemos preferido definir las como iguales. Se tomó, en todos los casos, una tasa alfa de error tipo I igual a 0.15, para generar F de entrada y F de salida. El Cuadro 4 muestra el único paso llevado a cabo por el procedimiento. La prueba t es significativa con un nivel de significación de 0.05, luego el índice NDVI es capaz de explicar la variable respuesta.

### Regresión con componentes principales (RCP)

En el Cuadro 5 brinda los coeficientes de la regresión con la primera componente principal, la que está explicada, como se mencionó anteriormente, fundamentalmente por los índices de vegetación SAVI y SR. El modelo se ajusta a los datos puesto que la prueba t es significativa con un nivel de significación de 0.05.

En el Cuadro 6 se resume la comparación entre los modelos, según su capacidad predictiva. Puede observarse que el modelo con mayor capacidad predictiva para estos datos, es la regresión paso a paso con el *PREES* menor y el  $R^2_{prediccion}$  más alto. Este modelo seleccionó como mejor predictor el índice NDVI. Como pudo apreciarse antes, este índice predomina en la segunda componente principal, que explica casi la tercera parte de la variabilidad total del fenómeno.

	Regresión OLS	RegresiónRPP	RegresiónRCP
PRESS	10498.3	9285.22	10073.5
$R^2$	0.73 %	12.2 %	4.74 %

Cuadro 6: Comparación de tres modelos de Regresión para predecir COS

## 6. Conclusiones

Se ha llevado a cabo un estudio comparativo entre los métodos de regresión OLS, RPP y RCP en el estudio de predicción de COS en función de índices de vegetación obtenidos a través de teledetección en la zona de Tetelilla, Teziutlán, Puebla, México. La comparación se realizó bajo el criterio de mejor capacidad predictiva. La RPP mostró la estadística  $PREES$  más baja y el  $R^2_{prediccion}$  más alto entre los tres modelos considerados. Este modelo es la regresión lineal simple de COS con el índice de vegetación NDVI.

Investigaciones posteriores serán necesarias para estimar los niveles de COS a gran escala, así como, utilizar otras técnicas que pudieran resultar más adecuadas.

## Referencias

- [1] Birth, G.S. and Mcvey, M.D., (1968) *Measuring the colour of growing turf with a reflectance spectroradiometer*. Agronomy Journal, 60, pp. 640 a 643.
- [2] Chuvieco, E. (ed.) (2007). *Earth Observation and Global Change*. Nueva York: Springer.Geosciences.
- [3] Huete, A.R., (1988). *A soil-adjusted vegetation index (SAVI)*. Remote Sensing of Environment, 25, pp. 53 a 70.
- [4] Linares Fleites, G., (2006). *Análisis de Datos Multivariados*. Editorial Benemérita Universidad Autónoma de Puebla. Facultad de Computación. MÃ©xico. 277p (ISBN 968 9182 15 3).
- [5] Linares, G., Saldaña A., Rivera M. y Reyes, H., 2011. *Comparación de métodos de regresión en la predicción de dióxido de carbono*. Revista Ciencias Matemáticas. Vol. 25. ISSN: 0256-5374
- [6] Martínez J. (2004). *Cambio Climático: una visión desde México* INE-SEMARNAT. MINITAB Release 15 (2005). Statistical Software. Minitab. Inc.
- [7] MINITAB Release 15 (2005). *Statistical Software*. Minitab. Inc.
- [8] Montgomery, D.C., Peck, E.A. y Vining, G.G. (2004). *Introducción al Análisis de Regresión Lineal*. Compañía Editorial Continental. México. 588 pp
- [9] Moslem Ladoni<sup>1</sup>, Seyed Kazem Alavipanah<sup>2</sup>, Hosein Ali Bahrami<sup>1</sup>, and Ali Akbar Noroozi<sup>3</sup>, 2010. *Remote Sensing of Soil Organic Carbon in Semi-Arid Region of Iran*. Arid Land Research and Management, 24, pp.271 a 281
- [10] Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W. and Harlan, J.C., (1974). *Monitoring the vernal advancement of retrogradation of natural vegetation*. NASA/GSFC, Type III Final Report, Greenbelt, MD, USA, 371 pp.
- [11] Sandoval Estrada, M., Stolpe Lau, N., Zagal Venegas, E., Mardones Flores M. y Junod Montano, J., 2003. *El secuestro de carbono en la agricultura y su importancia con el calentamiento global*. Theoria, Vol. 12, pp. 65 a 71.

- [12] Villers L, y Trejo. I., 2004. *Evaluación de la vulnerabilidad en los ecosistemas forestales. En: Cambio climático: una visión desde México*. Compiladores: Martínez, J. y Fernández, A. Instituto Nacional de Ecología. Secretaría del Medio Ambiente y Recursos Naturales. México, pp. 239 a 254.

# Una aplicación del índice Dcova a la Diversidad de escarabajos del bosque mesófilo de montaña

Rosas Nava Mariela Rubi<sup>1</sup>  
Covarrubias Melgar Dante<sup>1</sup>  
Deloya Cuauhtemoc<sup>2,3</sup>  
B.Maired Bernabé<sup>1</sup>

Línea de investigación: Estadística

<sup>1</sup>Unidad Académica de Matemáticas-AUG.

<sup>2</sup>Red Interacciones Multitróficas, Instituto de Ecología, A.C.  
(Centro Público de Investigación, CONACYT), Xalapa, Veracruz.

<sup>3</sup>Cuerpo Académico Biodiversidad y Gestión Ambiental Sustentable,  
U. A. de Ciencias Químico Biológicas.

*Av. Lázaro cárdenas s/n, Col. Haciendita, Ciudad Universitaria  
Chilpancingo, Gro.39060, México  
e-mail:ibur.aleiram@gmail.com*

## Resumen

Se estima que menos del 1 % del territorio nacional está ocupado por vegetación primaria de bosque mesófilo de montaña (BMM) (8,809 km<sup>2</sup>; serie III de INEGI 2005), y actualmente se estima que solo queda un 50 % de la superficie original, la cual ha sido reemplazada por otros tipos de cobertura (Challenger, 1998). Los análisis fueron realizados utilizando la base de los escarabajos copro-necrófagos en el BMM en un transecto altitudinal por Deloya et al. (2007) en el centro de Veracruz, y se compara el comportamiento de los estimadores de los índices de diversidad. Se encuentra que el índice Dcova es mejor para la comparación de dos comunidades diferentes.

**Palabras clave:** diversidad, comunidad, necro-trampas y copro-trampas

## 1. Introducción

Si bien no existe una evaluación reciente que permita determinar con exactitud el área que cubre el bosque mesófilo de montaña (BMM) en México, muy probablemente se trata del ecosistema más amenazado en el país (Challenger, 1998), y el ecosistema tropical que ocupa menos superficie a nivel mundial (Mulligan y Burke, 2005). El BMM se caracteriza principalmente por la presencia frecuente o persistente de nubes a nivel de la vegetación

(Hamilton, 1995). Esta definición basada en el clima refleja la importancia de las nubes o niebla para la ecología de este ecosistema. De ahí que también se le conozca como bosque de niebla, selva nublada, bosque nebuloso y bosque nublado.

Los bosques mesófilos son un grupo de comunidades distribuidas en las montañas, las cuales poseen estructura, afinidad florística y composición de especies diversas. Dado su complejo origen y naturaleza, la clasificación y definición del BMM posee por lo tanto dificultades y ambigüedad. No obstante, podemos decir que el BMM en México se caracteriza por presentar en su resguardo una composición de especies. En una visión hacia las políticas actuales de los países que permiten medir el nivel de importancia, que ha cobrado la ecología en la actualidad y el creciente interés por el tema en el mundo y el apoyo de diferentes organismos internacionales manifiestan una verdadera preocupación por conocer y preservar nuestros recursos naturales.

De acuerdo con Smith, 2002 a lo largo de las últimas cuatro décadas se han registrado avances importantes en los estudios ecológicos y particularmente de estadística en la ecología, ejemplos de ello son las contribuciones al estudio de las distribuciones espaciales de organismos y los diseños de estudios ecológicos. La implementación de políticas que incorporan caracterizaciones del estado que guarda una comunidad ecológica, hacen uso de indicadores propuestos varias décadas antes, los índices de riqueza, equilibrio y diversidad. Patil y Taille (1982) consideran a la diversidad como una propiedad medible intrínseca de la comunidad, definida como el promedio de la rareza de especie. Es importante señalar que si bien, el cálculo de índices de diversidad es relativamente sencillo, aún desde un conocimiento rudimentario, es fundamental al utilizarlos considerar atentamente sus limitaciones para poder interpretar adecuadamente su significado en cada caso particular.

## 2. Objetivo General

Estudiar el comportamiento de los estimadores de los diferentes índices de diversidad. Conocidos como el Simpson, Fager y el Dcova. Comparar los métodos de recolección de escarabajos Necro trampas permanentes (NTP), Necro trampas temporales (NTT) y Copro trampas temporales (CTT).

## 3. Marco Teórico

La diversidad específica es una propiedad emergente de las comunidades biológicas que se relaciona con la variedad dentro de ellas. Este atributo es la expresión de dos componentes, el primero de ellos es el número de especies presentes en la comunidad que es denominado riqueza de especies. El segundo componente es la equitabilidad, y describe cómo se distribuye la abundancia. Para estimar la diversidad se debe: 1. tener un buen conocimiento de la composición taxonómica. Es raro que se estime la diversidad de toda la comunidad, por lo general, se mide la diversidad en un fragmento de la misma que se denomina taxocenosis. 2. considerar que todos los individuos asignados a una clase (especie) son idénticos. Es decir, no se reconoce la variabilidad que puede existir entre, la presencia de una especie. Si es una variable nominal, las categorías son las especies, siendo imposible calcular un promedio, pero si se puede medir la dispersión, la distribu-

ción de las observaciones entre categorías. Numerosos índices han sido propuestos para caracterizar la riqueza de especies y la equitabilidad, denominados índices de riqueza e índices de equitabilidad, respectivamente. Los índices que combinan tanto la riqueza de especies como la equitabilidad en un solo valor se denominan índices de diversidad. Una de las principales críticas a estos índices es que combinan y, por lo tanto, confunden las variables que caracterizan a la estructura de la comunidad: (a) el número de especies (riqueza específica), (b) la abundancia relativa de las especies (equitabilidad), y (c) la homogeneidad y el tamaño del área muestreada.

### Índices de Diversidad.

Los índices de diversidad incorporan en un solo valor a la riqueza específica y a la equitabilidad. En algunos casos el valor del índice de diversidad estimado puede provenir de distintas combinaciones de riqueza específica y equitabilidad. Es decir, que el mismo índice de diversidad puede obtenerse de una comunidad con baja riqueza y alta equitabilidad como de una comunidad con alta riqueza y baja equitabilidad. Esto significa que el valor del índice aislado no permite conocer la importancia relativa de sus componentes (riqueza y equitabilidad).

### Clases de índices de diversidad.

Si bien es cierto que los índices expresados anteriormente son algunos de los más representativos, cada ecólogo puede definir uno, por lo que la literatura apunta la existencia de una gran cantidad. Patil y Taille (1982) proponen clasificar los índices de acuerdo a la medida de la rareza utilizada, entonces tenemos índices dicotómicos y de rango, para el caso dicotómico la rareza de especie definido como.

$$\pi^{\#} = \{\pi_1^{\#}, \pi_2^{\#}, \dots, \pi_k^{\#}\} \quad (1)$$

sin perder en generalidades tomemos

$$\pi_1^{\#} \geq \pi_2^{\#} \geq \dots \geq \pi_k^{\#}. \quad (2)$$

Los datos provienen del llamado vector de abundancias relativas. Sin embargo puede considerarse que el problema teórico se puede analizar tomando en cuenta otros criterios (Bouza y Covarrubias, 2005a) por lo que se proponen dos clases generales para los índices a partir de su estructura: la clase aditiva y la multiplicidad. Esto permite hacer un tratamiento unificado del problema de la estimación. La siguiente definición fija una hipótesis que pretende crear un marco inferencial alternativo para el estudio de los índices de diversidad. Se definen dos clases para los índices de diversidad,

- la clase aditiva.

$$\wedge^+ = \lambda = \sum_{i=1}^k \{\tau_i R(i, \pi_i)\} \pi_i \quad (3)$$

$$\pi_i = \frac{N_i}{N}, N > 0, i = 1, \dots, k \quad (4)$$

- y la clase multiplicativa

$$\Lambda^{\times} = \lambda = \prod_{i=1}^k \{\tau_i R(i, \pi_i)\} \pi_i \quad (5)$$

$$\pi_i = \frac{N_i}{N}, N > 0, i = 1, \dots, k. \quad (6)$$

Donde  $\tau_i$  es el rango que se le asigna a la especie  $i$ .  $R(i, \pi_i)$  la medida de la rareza de especie y  $\pi_i$  la abundancia relativa de la especie  $i$ , son las diferencias básicas propuestas por Patil y Taille (1982) y la propuesta por Bouza y Covarrubias (2005b).

## 4. Aplicación

Se comparan tres técnicas de muestreo para capturar escarabajos copro-necrófagos, con la siguiente clasificación: Técnica I: copro-trampas temporales (CTT) y necro-trampas temporales (NTT). b) Técnica II: necro-trampas permanentes (NTP).

### 4.1. Metodología

Para realizar las comparaciones entre índices de diversidad se utilizó la base de datos del estudio realizado con escarabajos copro-necrófagos en el BMM en un transecto altitudinal por Deloya et al. (2007) en el centro de Veracruz, mediante la captura realizada con trampas en 10 fincas de Veracruz. Para ello se desarrollaron experimentos de Monte Carlos. Se pretendió conocer cuál de los índices puede dar una interpretación biológica más clara al mismo fenómeno.

### 4.2. Modelo

Los índices de diversidad utilizados en este estudio son:  
 Índice de Simpson (1948). En las investigaciones mediante técnicas de diversidad indiscutiblemente el índice de Simpson es uno de los más utilizados, es también uno de los primeros en ser aplicados en la ecología al ser propuesto a finales de los cuarenta. La construcción del índice de Simpson permite una interpretación sencilla de los resultados ya que estos están acotados en el intervalo  $[0,1]$ , esto es de gran ayuda para los investigadores ya que es posible ver la diversidad es alta si el índice se aproxima a 1.  
 Definición 1:

$$\lambda_\rho = \sum_{i=1}^k \{1 - \pi_i\} \pi_i = 1 - \sum_{i=1}^k \pi_i^2. \quad (7)$$

El índice de Shannon (1949). Es una de las medidas de diversidad relacionadas con la teoría de información. Estas medidas parten del supuesto de que una comunidades análoga a un sistema termodinámico en la cual existe un número finito de individuos, los cuales pueden ocupar un número -también finito- de categorías (especies). La estadística para describir esta situación: un sistema con un número finito de individuos y de categorías (especies); sin restricciones en cuanto al número de especies ni de individuos por categoría (especie), esto equivale a la incertidumbre acerca de la identidad de un elemento tomado al azar de una colección de N elementos distribuidos en k categorías, sin importar el número de elementos por categoría ni el número de categorías. Dicha incertidumbre aumenta con el número de categorías (riqueza) y disminuye cuando la mayoría de los elementos pertenecen a una misma categoría.

Definición 2:

$$\lambda_{sh} = \sum_i^k \{-\log \pi_i\} \pi_i = - \sum_i^k \pi_i \log \pi_i. \quad (8)$$

Índice de Fager (1972). Un índice de diversidad que tiene características estadísticas que lo hacen ideal para su estudio es el índice NM (número de movimientos) propuesto por Fager en 1972, la idea es conocer el número de transferencias que se deben hacer para equilibrar de abundancias, el índice de fager es basado en el número de transferencias que se debe de hacer para convertir una distribución de individuos observados en una distribución equilibrada. El índice de Fager es definido como:

Definición 3:

$$\lambda_{fag}^* = \frac{N(k+1)}{2} - \sum_i^k N_i \tau_i = \lambda_0^* - \sum_i^k N_i \tau_i. \quad (9)$$

Donde k es el número de especies  $N_i$  el número de individuos en la especie i en la comunidad, N el número total de individuos.  $\tau_1, \dots, \tau_k$  Son los rangos de la especie en orden decreciente de su importancia, es decir que el rango para la especie de mayor abundancia se le asigna rango 1. Note que se presenta un problema cuando se desea proponer un estimador del índice (1), ya que tenemos que estimar N, por ello Bouza y Schubert (2003) proponen una transformación del índice de Fager que permite estimar las abundancias relativas  $\pi_i$ . El índice transformado estaría dado por el estimador insesgado del índice Dcova, su varianza y el estimador de la varianza. El estimador en el MAScr, suponiendo n fija están dados por:

Definición 4:

$$\hat{\lambda}_{dcova} = \frac{1}{\lambda_0} \sum_i^k \hat{\pi}_i R_i \quad (10)$$

			FINCA									
			EMMCAÑ	ABMM	PASTI	XBMM	CCBS1	TCSol	MBMM	MCBS2	VCSE	ONZ
METOD	NT	INDIV	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
	CT	INDIV	3	21	21	0	65	25	126	143	79	
	NTP	INDIV	2	4	8	1	9	16	26	20	39	
		INDIV	13				74	27		152	208	92

Figura 1: Comparación técnicas de muestreo para capturar escarabajos (CTT, NTT, NTP).

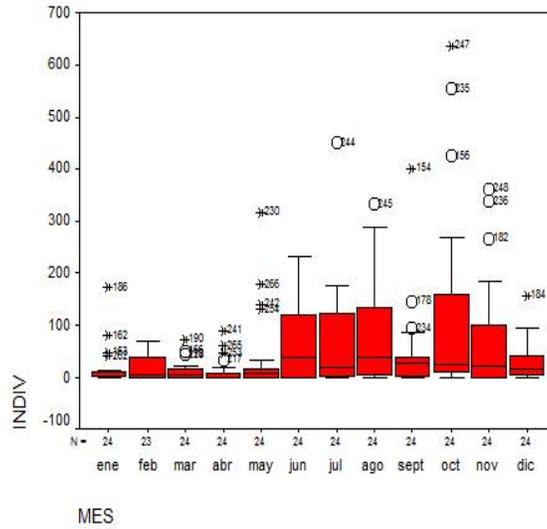


Figura 2: Gráficos de dispersión de los muestreos

$$var\{\lambda_{docova}\} = \frac{1}{n\lambda_0^2} \left[ \sum_{i=1}^k \pi_i R_i^2 - \left( \sum_{i=1}^k \pi_i R_i \right)^2 \right]. \quad (11)$$

## 5. Resultados y discusión

Para realizar el comparativo de los métodos de captura se realizó un análisis de varianza, Figura 1, Figura 2.

Se observa que el método de (NTP) tiene mayores unidades de escarabajos con una media de 94.32, Figura 3.

Se observa que el método de necro-trampas permanentes es significativo, Figura 4.

Observamos que la finca con mayor captura de escarabajos fue la finca 9 (VCSE=

**Report**

INDIV

MES	Mean	N	Std. Deviation
ene	17.83	24	38.13
feb	19.91	23	23.76
mar	13.79	24	19.74
abr	11.25	24	23.52
may	38.00	24	76.95
jun	70.75	24	81.10
jul	70.62	24	101.27
ago	76.63	24	96.34
sept	46.83	24	83.54
oct	117.17	24	180.32
nov	73.42	24	109.38
dic	30.67	24	36.83
Total	49.01	287	89.75

Figura 3: Reporte con las medidas centrales y de dispersión con los individuos

**ANOVA**

INDIV

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	284215.760	2	142107.880	19.986	.000
Within Groups	2019320.2	284	7110.282		
Total	2303536.0	286			

Figura 4: Comparación de técnicas de muestreo por finca

**Multiple Comparisons**

Dependent Variable: INDIV  
Tukey HSD

(I) METOD	(J) METOD	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
NT	CT	40.20*	11.502	.001	13.25	67.16
	NTP	-40.29*	12.853	.005	-70.42	-10.17
CT	NT	-40.20*	11.502	.001	-67.16	-13.25
	NTP	-80.50*	12.829	.000	-110.56	-50.43
NTP	NT	40.29*	12.853	.005	10.17	70.42
	CT	80.50*	12.829	.000	50.43	110.56

\*. The mean difference is significant at the .05 level.

Figura 5: Comparación de técnicas de muestreo por mes (escarabajos copro-necrófagos)

cafetal con sombra especializada) y fueron capturados por el método de necro-trampas permanentes (NTP).

Observamos que los meses donde hubo mayor captura de escarabajos fueron de julio-octubre, Figura 5.

Se observa que el mes donde hay mayor captura de escarabajos es en octubre, Figura 6 y Cuadro 1.

**comunidad    Número de    abundancia    dcova    shanon    simpson**  
**especies**

todas            24                    7388                    0.3193    2.2360    0.8598

Cuadro 1. Datos generales

Se observa un valor alto en la dominancia (Simpson) al interior de unas pocas especies, resultado similar se observa en la diversidad (Shannon), sin embargo el resultado obtenido con Dcova es menor, debido al desequilibrio o inequidad observado entre la abundancia y la riqueza de especies, Figura 7.

Comunidad	Nº.de esp.	Abundancia	Vegetacion	Altitud	Dcova	Shanon	Simpson
Cañadas BMM	6	56	BMM	1400	0.2857	1.1163	0.5818
Cañadas ABM	7	304	Acahual	1400	0.2500	0.4488	0.1795
Cañadas pastizal	12	355	Pastizal	1400	0.1538	1.6131	0.7029
Xalapa BMM	4	14	BMM	1200	0.4000	10.7067	0.5824
Coatepec CBS	11	882	CBS	1200	0.1667	1.5598	0.7491
Teocelo CSOL	8	476	CSS	1200	0.2222	1.3438	0.6962
Mirador BMM	15	1815	BMM	1000	0.1250	1.6738	0.7047
Mirador CBM	16	2078	CBM	1000	0.1176	2.1571	0.8509
Vequia CSE	16	1414	CBSE	1000	0.1176	1.5392	0.6885
TOTAL	.	7388					

Comparando los índices se observa que existen lecturas diferenciadas entre los índices, con relativamente "pocas especies".

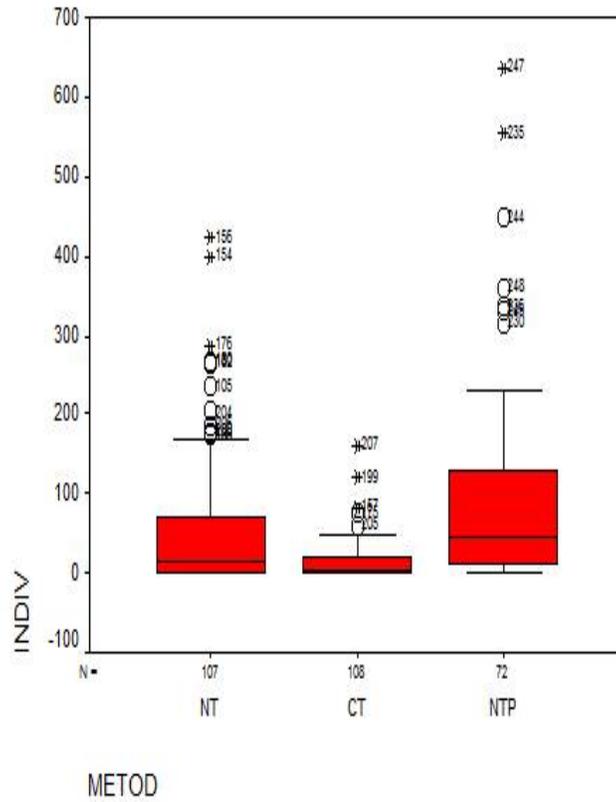


Figura 6: Datos generales

**Report**

INDIV

METOD	Mean	N	Std. Deviation
NT	54.03	107	84.34
CT	13.82	108	24.63
NTP	94.32	72	130.02
Total	49.01	287	89.75

Figura 7: Análisis comparativo respecto a los índices

## 6. Conclusiones

Existe una diferencia significativa entre las técnicas y métodos de captura utilizados, sugiriendo el uso de las necro-trampas permanentes por ser más eficientes en la captura y en esfuerzo de muestreo. Es recomendable el uso del índice Dcova, debido a que permite la comparación entre dos comunidades distintas. El valor bajo obtenido por Dcova sugiere la presencia de especies dominantes en la comunidad, más que una alta diversidad y por lo tanto se estaría observando un desequilibrio en la riqueza específica.

## Referencias

- [1] Bouza C. y Covarrubias D.(2005a), *Estimación de índices de Simpson en m Sitios de Muestreo*, Investigación Operacional 26: N° 2, 187-195.
- [2] Bouza C. y Covarrubias D. (2005b), *Estudio del índice de biodiversidad de Fager*, Presentado en el XV Simposio Internacional de Métodos Matemáticos aplicados a las Ciencias, San José, Costa Rica.
- [3] Bouza C. Y Schubert L. (2003), *The Estimation of Bio-diversity study and the characterization of the dynamics: an application to the study of a pest*, Rev. Matemática e Estadística.2, 85-98.
- [4] Covarrubias D., Bouza C. y Pimentel (2009), *Estudio comparativo de estimadores de índices de diversidad: una corrección al índice de fager*, Presentado en el 2º encuentro de Biometría y la V Reunión de la Región Centro y del Caribe de la Sociedad de Biometría, Boca del Río, Veracruz, 7-10.
- [5] Challenger (1998), *Utilización y conservación de los ecosistemas terrestres de México Pasado, presente y futuro*
- [6] Deloya et al. (2007), *Fauna de coleópteros Scarabaeidae Laparosticti y Trogidae (Coleoptera: Scarabaeoidea) asociados al Bosque Mesófilo de Montaña, cafetales bajo sombra y comunidades derivadas en el Centro de Veracruz, México.*
- [7] Fager E.W (1972), *Diversity: A Sampling Study.*, American Naturalist, 106, 293-310.
- [8] Hamilton L.S. (1995), *A Campaign for Cloud Forests International Union for the Conservation of Nature Gland, Switzerland*
- [9] Magurran, A. (2004), *Measuring Biological Diversity*, Blackwell publishing. U.K.
- [10] Milligan M.R. (1990), *Myakka river basin biological study: Down's Dam to Snook Haven*, Sarasota County Ecological Monitoring Division.
- [11] Mulligan and Burke, S.M. (2005), *FIESTA: Fog Interception for the Enhancement of Streamflow in Tropical Areas.*, Report to DfID.
- [12] Patil G. P. and Taille C. (1982), *Diversity as a Concept and its Measurement* , Journal of the American Statistical Association 77.
- [13] Pielou E. C. (1997), *Mathematical Ecology*, Wiley, New York.

- [14] Shannon C.E. (1949), *A mathematical Theory of Communication*, Then Bell System Technical Journal, 27, 379-423, 623-656.
- [15] Simpson E.H (1948), *Measurement of Diversity* , 163:688.

Matemática Estadística y Medio Ambiente.

Se terminó de imprimir en Junio de 2013  
en la Facultad de Ciencias Físico Matemáticas  
El tiraje consta de 40 ejemplares  
Con domicilio en Av. San Claudio y 18 Sur, Col. San Manuel, Ciudad  
Universitaria, Puebla. Pue., CP 72570, Tel. (222) 2295500  
Fax 22295636 2012