



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y
EN SISTEMAS

“APLICACIÓN DE UN ANÁLISIS DE DISTRIBUCIÓN ESPACIAL PARA LA DETERMINACIÓN DE LA UTILIZACIÓN DE UN MÉTODO DE MUESTREO”

T E S I S
QUE PARA OBTENER EL GRADO DE :
ESPECIALISTA EN ESTADÍSTICA APLICADA

P R E S E N T A:
MARÍA DE JESÚS HERNÁNDEZ LEÓN

ASESOR:
DR. CARLOS DÍAZ ÁVALOS



CIUDAD UNIVERSITARIA

OCTUBRE 2013



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Al Dr. Carlos Díaz Ávalos por su dirección, enseñanza tanto como director así como profesor durante los cursos de la especialización; también por su tiempo e inmensa paciencia, para la realización de éste trabajo.

A la M. en C. Patricia Isabel Romero Mares, por su encauzamiento que facilito el planteamiento y realización del presente trabajo, así como de sus comentarios y recomendaciones sobre el mismo; además de los conocimientos compartidos durante las clases.

A la M. en C. Leticia Eugenia Gracia-Medrano Valdelamar, por su iniciación en las técnicas multivariadas, su revisión, comentarios y sugerencias para la culminación de este trabajo.

Al Act. Miguel Ángel Chong Rodríguez por su revisión y sugerencias sobre el trabajo y su presentación.

A la Dra. María Guadalupe Barajas Guzmán, por su revisión, sugerencias y oportunas recomendaciones sobre la tesina.

A mis demás profesores de los cursos de la especialización en estadística aplicada, Dr. Ignacio Méndez Ramírez , Mat. Luis Gonzalo León Novelo y Act. José Salvador Zamora Muñoz

Al M.en C. Margarito Álvarez Rubio, por su enseñanza y aterrizaje de las metodologías estadísticas dentro del campo biológico.

Al Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas de la Universidad Nacional Autónoma de México, donde estudié la Especialización en Estadística Aplicada.

RESUMEN

Dentro del quehacer biológico de campo existe la problemática de la aplicación de métodos estadísticos en el muestreo, algunas veces por desconocimiento de los procedimientos o por la confusión en el manejo de los conceptos, sin embargo, es posible la aplicación de métodos de muestreo aplicando los lineamientos estadísticos.

El objetivo de éste trabajo fue realizar una aplicación de un análisis de distribución espacial, para demostrar como dichos métodos permiten una visión objetiva del comportamiento de los datos y facilitan la aplicación de la metodología estadística para el muestreo; para ello primero se utilizaron los datos de avistamiento de la nutria de río (*Lontra longicaudis annesctens*) en la cuenca Apatlaco-Tembembe, tomados de trabajos ya publicados. Se delimito un polígono del área de trabajo, el cual se dividió en polígonos de 10 X 10 km a los cuales se les determino la densidad de avistamientos, asimismo se hizo un análisis de la distribución espacial, utilizando la función K de Ripley, el cual mostró que no había una distribución homogénea, sino más bien de tipo agregada; por lo cual se procedió a recalcular la densidad mediante la función Kernel y un análisis con un modelo lineal generalizado. Los resultados de la función kernel mostraron las zonas de mayor densidad, las cuales están asociados a cauces y áreas de selva caducifolia. El modelo lineal generalizado mostró que los tipos de suelo bosque de encino y área sin vegetación son restrictivos, favoreciendo el agrupamiento de nutrias en áreas de selva caducifolia. Estos resultados son relevantes para el planteamiento de estudios posteriores, ya que sabiendo que hay una zonas restrictivas y zonas de mayor prevalecía de la nutria de río, se puede plantear un muestreo por estratos y con ello se puede hacer más eficiente la utilización de recursos, así como tener resultados más confiables respetando las condiciones de un muestreo estadístico.

Palabras clave: Nutria de Río, muestreo, densidad, distribución espacial, función K de Ripley, la función Kernel y modelo lineal generalizado.

ÍNDICE

RESUMEN	i
INTRODUCCIÓN	1
CAPÍTULO 1: MUESTREO EN ESTUDIOS BIOLÓGICOS	2
1.1 Generalidades	2
1.2 El cuadrante	3
1.3 El transecto	5
CAPÍTULO 2: ESTADÍSTICA ESPACIAL	13
2.1 Análisis de patrones puntuales	13
2.2 Vecino más cercano	17
2.3 Autocorrelación espacial	18
2.4 Regresión lineal	18
2.5 Modelo lineal generalizado	20
CAPÍTULO 3: APLICACIÓN DE MÉTODOS DE ANÁLISIS DE DISTRIBUCIÓN ESPACIAL EN UNA POBLACIÓN DE NUTRIAS DE RÍO (<i>Lontra longicaudis annectens</i>) DENTRO DE LA CUENCA APATLACO- TEMBEMBE	22
3.1 Objetivo	22
3.2 Procedimiento	22
CAPÍTULO 4: RESULTADOS Y ANÁLISIS	30
CAPÍTULO 5: DISCUSIÓN Y CONCLUSIONE	40
GLOSARIO	43
ANEXO	44
BIBLIOGRAFIA	46

INTRODUCCION.

El quehacer biológico, tiene diversas problemáticas, sobre todo para la obtención del material de estudio, muchas de éstas recaen en la conceptualización y terminología de la biología y la estadística. El conflicto hace que repetidamente el biólogo, con desconocimiento de los métodos estadísticos, se enfoque en cómo va a obtener cada elemento de la muestra y considere ese su método de muestreo; ejemplo de ello es el escuchar *muestre usando trampas de caída, redes de niebla o, use cuadrante o transecto o el muestreo se realizó entre las 07:00 y las 14:00. En cada área o tramo de playa se utilizó como unidad de muestreo un cuadrante de 1 m², el cual fue lanzado al azar cada 10 pasos sobre la superficie de la playa* (Baena, et al, 2009), bajo la creencia de que, usar estos métodos automáticamente es una forma estadística, sin haber reflexionado en que, la utilización de técnicas para la obtención de los elementos muestrales, requieren de una planeación con rigor estadístico del muestreo (Skellan, 1958; Sen, et al, 1974); el cual está relacionado con procedimientos específicos de análisis. Además el trabajo se complica cuando se utilizan métodos no probabilísticos (Azorín-Poch, 1972, 1986), como son: bola de nieve, de experto, por cuotas, por voluntarios y por disponibilidad, para luego tratar de realizar análisis estadísticos, lo cual nos lleva a errores por la incongruencia del proceso muestrales y el análisis.

Otro problema menos visible, es que, en muchos casos se desconoce la ubicación de los organismos de la población en la zona de interés; cierto es que, ocasionalmente algunos estudios son terreno virgen, sin embargo, incluso al hacer un estudio piloto se puede hacer bajo métodos estadísticos y probar que tipo de distribución presenta la población. Ahora bien, la misma palabra distribución implica un conflicto entre biología y estadística, ya que, biológicamente hablando nos referimos a la distribución como la ubicación de los organismos, no así en estadística. La importancia del conocimiento de la distribución espacial radica en que nos permite conocer tanto el arreglo espacial así como el arreglo estadístico de los datos; además nos ayuda a elegir un procedimiento que facilite la conformación de la muestra, así como nos hace más eficiente el manejo de recursos (humanos, económicos y temporales) y nos brinda un análisis con resultados de alta confiabilidad.

Algunos métodos estadísticos frecuentemente utilizados en el trabajo de campo biológico son, el muestreo aleatorio simple, el muestreo sistemático, el estratificado y el por conglomerados; así como la combinación entre algunos de ellos, permitiendo un análisis más confiable, pero éstas combinaciones dependerán del objetivo fundamental del estudio.

En este trabajo se presentará en el primer capítulo una breve revisión de los métodos de estudio utilizados en el quehacer biológico con fauna silvestre. En el capítulo dos se plantearán las características y relevancia de métodos de análisis dentro de la estadística espacial. En el tercer capítulo se plantea un ejercicio de aplicación en el análisis de estadística espacial en un modelo biológico. En el penúltimo capítulo se presentarán los resultados y el análisis. Finalmente en el capítulo cinco se presenta la discusión y conclusión sobre la aplicación del análisis de estadística espacial desarrollado.

Capítulo 1: MUESTREO EN ESTUDIOS BIOLÓGICOS

1.1 GENERALIDADES

Cuando se realiza un estudio biológico en el campo, existen varios aspectos que se deben considerar para la realización del muestreo, más si se desea que se realice bajo lineamientos estadísticos, los cuales nos permitirán y facilitarán el análisis de los datos, además de brindarnos una confiabilidad en los resultados y la posibilidad de hacer inferencias.

Sin embargo, en muchos de los casos se realizan los trabajos aplicando técnicas o procedimientos de captura de organismos o cuantificaciones o sólo determinaciones merísticas y se considera a estos métodos como los métodos de muestreos, por ello es frecuente escuchar “se utilizó una trampa de arena de 1m^2 para registro de huellas” o “se usaron estaciones fijas” cuando en realidad la disposición de estas trampas es lo que nos habla del tipo de muestreo y no la técnica en particular para registro del elemento. Por tal motivo es importante tener claro que al hablar de una técnica nos estamos refiriendo a la manera en que se utilizará el equipo que nos permitirá la obtención del registro del elemento que conformará la muestra; y que al hablar de un método de muestreo nos referimos a los procedimientos estadísticos que nos permiten la disposición del equipo que nos seleccionará a los elementos que conformarán la muestra. Los métodos de muestreo más comúnmente utilizadas en estudios biológicos son: muestreo aleatorio simple, muestreo sistemático, muestreo estratificado y muestreo por conglomerados; así como las posibles combinaciones que se puedan hacer con ellos, cuando se va haciendo a diferentes niveles (Zar, 1999). La elección depende del contexto del problema de estudio, sin que se pueda afirmar que exista una forma de muestreo óptima. La elección del método a utilizar dependerá del problema o hipótesis planteadas, de los supuestos involucrados, la experiencia adquirida y la facilidad o familiaridad de uso de procedimientos estadísticos empleados.

La problemática del muestreo se enfoca en la elección de la técnica adecuada para coleccionar los elementos que conformarán la muestra, además de la determinación del tamaño adecuado para ésta.

El muestreo estadístico se caracteriza por: a) tener una probabilidad asociada a cada uno de los posibles elementos que pueden conformar una muestra, b) permitir el cálculo del error asociado al mecanismo de muestreo al estimar cantidades de interés, c) permitir hacer procesamientos con los resultados observados que brinden una descripción apegada a la realidad, y d) permitir hacer inferencias de los parámetros en muchos de los casos (Lorh, 2000).

Sin embargo, no sólo la interpretación de conceptos por los campos estadísticos y biológicos ponen en conflicto la aplicación correcta de los métodos y técnicas, sino que, dentro del propio campo de la biología, dependiendo del autor, de cuándo y quién haya hecho la traducción de los procedimientos para la obtención de las muestras, hay una interpretación diferente de los procedimientos y conceptos a seguir; por ejemplo al hablar del cuadrante se tienen dos interpretaciones una haciendo referencia al plano cartesiano, el cuál sólo consta de cuatro cuadrantes y no se considera una unidad con área, por otro lado, se habla de cuadrante (Claphan, 1932; Seber, 1982; Krebs, 1989) como una unidad de área que está definida por una forma geométrica; particularmente en este trabajo, hablaremos del cuadrante refiriéndonos a ésta última concepción. Otro concepto que se encuentra en esta misma situación es el transecto, que algunos lo mencionan como transecta y la utilizan como un área delimitada y definida, que además se divide de acuerdo al gradiente ambiental que presenta, y en cada parte de la división se toman algunos elementos para conformar la muestra. La otra forma de ver el transecto es una

línea (conocida por algunos como línea de Canfield (Bautista-Zúñiga, et al, 2004), de la cual se toman puntos al azar para hacer los registros, ya sean perpendiculares a ella o angulares y que a ambos lados de la línea se considera un margen (ancho de banda). En éste trabajo se considerará el concepto de transecto y no de transecta. Tanto el cuadrante como el transecto se describirán en apartados más adelante.

La elección de la técnica para la obtención de los elementos de la muestra dependerán de la biología del organismo así como del objetivo. En el caso de que el organismo no se desplaza o lo que se va a utilizar como registro de él es su madriguera o si las áreas para registro superan la capacidad de desplazamiento de los organismos o algún otro tipo de elemento que sea fijo (sésil) se usará el cuadrante; o si se desplaza continuamente (vágil) y la velocidad de desplazamiento o la distancia recorrida es amplia referente al organismo, ejemplo los cetáceos (Ballenas, delfines, marsopas y zifios) se usará el transecto.

Otro aspecto a considerar de la biología de la especie, es el conocimiento de su distribución; aquí cabe mencionar que, en biología, cuándo se habla de distribución se hace referencia a la ubicación de los organismos y no al concepto de distribución estadístico; sin embargo, para distinguir los conceptos en éste trabajo cuando hablemos de la ubicación hablaremos de la distribución espacial y cuando se mencione distribución será sobre la conceptualización estadística.

La interpretación biológica de la distribución que presentan los organismos se ve representada, básicamente por tres tipos: azarosa, agregada o uniforme; cada una de ellas nos habla del medio o bien de la conducta de los organismos. Así la azarosa nos habla de un medio altamente homogéneo o bien de organismos generalistas; la distribución agregada nos indica un medio heterogéneo o de la disposición de los recursos en parches (también agregado) y finalmente una distribución homogénea nos habla de una competencia intraespecífica o bien de una territorialidad (Begon, et al, 1997). La relevancia del conocimiento de la distribución nos puede ayudar a la elección de las zonas de muestreo y del método a aplicar en las zonas, para que lograr un mejor análisis, acorde con los objetivos del trabajo.

1.2 EL CUADRANTE

El cuadrante es frecuentemente usado en organismos sésiles o en aquellos que son prácticamente estacionarios o tienen guaridas, panales o madrigueras fijas, como es el caso de los pequeños mamíferos. Cabe mencionar que en este caso, el tamaño y límite del cuadrante puede tener un efecto en el muestreo conocido como efecto de borde (Seber, 1982), el cual está relacionado con el tamaño y forma del cuadrante, así como con el tamaño y forma de la unidad de estudio (u.e); ya que si no están en la misma dimensión y el cuadrante es menor que el tamaño de la u.e, se puede dejar fuera del área de registro elementos que en realidad forman parte de la u.e. (cuando hablamos de agregados), o bien hacer un registro de ausencia por no tener la capacidad de contenerlo (en casos de elementos individuales); o bien si el cuadrante es de mayor talla que la u.e., en el cuadrante se pueden incluir más elementos de los que realmente deben incluirse, y llevar a conclusiones erróneas, como de que se tratará de agregados, cuando en realidad es sólo que no es la correcta dimensión (Clapham, 1932; Seber, 1982; Krebs, 1989; Dixon, 2002).

Se ha visto que el efecto de borde en el cuadrante es mínimo en la forma circular, luego en el cuadrangular y por último en el rectangular (Krebs, 1989), aquí lo que implica es que: al ser mayor el área perimetral, respecto al área total, mayor puede ser el error y llevar a sobre o subestimación del contenido del cuadrante; sin embargo, en ocasiones el uso del cuadrante tiene

menor error si es de forma rectangular, esto cuando el área de estudio presenta una alta heterogeneidad, ya que, el rectángulo por ser alargado puede atravesar mayor número de áreas (Clapham, 1932; Seber, 1982; Krebs, 1989). Por otro lado, con el rectángulo existe una mayor probabilidad de error por unidad de área para determinar si los individuos se encuentran dentro o fuera del borde del área, debido a las características de relación del perímetro (Seber, 1982).

El efecto de borde aumenta cuando el tamaño del cuadrante disminuye, así que se busca una mayor área considerando el menor efecto de borde (Seber, 1982). El efecto del cuadrante tiende a reducirse, cuando el cuadrante es menor que el tamaño del agrupamiento, cuando este no tiene una distribución uniforme (Seber, 1982; Diggle, 1983). En otras palabras, hay menor efecto del cuadrante cuando éste está cercano al tamaño del organismo o a la agrupación, y es mayor el efecto si excede la talla del organismo o a la agrupación. En general el tamaño del cuadrante es el adecuado cuando no interfiere en el establecimiento de la conformación de la u.e.; y puede evidenciarse a los agrupamientos o el número de organismos muestreados por área, es decir cuando la probabilidad no sub o sobrestima (Clapham, 1932).

Si se tiene alguna referencia de la densidad posible en un área dada y se considera que las densidad por área es constante es recomendable que todos los cuadrantes tengan el mismo tamaño (Seber, 1982).

De forma general, para el uso del cuadrante se pueden considerar los siguientes pasos para la planeación y elección del tipo de cuadrante.

1.- Forma del cuadrante: considerando tamaño del animal y movilidad, sin perder de vista la heterogeneidad del medio. La forma geométrica puede ser de cualquier tipo para dos dimensiones, circular, cuadrada, hexágona, etc.

2.- Tamaño del cuadrante:

2.1 Saber como es el arreglo de los organismos; gregario o solitario.

2.1.1 Si se agrupan: ¿Cómo?, ¿De que tamaño es el agrupamiento?. Entonces al agrupamiento en el momento de calcular el número de cuadrantes se le considera como el elemento a muestrear (si es que queremos conocer la densidad)

2.1.2 Si no se agrupan: Uniformemente o con una probabilidad desigual variable

3.- Selección del número de cuadrantes para muestrear; puede determinarse conociendo la probabilidad de encontrar un organismo por el área total a muestrear entre el tamaño del área del cuadrante.

$$m = \frac{pA}{a} , \dots\dots\dots (1.1)$$

donde m = número de cuadrantes de la muestra, p = probabilidad de encontrar a un organismo, A = área total a muestrear, a = área del cuadrante (Seber, 1982).

1.3 EL TRANSECTO

El transecto se utiliza cuando la población que se pretende muestrear es altamente dispersa (tiene un arreglo espacial amplio), cuando el terreno es de difícil acceso o el organismo tiene una alta movilidad y/o el área de muestreo tiene forma irregular (Smith, 1979; Seber, 1982).

Teoría general

La técnica del transecto consiste en tender una línea o líneas de longitud L , que no se cruzan ni se traslapan, sobre la línea del transecto se marca un ángulo recto y se mide la distancia al objeto visualizado y sobre la línea, Figura 1.1, la línea está dividida en secciones y en diferentes puntos sobre la línea se hacen los registros de los elementos que conformarán la muestra; se considera que la línea del transecto atraviesa a la población. Los registros se hacen en cualquiera de los lados de la línea, el tipo de datos que se pueden registrar a lo largo de la línea son número de individuos vistos (n), el área que ocupan, con que otros organismos se asocian, la posición del registro respecto a la línea del transecto, cuando se tiene el avistamiento inicial se registra uno o más de las referencias estadísticas de la Figura 1.1 (Sen, et al, 1974; Buckland, 1985).

El transecto se ha aplicado a diversos hábitats para la determinación de la densidad, aún considerando que los organismos se desplazan a lo largo del transecto y que pueden alejarse o acercarse a la línea del transecto, en la utilización del transecto básicamente lo que se hace es la medición de la distancia de un animal a una línea de referencia ya sea perpendicularmente o radialmente. Skellam (1958) realizó un trabajo en el que se plantean los fundamentos matemáticos del uso del transecto, en el cual considera el cálculo de la densidad en función del número de avistamientos, la velocidad del movimiento del organismo respecto al observador, la distancia a la que es avistado el organismo respecto a la línea del transecto, las velocidades a las que se desplazan tanto el organismo como el observador, así como la adaptación que se debe hacer cuando la zona de muestreo es heterogénea; ecuaciones 1.2, 1.3 y 1.4

$$D = \frac{z}{2RV}, \dots\dots\dots (1.2)$$

donde: D = densidad de la población estudiada

z = número de elementos contados por unidad de tiempo

V = Velocidad promedio del organismo relativa al movimiento del observador

R = Distancia lineal o radial en el que un organismo se aproximará al encuentro con el observador

$$V^2 = \bar{u}^2 + \bar{w}^2, \dots\dots\dots(1.3)$$

donde : \bar{u} = Velocidad promedio del organismo

\bar{w} = Velocidad promedio del observador

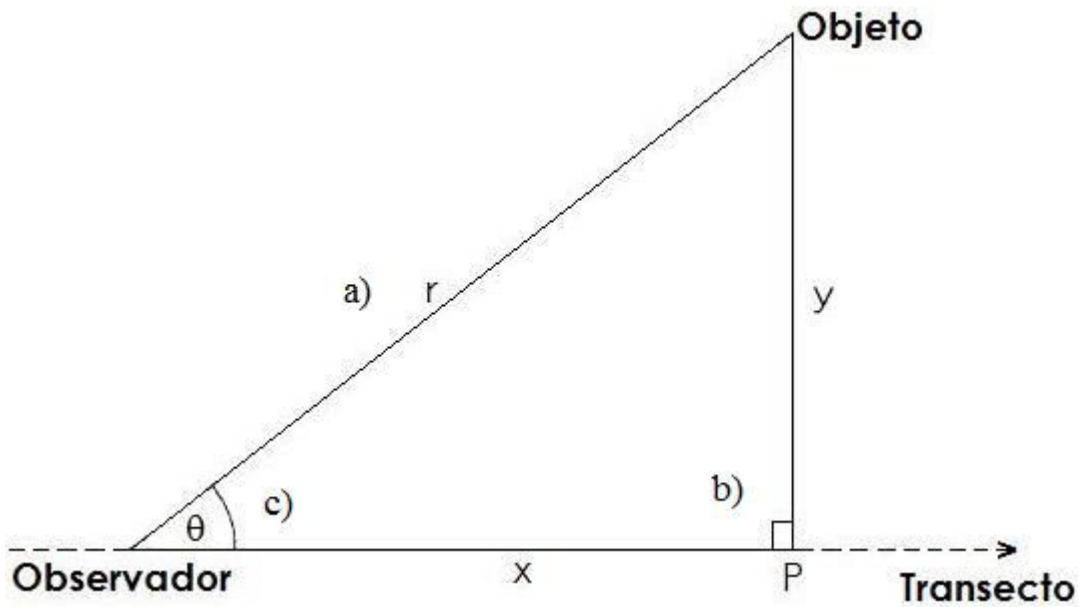


Figura 1.1. Esquema de medición del avistamiento de un objeto de forma perpendicular, (P – Objeto) cuando el observador está a una distancia y , así como cuando está a una distancia y radial al transecto ($\overline{P\theta O}$). a) La distancia radial hacia el animal observado r_i ($i= 1, 2, \dots, n$), b) El ángulo recto a la distancia y_i del animal visto hacia el camino del observador, c) El ángulo de visualización θ_i del observador hacia la posición del animal donde fue visto inicialmente. (Buckland, 1985).

Ahora bien si la zona donde se realiza el muestreo es heterogénea se tiene que hacer un ajuste en el que se implica las densidades de los diferentes segmentos que constituyen la zona total de muestreo

$$\varepsilon_o \{n\} = DHTV, \dots\dots\dots(1.4)$$

donde $T = \sum Ti$; tiempo que el observador invierte en el transecto.

H = Diámetro medio de la región del observador a la detección.

Trabajos posteriores respecto a la utilización del transecto llevaron al planteamiento de supuestos (Smith, 1979; Seber, 1982) que facilitan el cálculo de la densidad y en general los muestreos con el uso del transecto, sin que se vea afectado el resultado por el desplazamiento del organismo sobre el transecto. Los supuestos son:

- 1.- Los animales se distribuyen al azar de manera independiente en el área de la población
- 2.- El avistamiento del animal es independiente del avistamiento de otro
- 3.- Ningún animal es contado más de una vez
- 4.- El animal está inmóvil antes de ser avistado
- 5.- El comportamiento de respuesta de la población en su conjunto no se afecta a lo largo del transecto

6.- Los individuos son homogéneos con respecto a su conducta de respuesta de acuerdo a sexo, edad, etc.

7.- La probabilidad de que un animal sea visto, es una función de la distancia al transecto. La probabilidad de detectar un animal que esta directamente sobre el transecto es 1, ejemplo $p(x=0)=1$ donde x es la distancia al transecto.

8.- No hay error de medición

Ahora bien otras de las consideraciones que hace Smith (1979) retomando el trabajo de Skellam (1958) respecto al movimiento de los animales objeto son que:

- i) Los objetos están distribuidos azarosamente sobre la región geográfica de interés antes de empezar o ser afectados por el observador
- ii) Los objetos están inmóviles antes de ser sorprendidos, pero pueden moverse a una nueva locación donde serán detectados o dejan de moverse
- iii) La probabilidad de detectar a un objeto es una función de la distancia al transecto

Estas consideraciones no se contraponen con los supuestos que se mencionaron anteriormente sobre el uso del transecto, la principal diferencia entre cuando se mueven y cuando no, es ii; esto es valido para anidaciones o pequeños mamíferos o animales con guaridas o madrigueras. El que los animales no se hayan detectado antes puede responder a que se encuentran en las madrigueras o se esconden de depredadores o el observador está ocupado con otro animal.

Pero si consideramos la velocidad de desplazamiento del organismo y del observador se puede utilizar la fórmula (1.5).

$$\hat{D} = \frac{n}{(HTV)}, \dots\dots\dots (1.5)$$

donde n = número de animales observados

H = el diámetro medio de la región del observador a la detección

V = la velocidad promedio a la que se mueve el objeto

T = tiempo que el observador invierte en el transecto.

O bien si se considera inmóvil el objeto se puede estimar

$$\hat{D} = \frac{n}{(2wL)}, \dots\dots\dots (1.6)$$

donde n = número de animales observados

w = la distancia máxima a la que es visible el objeto para el registro del avistamiento

L = largo del transecto

Si se supone que la distancia se mide perpendicular siempre al transecto podemos decir que, para la determinación de la densidad se puede utilizar la diferencia de las n_s respecto a $2L$, es decir:

$$D_2 - D_1 = \frac{E(n_2) - E(n_1)}{2L}, \dots\dots\dots (1.7)$$

donde la diferencia entre $n_2 - n_1$ sólo se puede utilizar para determinar el cambio independiente significativo de la visibilidad y el movimiento.

La distancia al transecto puede ser una causa de un mal registro de elementos en él, esto debido a que ocasionalmente en una distancia corta al transecto o incluso si esta sobre la línea del transecto no se llegan a registrar bien los objetos debido a que el observador usualmente busca a una distancia del transecto y no sobre él. Por otro lado los objetos no serán detectados si están más allá de la distancia de alcance del transecto o mejor dicho de la distancia focal del observador. Esto se soluciona si se establece una distancia máxima y otra mínima para la detección del objeto respecto al transecto (Smith, 1979). Otro factor de conflicto en los registros es el de contar múltiples veces a un mismo elemento, esto se vera reducido considerablemente, ya que en una línea de transecto el objeto será poco probable volverlo a contar, sí se restringe arbitrariamente las regiones donde se colectarán los datos desde un principio, sí el transecto se divide en varios segmentos que no se cruzan y que no están muy juntos (Smith, 1979).

La utilización del transecto tiene diversas modalidades; las modalidades pueden ir desde la forma más simple que es un transecto lineal, hasta transectos en radiales (Buckland, et al, 1993). El caso del transecto lineal, es la base para la construcción o elaboración de cualquier otra disposición de transectos, incluso de la determinación con distancias radiales. Una consideración importante para el empleo de transecto, es que hacia ambos lados de él hay un límite de detección de los objetos en cuestión, así que al hablar de los avistamientos se debe considerar que éste puede ser de cualquier lado de la línea. Sí se considera simétrica el área a ambos lados de la línea del transecto, se puede estimar la desviación estándar y la varianza como si la distribución a la línea fuera normal (bajo modelos polinominales), (Buckland, 1985). Para estimar la densidad en el transecto no se tiene que asumir distribución uniforme al ángulo recto (Hiby, et al, 2001).

En el uso del transecto es importante el detectar agrupamientos para evitar la sobreestimación de la densidad cuando se hacen registros de animales (Diggle, 1983; Buckland, 1985). Se recomienda agrupar los datos cuando se considera que puede haber errores en las mediciones, pero que estas son pequeñas o bien si hay registros de ceros.

El modelo radial se usa generalmente para animales móviles (Hayes, et al, 1983), y consiste en registrar el ángulo (θ) desde el transecto al objeto avistado y trazar el ángulo recto desde el objeto al avistamiento, para este tipo de mediciones es usual el auxiliarse de los teodolitos u otros aparatos telemétricos (Buckland, 1985). Si el animal se mueve y no se desea o no se puede calcular la velocidad de desplazamiento la opción es, tomar el radio a la línea central y el observador se mueve al centro (Sen, et al, 1974). Si el radio es grande y el ángulo tiende a cero el error de redondeo puede causar una distancia perpendicular relativamente más grande que se redondea a cero, el agrupamiento de observaciones reduce ésta dificultad. (Buckland, 1985)

- La utilización del modelo radial parte de los siguientes supuestos (Hayes, et al, 1983):
- a.- Los animales están distribuidos azarosamente e independientemente en todas las áreas de estudio
 - b.- Los avistamientos son eventos independientes

- c.- Los animales se mantienen en un lugar, al menos hasta antes de ser detectados, hasta antes de que puedan huir
- d.- no se cuenta ningún animal más de una vez
- e.- la distancia es medida sin error
- f.- las condiciones de avistamiento se mantienen estables a lo largo de todo el estudio
- g.- el avistamiento de los animales en la línea del transecto se pueden ver con una probabilidad de uno.

Se puede apreciar que los supuestos, son los mismos que se tiene para la utilización y detección de objetos cuando se utiliza el modelo perpendicular que el radial. Sin embargo, el modelo de ángulo recto requiere supuestos menos estrictos que el modelo radial, debido a que se considera que el animal es detectado si está en el plano de la circunferencia de detección, ya que, si el organismo se encuentra por arriba o por abajo del plano de detección, se puede modificar la inclinación y por lo tanto, la determinación del ángulo, porque además habría que hacer la corrección al plano, por lo cual se considera mejor estimador al del ángulo recto (Hayes, et al, 1983).

Otra de las aplicaciones del transecto es cuando se utilizan las veredas como transectos, cabe mencionar que en la mayoría de los casos estas veredas son sinuosas. El usar caminos puede tener el inconveniente de que se sesga el muestreo, ya sea para sobreestimar o subestimar, debido a que, el supuesto de una distribución uniforme de la distancia al ángulo recto no se mantiene para transectos curvos, ya que, se requiere que la distancia a la línea (w) sea menor al radio mínimo de la curvatura, Figura 1.2, esto debido a que, en general, un objeto puede estar en ángulo recto con más de un punto en un transecto curvo, por lo cual la distancia de un ángulo recto se debe especificar como la distancia al más cercana de los puntos si es que se define previamente los puntos donde se va hacer el registro de los datos o bien limitar la distancia a la línea central si se van a registrar toda ocurrencia para todos los objetos visibles (Hiby, et al, 2001).

El registro de avistamientos en un transecto curvo tiene que medir la distancia mínima al transecto de donde se ubica el objetivo cuando es detectado, cuando se detecta un animal móvil se debe de mantener en la mira hasta acercarse lo más posible a la distancia más corta al transecto del punto en que fue visto inicialmente. La detección de los objetivos es imposibles cuando están más allá de la distancia w de la ruta Figura 1.2. (Hiby, et al, 2001).

El uso de transectos curvos o veredas puede presentar más problemas que facilidades. En caso de utilizarse, no se deben de usar transectos cortos, ya que hacen que la probabilidad de utilización por parte de la especie objeto sea muy baja. La red del camino puede fallar en la representatividad de la muestra de los hábitats disponibles, dado que existen veredas que son más difíciles de transitar o que sufren cambios en la cercanía del hábitat, lo cual lleva a fraccionar el área y es mejor hacer el muestreo considerando los diferentes estratos para así evitar sesgos; así que, el uso de veredas debe hacerse bajo suposiciones razonables que dependerán de la especie y el uso del camino tanto por el organismo como por los lugareños (humanos y no humanos, depredadores, por ejemplo) que se encuentran en la zona. Un aspecto favorable en el uso de veredas o transectos curvos naturales es posible para evitar cualquier riesgo con la correlación del objetivo, por ejemplo seguir los cursos de agua (Hiby, et al, 2001).

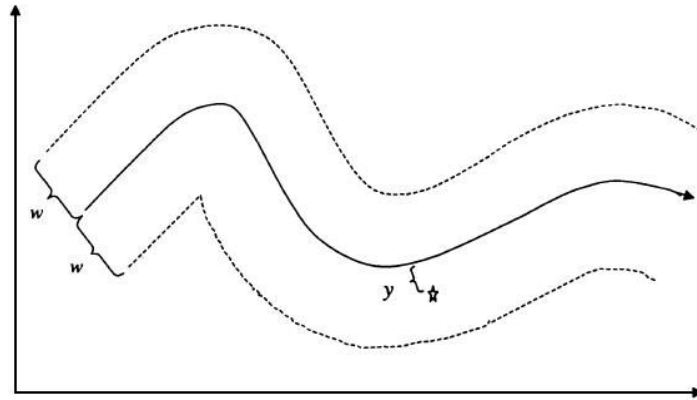


Figura 1.2. Esquema de la disposición de un transecto curvo en el cual se tiene para ambos lados del transecto la distancia máxima de avistamiento desde el transecto (w), y donde se observa la distancia mínima del transecto donde no se tiene efecto por el ángulo de curvatura del transecto. (Hiby, et al, 2001).

De cualquier forma en el uso de cualquier transecto sin una genuina aleatoriedad en el transecto, hace imposible evitar el riesgo de un sesgo, y esto debe de ser considerado. Existe el riesgo de que algunos animales se vean cruzar el transecto muy por delante del observador y por lo tanto se le asignará una muy pequeña o incluso nula distancia perpendicular. Esto amplifica el efecto del movimiento al azar del objetivo, que se considera normalmente insignificante, y puede causar graves sesgo al azar. (Hayes, et al, 1983; Buckland, 1985; Hiby, et al, 2001)

Tipos de transecto

Existen diversas formas de tender un transecto tales como: transecto en banda Figura.1.3, en línea Figura 1.4, de conteo por puntos Figura 1.5, puntos al transecto Figura 1.6, por pistas, por marcas, en red Figura 1.7 (Krebs, 1989; Zar, 2000; Buckland, et al, 1993).

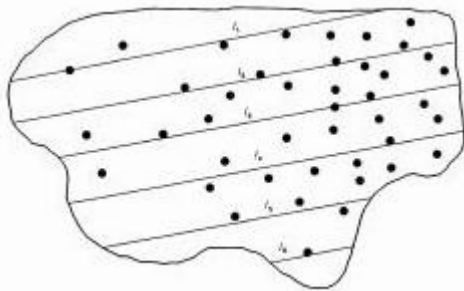


Figura 1.3. Transecto en banda

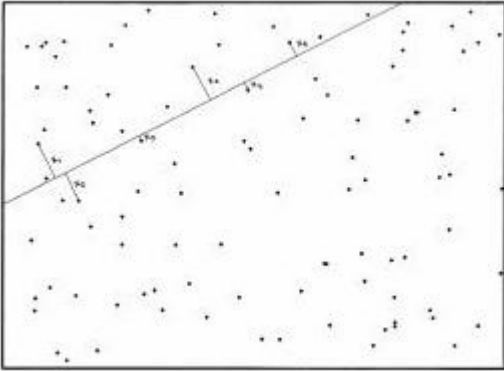


Figura 1.4. Transecto en línea.

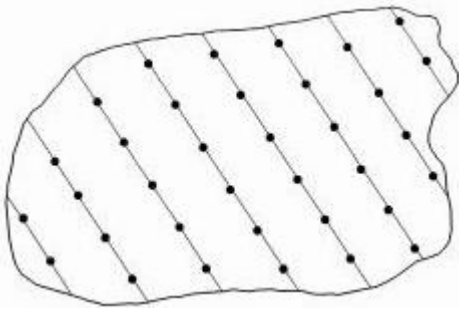


Figura 1.5. De conteo por puntos

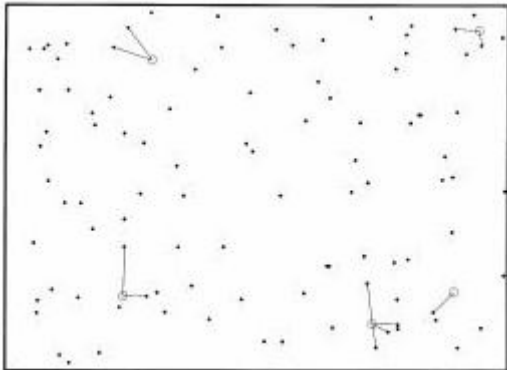


Figura 1.6. Puntos al transecto

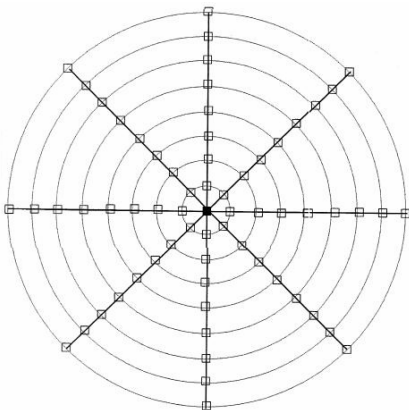


Figura. 1.7 Transecto en red

La utilización del transecto debe hacerse cuidando que se realicen las premisas estadísticas, ya que en algunas ocasiones el investigador toma el nombre de un tipo de muestreo de transecto, pero no cubre las premisas y hace la colecta con todo lo que se encuentre a lo largo del transecto. De hecho, la elección de que tipo de transecto se va a utilizar es sólo una parte del diseño que se debe hacer para realizar el muestreo.

Es recomendable que si se van a hacer registros por más de una persona, es conveniente realizar un muestreo piloto, con el que se pueda estandarizar el alcance visual de los participantes (Krebs, 1989; Buckland, et al, 1993).

Dentro del diseño también se debe considerar si el área de trabajo es terrestre o acuático y en cada uno de los casos las posibilidades de condiciones climáticas, estacionales y físicas. En el caso de terrestre además la latitud, relieve, tipo de substrato y la accesibilidad del terreno para el muestreo. En el caso del acuático, hay que considerar, la profundidad, visibilidad y corrientes, si es marino, ya sea si es costero u oceánico. Si es dulceacuícola, si es ribereño, lagunar, estuarino, presa o de temporal. Estas características nos permitirán definir si se tiene que hacer por estratos, conglomerados, etapas o incluso ir un poco más lejos por bloques.

Capítulo 2: ESTADÍSTICA ESPACIAL

La estadística espacial es una metodología que nos permite el análisis de variables que presentan interacción espacial y temporal, en el análisis de este tipo de variables, el índice, representado por las coordenadas donde se hizo la medición, es una pieza clave de información que nos facilita la visualización de comportamientos asociados de la variable de interés. Las variables con dependencia espacial muestran que dicha dependencia decrece con la separación entre las observaciones, es decir, cuando hay dependencia espacial, los sitios cercanos tienen valores más similares que los distantes. Por ejemplo en el registro de la presencia o ausencia de un organismo no sólo se debe de considerar las frecuencias de observación, sino las características del entorno donde se hace el avistamiento.

Los métodos cuantitativos de análisis espacial tienen como objeto la detección y descripción de patrones espaciales, permitiendo evaluar hipótesis sobre los procesos ecológicos que han causado el patrón observado (Murguía, 2005).

Desde un punto de vista más formal, la realización de un estudio en estadística espacial, implica la reunión de una serie de metodologías apropiadas, para el análisis de datos que corresponden a la medición de variables aleatorias en diversos sitios (puntos del espacio o agregaciones espaciales) de una región. $\{Z(s): s \in D \subset R^P\}$; en el que s es la ubicación en el espacio Euclidiano P -dimensional y $Z(s)$ es una variable aleatoria en la ubicación s . Donde las variables aleatorias $Z(s)$ pueden ser continuas y $D \subset R^P$ es un conjunto continuo (Giraldo-Henao, 2011). El estudio espacial se puede clasificar por tres tipos:

- I.- Geoestadístico: estudia procesos estocásticos $D \subset R^P$ en los que el espacio es continuo, la selección de los sitios depende del investigador.
- II.- Áreas o Regionales: los patrones del proceso estocástico $D \subset R^P$ son de tipo discreto, la selección de los sitios depende del investigador.
- III.- Patrones puntuales: el conjunto de datos $D \subset R^P$ es aleatorio. La decisión de dónde se hace la medición de los elementos de la muestra no depende del investigador; el conjunto puede ser discreto o continuo, pero la ubicación de los sitios dependen del fenómeno en estudio, es decir, está dado por la naturaleza fenomenológica. Este caso particular de análisis tiene como objetivo el determinar la distribución dentro de la región; es decir, si es aleatoria, agregada o uniforme.

2.1 ANÁLISIS DE PATRONES PUNTUALES

Un supuesto común es que si un patrón espacial es completamente aleatorio, entonces estamos hablando de un proceso Poisson homogéneo; esto implica que cualquier región del plano tiene la misma probabilidad de contener un evento o punto, sin embargo, esto no siempre sucede, ya que tenemos patrones aleatorios no homogéneos, los cuales los podemos clasificar de acuerdo a su función de intensidad, la cual representa el número de eventos que ocurren por unidad de área, para áreas infinitamente pequeñas. Así, tenemos que un proceso puntual se llama completamente aleatorio si: (Giraldo-Henao, 2011).

- 1.- El promedio de eventos por unidad de área, la intensidad $\lambda(s)$ es homogénea a lo largo de D
- 2.- El número de eventos $n(A1)$ y $n(A2)$, en dos subregiones $A1$ y $A2$ que no se traslapan, son independientes
- 3.- El número de eventos en cualquier sub-región A , sigue una distribución de Poisson, con intensidad $\lambda(A)$

Ahora bien cuando la función de intensidad no es constante se habla de procesos puntuales no homogéneos los cuales pueden o no presentar algún tipo de interacción entre eventos.

La intensidad se puede estimar por diversos métodos:
Si el patrón es completamente aleatorio

$$\lambda(s) = \lambda = \frac{n}{|A|}, \dots\dots\dots(2.1)$$

donde n = número de eventos y A = área de donde se registran los eventos.

La función K de Ripley es una herramienta para el análisis de los datos de proceso puntuales en el espacio totalmente aleatorios; los cuales comprenden un área predefinida. La función se puede utilizar para resumir un patrón de puntos; para realizar una prueba de hipótesis sobre el patrón de puntos, o para estimar los parámetros y el ajuste del modelo (Dixon, 2002); permite analizar patrones tanto a pequeña escala como a gran escala espacial, incluyendo las distancias entre agregados, ya que se fundamenta en el análisis de las distancias entre cada punto y todos los demás puntos del patrón. (Rozas y Camarero, 2005).

La k de Ripley define el grado en que un conjunto de puntos se separan el uno del otro, para ello se utiliza la distancia promedio entre los vecinos más cercanos.

$$K(t) = \lambda^{-1} E[N(b(0,t))] \dots\dots(2.2)$$

donde λ es la densidad o intensidad de los eventos (número de eventos por área) y E es el número de eventos adicionales en distancia t de un caso elegido al azar. $K(t)$ describe las características de los procesos puntuales en muchas escalas de distancia.

La función K se define en términos de un evento elegido al azar o bien cuando se conoce el grado de heterogeneidad en la distancia t entonces se puede ajustar.

La función $K(t)$ puede ser analíticamente evaluada; y se puede escribir de una manera reducida; la más simple, y más comúnmente usada es la función $K(t)$ para procesos Poisson homogéneos conocido también como procesos espaciales completamente al azar (CRS, por sus siglas en inglés), (Dixon, 2002).

$$K(t) = \pi t^2, \dots\dots\dots(2.3)$$

Ahora bien la estimación de la función K de Ripley debe de considerar el efecto de borde, lo que representa que, como el área de estudio se define de manera arbitraria, existen elementos que pertenecen al conjunto de datos y que no son incluidos en el estudio. Por ello se recurre a una corrección del efecto de borde, ecuación 2.4

$$K(d) = n^{-2} A \sum_{i=1}^n \sum_{j=1}^n w_{ij}^{-1} I_{ij}(d), \dots\dots\dots(2.4)$$

para $i \neq j$, donde: n es el número de puntos en la muestra, A es el área de la región de interés, w_{ij} es el patrón de corrección de borde o función de peso, es decir, la proporción de la circunferencia de dicho círculo que cae en el área de estudio. En otras palabras, es la función de peso, que proporciona la corrección de borde; tiene el valor de 1 cuando el círculo está centrado en i y pasa por el punto j que está completamente dentro de la zona de estudio. I es el indicador de la función, que asume el valor de 1 cuando es cierta la distancia y 0 en cualquier otro resultado, d es la distancia entre los puntos i y j , es decir:

$$I_{ij}(d) = \begin{cases} 1 & \text{si } d_{ij} \leq d \\ 0 & \text{si } d_{ij} > d \end{cases} \dots\dots (2.5)$$

La función $K(d)$ es sensible a la falta de homogeneidad proporcionando funciones de distribución sesgadas cuando estos requisitos no se cumplen, aunque es perfectamente interpretable para procesos heterogéneos debido a que está definida en términos de aleatoriedad (Dixon , 2002). Además la función K de Ripley recupera mayor cantidad de información referente al patrón de puntos. El sesgo observado cuando analizamos un patrón heterogéneo mediante la función K , puede ser corregido delimitando una subregión poligonal. Finalmente la función K se define como el número esperado de puntos adicionales dentro de una distancia s de un punto arbitrario, dividido por la densidad total de los puntos. En la práctica, un borde de corrección es necesario para evitar sesgar la estimación debido al no registro de los puntos fuera del polígono.

De manera concreta decimos que la función K de Ripley refleja el tipo, intensidad y rango del patrón espacial mediante el análisis de las distancias existentes entre todos los puntos, donde podemos distinguir ciertos valores en circunstancias específicas.

Bajo aleatoriedad total

$$K(d) = \frac{\lambda \pi d^2}{\lambda} = \pi d^2, \dots\dots\dots(2.6)$$

donde

Si $\hat{K}(d) < \pi d^2$ hay regularidad

Si $\hat{K}(d) > \pi d^2$ hay agregación

Por lo cual es un método que tiene un alto grado de confiabilidad en su aplicación para el análisis de la distribución de los puntos y por lo tanto su clasificación.

Además de que la función K de Ripley tiene la particularidad de que puede hacer un análisis gráfico que es indicativo del comportamiento de los valores analizados, Figura 2.1. La valoración gráfica nos facilita la identificación de los patrones puntuales; ya que, en algunas ocasiones la estimación de un número adimensional puede ser un poco confuso, sobretodo cuando se tiene poca práctica en el análisis de patrones puntuales. Por otro lado, el resultado del análisis no siempre es un arreglo al azar homogéneo; en estos casos se recurre a otros métodos que nos permitan definir la heterogeneidad de la distribución de los puntos, tal es el caso de las estimaciones núcleo (kernel).

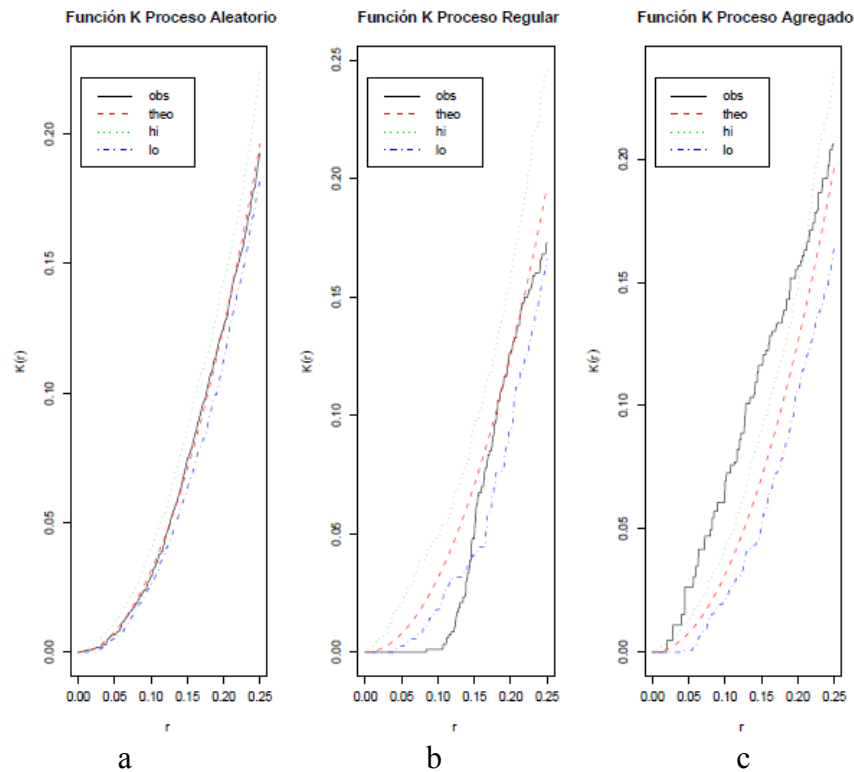


Figura 2.1. Gráficos generados a partir de la función K de Ripley, a) muestra un proceso de tipo aleatorio, se puede apreciar como tanto la línea de los observados como la de los teóricos se encuentran dentro de las bandas de confianza superior (hi) e inferior (lo), de la prueba del modelo; b) un proceso de tipo regular, en la parte inferior de la gráfica los datos observados están fuera del área de bandas, y a partir de la parte media a superior ya se encuentran dentro de las bandas; en c) se observa que el valor calculado teórico (- - -) como el de las bandas (- . - ., - - -) se encuentran por debajo de los valores observados (----), lo cual hace evidente al proceso tipo agregado. (Giraldo-Henao, 2011).

La estimación por núcleo o kernel o función núcleo, es un método de estimación de intensidad (densidad), el cual se utiliza para poder visualizar la presentación de los datos originales y facilitar la búsqueda de patrones, cuando la densidad no es uniforme (Miñarro, A. 1998; Servy, E. et al. 2005, Giraldo-Henao, 2011). Tiene la característica de que considera la existencia de correlación entre los datos, o si hay valores extremos o atípicos.

En el cálculo de la estimación Kernel, se genera una ventana que depende de la cercanía que existe entre los datos, a ésta ventana se le denomina también ancho de banda, la determinación del tamaño de la ventana es crítico a igual que la definición del núcleo o kernel (K), las cuales pueden ser de diverso tipo, figura 2.2

Núcleo	$K(t)$	Rango
Epanechnikov	$\frac{3}{4}(1-t^2)$	$ t < 1$
Gauss	$\frac{1}{\sqrt{2\pi}}e^{-(1/2)t^2}$	$ t < \infty$
Triangular	$1 - t $	$ t < 1$
Rectangular	$\frac{1}{2}$	$ t < 1$
Biweight	$\frac{15}{16}(1-t^2)^2$	$ t < 1$
Triweight	$\frac{35}{32}(1-t^2)^3$	$ t < 1$
Arco coseno	$\frac{\pi}{4} \cos \frac{\pi}{2}t$	$ t < 1$

Figura 2.2. Algunas de las funciones núcleo más comunes utilizadas para le estimación por núcleos. (Miñarro,1998).

De manera general la estimación Kernel se presenta como en la ecuación 2.7, dónde la diferenciación se hace por la elección de la función K o núcleo.

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n k\left(\frac{x - X_i}{h}\right), \dots\dots\dots(2.7)$$

donde h_n es el ancho de banda, n el número de elementos, x el valor que asume en la función, X_i el valor observado, K función Kernel, la cual puede ser de diferente tipo (Miñarro, A, 1998; Giraldo-Henao, 2011).

Se puede interpretar de la función núcleo, como una suavización del resultado de la suma de las protuberancias, generadas por las diferencias de las intensidades provenientes de los valores de los datos.

2.2. VECINO MÁS CERCANO

La determinación del vecino más cercano nos permite determinar la menor distancia que existe entre una serie de puntos dados. En éste trabajo se utilizó la rutina `spDistsN1`, del programa R. El cálculo se realizó mediante la aplicación de la ecuación de distancias euclidianas.

$$d_{(i,j)} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}; \forall i = 1, \dots, n; \forall j = 1, \dots, n \dots\dots\dots(2.8)$$

La función devuelve un vector que estimo la distancia entre un punto y una matriz de puntos de dos dimensiones, donde la primera columna es la longitud, y la segunda es la latitud.

2.3 AUTOCORRELACIÓN ESPACIAL

Cuando los valores están relacionados en momentos diferentes en el tiempo espacial hablamos de una autocorrelación, ésta es una medida del grado de similitud para una variable cuantitativa o categórica z entre puntos vecinos en el espacio. La autocorrelación surge cuando los términos de error del modelo no son independientes entre sí, es decir, cuando: $E(u_i, u_j) \neq 0$. para todo $i \neq j$. Entonces los errores estarán vinculados entre sí. La presencia de autocorrelación invalida el supuesto de independencia en la que se apoyan muchas de las pruebas estadísticas clásicos. Entendemos como autocorrelación a aquella relación entre los valores de una variable, que se puede atribuir a la forma en la que los puntos de muestreo de esta variable están dispuestos en el espacio (Camero y Rozas, 2006)

Un correlograma es un gráfico que muestra cómo cambia la autocorrelación espacial, expresada por un índice, como el I , en función de la distancia (en abscisas). La interpretación de la autocorrelación expresa la cercanía o lejanía de los eventos en el espacio, siendo una autocorrelación positiva cuando los valores similares tienden a agregarse; una correlación negativa cuando los valores similares tienden a separarse y que no hay autocorrelación cuando la disposición es al azar.

La disposición espacial de los eventos es el resultado de la heterogeneidad espacial, la cual se puede presentar debido a la no homogeneidad de otras variables correlacionadas con la variable de interés, por ejemplo los patrones espaciales de abundancia de especies biológicas, presentan una heterogeneidad espacial debido a la no homogeneidad de su hábitat en una región amplia.

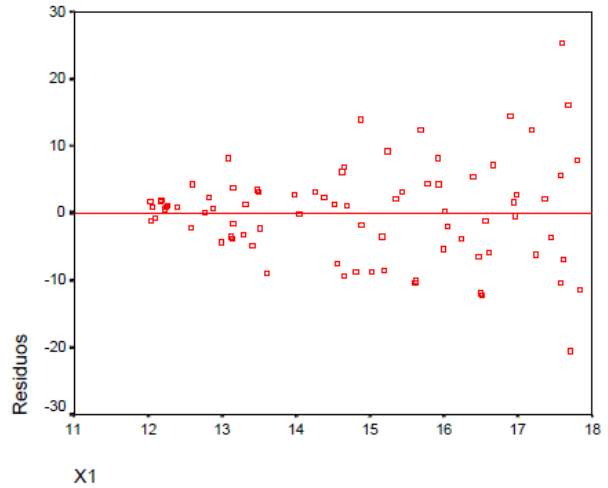
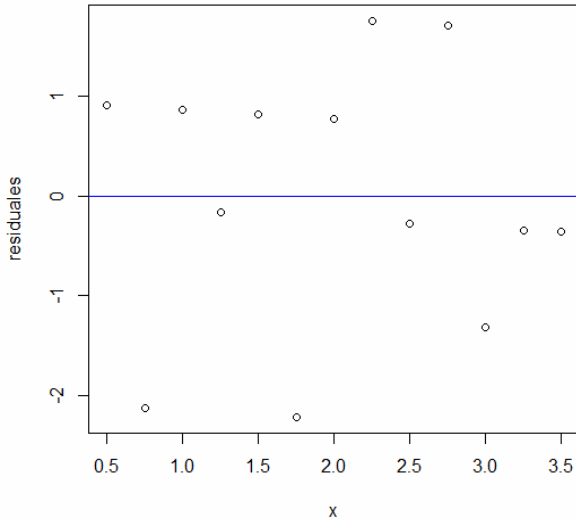
Es claro que, la utilización de la estadística espacial, nos ayuda a hacer un mejor análisis del comportamiento en la distribución espacial que se puede presentar en una especie determinada, ya que este tipo de estudios, incluye las dependencias existentes entre las variables que definen la disposición de los eventos. Por tal motivo para el ejercicio a desarrollar en este trabajo se aplicaran las herramientas de análisis de la estadística espacial. Lo que permitirá conocer la distribución espacial de las nutrias de río dentro de la cuenca Apatlaco-Tembembe; con lo cual se provoca una mejor planeación de muestreos para trabajos posteriores, ya que, la distribución de los organismos nos indican, en que zonas se debe aplicar un mayor esfuerzo y en cuales no, en caso de que no sea homogénea, además facilita la maximización de los recursos, da una mayor confiabilidad a los resultados obtenidos y permite la realización de inferencias.

2.4. REGRESIÓN LINEAL

La metodología de regresión lineal permite evaluar y modelar la relación entre variables, es decir pretende obtener una ecuación con la finalidad de estimar o predecir, el comportamiento de una variable respuesta respecto a una variable regresora en caso del modelo lineal simple, para el caso múltiple es una variable respuesta en relación de varias variables regresoras. Sin embargo, para ambas técnicas se parte de los supuestos que: los errores se distribuyen de manera normal, la varianza es constante y la variable respuesta se relaciona linealmente con la(s) variable(s) independiente(s). Para verificar éstos supuestos se vale de la utilización de gráficas de dispersión, análisis de residuos (Figura 2.3) y de la gráfica de cuantil-cuantil (qqplot), figura 2.4. En el análisis de los residuos, se debe de calcular el valor de los residuales, los cuales se obtienen mediante la diferencia entre los valores estimados por el modelo y los valores observados; en

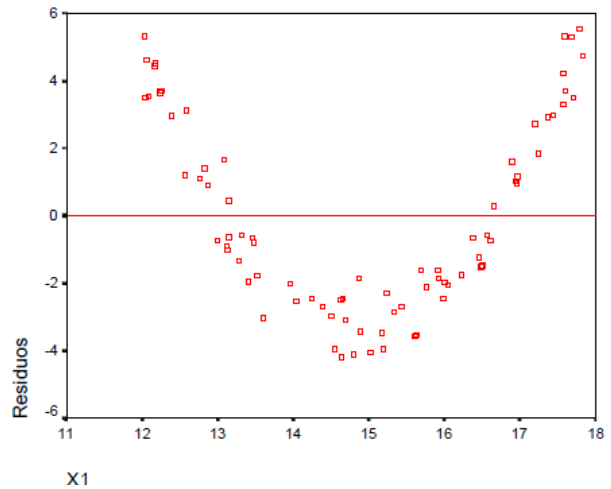
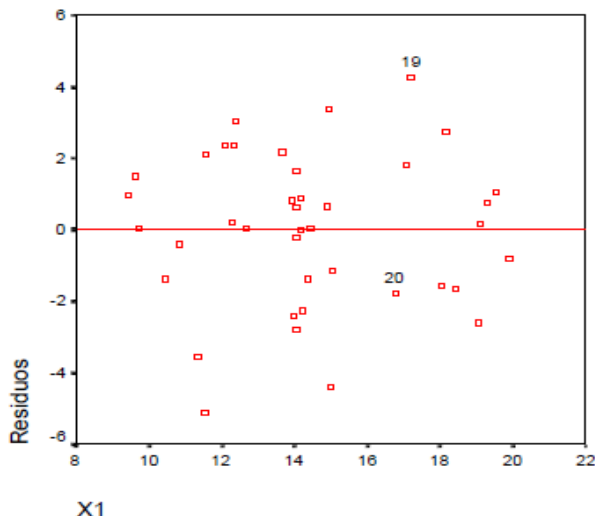
algunos casos se utilizan los residuos estandarizados, los cuales se calculan dividiendo por la desviación estándar de todos los residuos a cada i -ésimo valor salvo por el mismo y al graficarlos éstos tiene que seguir una distribución normal.

Las gráficas a evaluar se muestran en las figuras 2.3 y 2.4



a) la relación de los datos se puede interpretar como lineal.

b) el análisis de los res, muestra una tendencia de función creciente en la varianza



c) Se observa una variación, ya que primero empieza a aumentar y luego tiende a disminuir

d) Este tipo se considera de tipo no lineal, particularmente cuadrático.

Figura 2.3 Gráfica del análisis de los valores estimados vs. los residuales, en diferentes casos a) lineal; los tipos b y c muestran heterocedasticidad, la primera creciente y el segundo es variable; d) es de tipo no lineal.

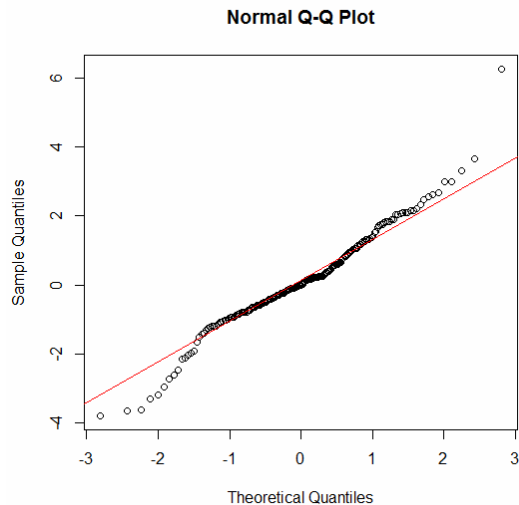


Figura 2.4 Análisis de normalidad de los datos, que permite contrastar la simetría de la distribución de los residuos.

Cuando se han realizado estos análisis y se observa que el comportamiento no es lineal, es usual recurrir a realizar una transformación de la variable respuesta o de las variables regresoras, ya sea por logaritmos o mínimos cuadrados ponderados, sin embargo, estas transformaciones no siempre corrigen la no normalidad, heterocedasticidad o no linealidad o bien si la corrigen, la interpretación de los datos no es posible dentro de la fenomenología del objeto de estudio, por ejemplo Si decimos que la abundancia de pino silvestre es función de la elevación tenemos una idea más o menos clara de lo que esto puede significar. Si la relación es positiva, un aumento de la elevación aumentaría la abundancia de esta especie. Pero ¿qué quiere decir que el logaritmo de la abundancia de pino silvestre es función de la elevación? (Cayuela, 2010); más aún si se hacen transformaciones de tipo exponencial o potencial, así que una alternativa en vez de las transformaciones es la aplicación de un modelo lineal generalizado

2.5 MODELO LINEAL GENERALIZADO

Los modelos lineales generalizados (GLM, por sus siglas en ingles), son una extensión de modelos lineales, pero permiten aplicarse a distribuciones no normales en los errores, es decir binomiales, poisson gamma, entre otras, además de ser aplicables con aquellas donde su varianza no es constante. Particularmente los modelos lineales generalizados con distribuciones donde se requiere realizar una función recíproca y se tienen errores de tipo gamma son útiles cuando existe un coeficiente de variación constante, es decir cuando la varianza aumenta según aumenta la media de la muestra de manera constante (Cayuela, 2010).

La cantidad de varianza explicada por el modelo lineal generalizado es la devianza, esta nos da una idea de la variabilidad de los datos. Para su cálculo se compara el valor de la devianza del modelo nulo (*Null deviance*, que es aquella que sólo tiene la constante) con la devianza residual (*Residual deviance*, es la que tiene la constante y la(s) variable(s)), con ello obtenemos una idea de cuanto de la variabilidad no es posible explicar por el modelo. Así que entre más pequeño el valor de la devianza mejor será el ajuste.

$$D^2 = \frac{\text{Devianza.modelo.nulo} - \text{Devianza.residual}}{\text{Devianza.modelo.nulo}} * 100 \dots\dots\dots 2.9$$

Cabe mencionar que en los modelos lineales generalizados se debe considerar que no hay un modelo que sea valido para tratar de explicar el fenómeno, y que lo importante es la evolución y criterios que se utilizan, para la toma de decisiones respecto al modelo que será el de trabajo. Así que la exploración de los datos al igual que en los modelos de regresión no se debe de pasar por alto. Adicionalmente puede utilizarse el análisis de los índices de Cook y el apalancamiento (Cayuela, 2010). Una vez que se han realizado estos análisis un punto a considerar y que debe de tener gran peso es el principio de parsimonia, es decir se busca un modelo tan simple como sea posible; ya que el modelo más simple, procurando eliminar las redundancias. Esto implica que hay que eliminar a las variables no significativas, el modelo debe de poder reportar una explicación lógica y que sea posible relacionar con la teoría.

Capítulo 3: APLICACIÓN DE MÉTODOS DE ANÁLISIS DE DISTRIBUCIÓN ESPACIAL EN UNA POBLACIÓN DE NUTRIAS DE RÍO (*Lontra longicaudis annectens*) DENTRO DE LA CUENCA APATLACO-TEMBEMBE

3.1 OBJETIVO

El objetivo de este trabajo es aplicar la metodología de la estadística espacial, para demostrar como la aplicación de dichos métodos abren camino a nuevos conocimientos y manejos de material, de una forma eficiente y confiable.

Para lo cual se estableció un objetivo biológico, el cual será resuelto mediante el análisis de la estadística espacial.

Objetivo biológico: Determinar la preferencia del hábitat de la nutria de río neotropical (*Lontra longicaudis annectens*) (Ver anexo), en la cuenca Apatlaco-Tembembe, México.

3.2 PROCEDIMIENTO

Se revisó la información con la que se contaba para la región de la cuenca Apatlaco-Tembembe. Los registros con los que se realizó el estudio de la nutria de río neotropical (*Lontra longicaudis annectens*) fueron reportados desde 1996 a 2004 para la cuenca Apatlaco-Tembembe, México; en el trabajo de Cirelli Villanova, 2005. Así mismo se obtuvieron los registros geográficos de los 32 municipios, en los cuales se ubicaron los avistamientos de la nutria de río, todos ellos pertenecientes a la Cuenca Apatlaco-Tembembe, los cuales son: Acamixtla, Acuitlapan, Amacuzac, Atlatlahuacan, Ayala, Buenavista, Coatlan, Cuernavaca, Emiliano Zapata, Huitzilac, Iguala, Ixcateopan, Ixtapan de la Sal, Jojutla, Jolapan, Juitepec, Mazatepec, Miacatlan, Pedro Ascencio, Ocuilan, Pilcaya, Puente de Ixtla, Taxco de Alarcón, Sultepec, Temixco, Tetipac, Teotlalco, Teloloapan, Tepalcingo, Tetecala, Tlaquitenango, Tlaltizapan, Tlayacapan, Tonalico, Xochitepec, Yautepec, Zacatepec, Zacualpan, Zumpanhuacan.

Por las características de los datos (retrospectivo, transversal, observacional y descriptivo) podemos clasificar éste estudio como una encuesta descriptiva; en el cual, al aplicar el análisis espacial nos permite generar una idea para estudios subsecuentes con objetivos específicos respecto a las preferencias de hábitat de la nutria de río.

Para la ubicación geográfica de todos los puntos, tanto del polígono como de los avistamientos y cauces, se realizó la conversión de unidades UTM (Universal Transverse Mercator), que es un sistema de coordenadas basado en la proyección cartográfica, es decir, es un sistema que establece una relación ordenada entre los puntos de la superficie curva de la tierra y los de una superficie plana (mapa). A un sistema de coordenadas geográficas, para lo cual se utilizó el elipsoide WGS84 (World Geodetic System 84) de la zona 14 del hemisferio norte. El WGS84 es un sistema de coordenadas geográficas mundial el cual consiste en un patrón matemático de tres dimensiones que representa la tierra por medio de un elipsoide y con el cual se permite localizar cualquier punto de la tierra; para realizar el cambio de unidades UTM a WGS84 se utilizó el convertidor de la página <http://www.atlascajamarca.info/conversor/>.

Una vez obtenidos los datos en forma geográfica se transformaron en unidades de kilómetros mediante las fórmulas 3.1, 3.2 y 3.3. La exploración de los datos se realizó utilizando el lenguaje estadístico R, particularmente el paquete `splancs` para el análisis espacial.

$$\cos^2 = (\cos[0.5(18.2 + LatOb)])^2 \quad (3.1)$$

donde; 18.2= latitud de la orilla del plano

LatOb= Latitud Observada

Distancia en X para el plano cartesiano del polígono

$$60 * \sqrt{(-100 - LongOb)^2 \cos^2} \quad (3.2)$$

donde; -100=Longitud base del plano

LongOb= Longitud observada

Distancia en Y para el plano cartesiano del polígono

$$60(18.2 - LatOb)^2 \cos^2 \quad (3.3)$$

Ejemplo

Originales		UTM	Convertidas	WGS84	Coordenadas para la distancia en kilómetros	
LongW	LatN	Estado	Longitud	Latitud	x	Y
E474459	2052557	Morelos	-99.24202	18.56358	17.73120909	3.173324395

Con los valores ya convertidos en kilómetros primero se realizó un mapa de la cuenca Apatlaco-Tembembe Figura 3.1; asimismo se ubicaron los puntos de avistamiento en el mapa con los cauces (Figura 3.2); posteriormente se realizó otro mapa en donde se incluyó un polígono con una zona de amortiguamiento (buffer) de 5 km a los 71 cauces (Figura 3.3).

Cuenca Apatlaco-Tembembe

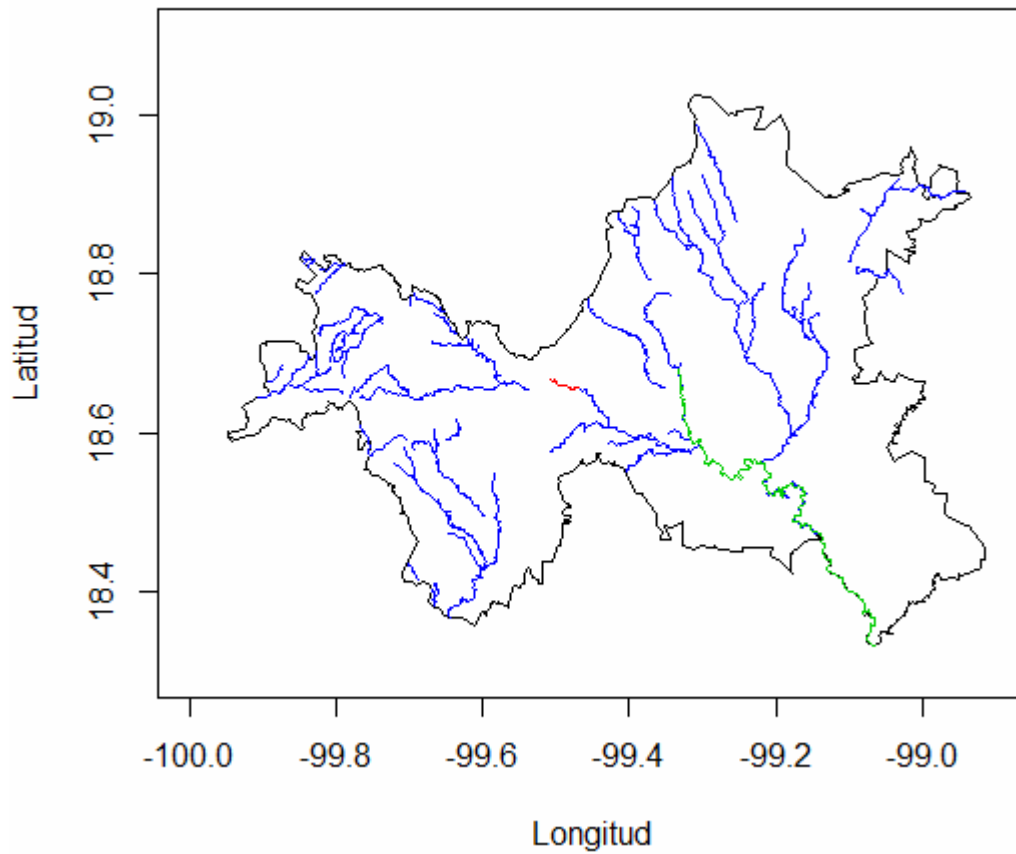


Figura 3.1. Mapa en coordenadas geográficas de la cuenca Hidrológica Apatlaco- Tembembe con los 71 cauces contenidos en los 32 municipios de los tres estados, Morelos, Estado de México y Guerrero. El cauce 71 se muestra en rojo y el cauce 9 en verde, ambos cauces son los que se asocian a un mayor registro de avistamientos de nutria, de 10 y 9 respectivamente.

Cuenca Apatlaco-Tembembe

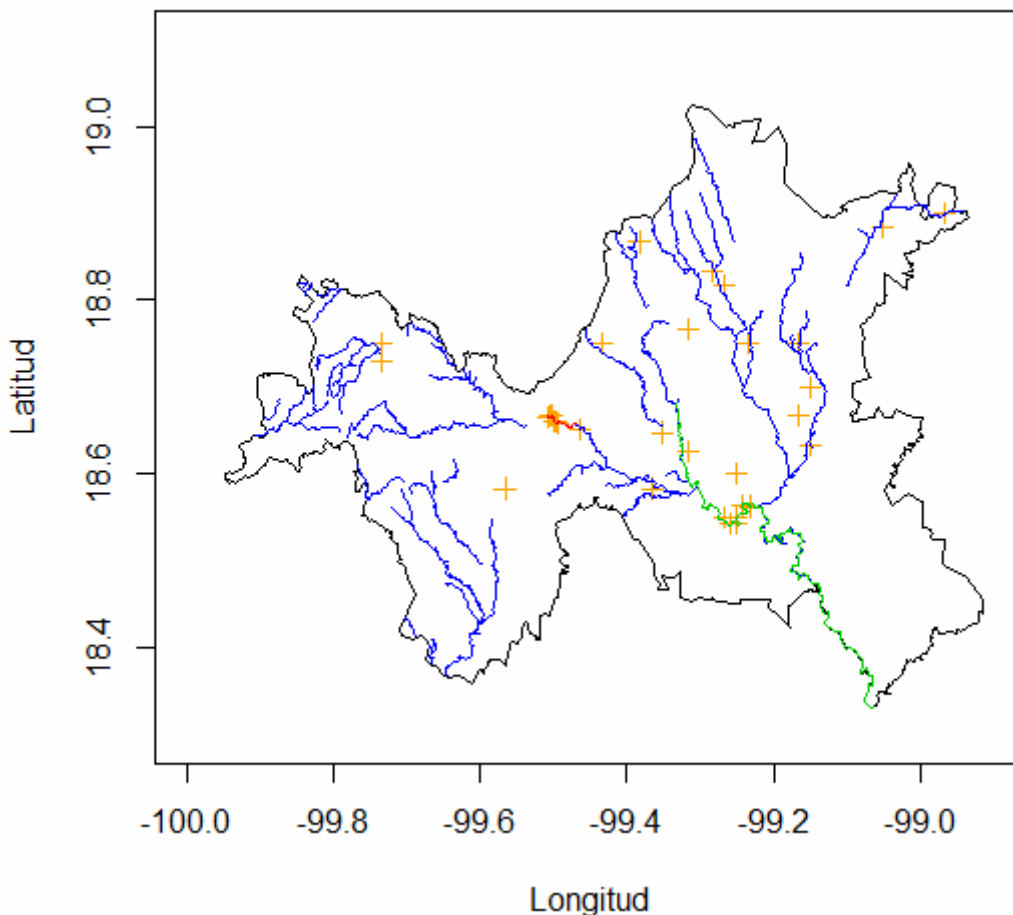


Figura 3.2. Mapa en coordenadas geográficas de la cuenca Hidrológica Apatlaco- Tembembe con los 71 cauces y los 38 avistamientos de nutrias (+), contenidos en los 32 municipios de los tres estados, Morelos, estado de México y Guerrero. El cauce 71 se reporta en rojo, y el cauce 9 se reporta en verde y las nutrias en naranja. Se observa que hay una mayor concentración de avistamientos cerca del cauce 71, en la parte extrema superior este, asimismo hay otra concentración apreciable de nutrias en la parte media del cauce 9.

Polígono Buffer de 5Km en cauces con avistamiento

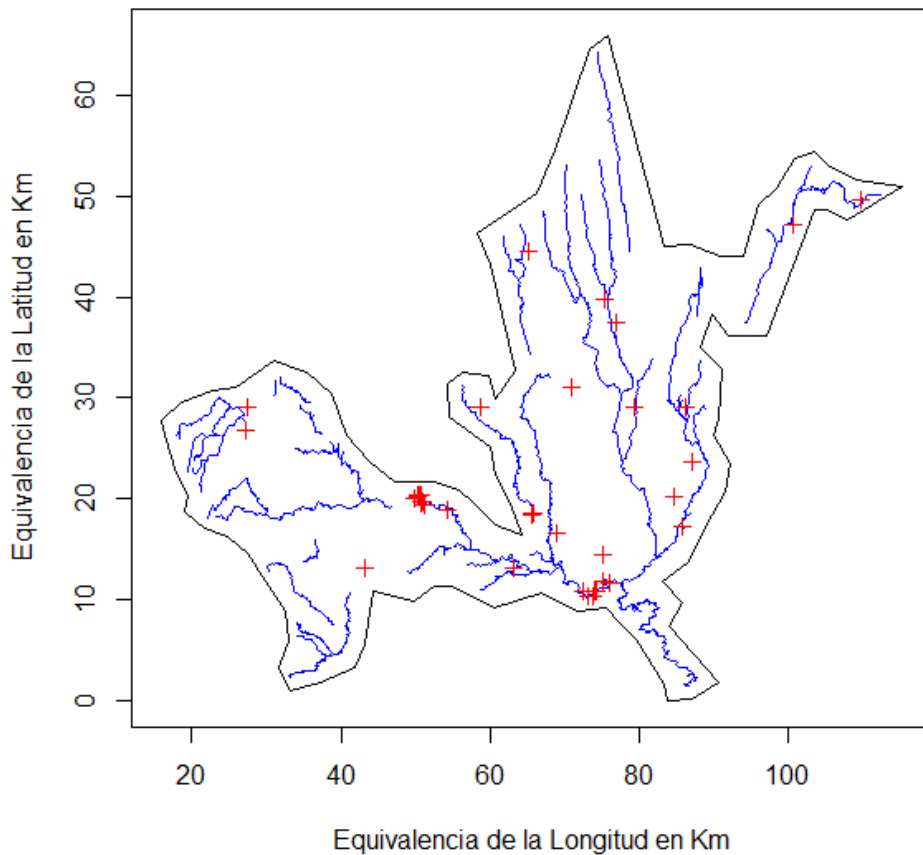


Figura 3.3, Mapa con la conversión de unidades geográficas a kilómetros del Polígono buffer, que contiene los 71 (—) cauces y los 38 avistamientos (+) de las nutrias de Río en la cuenca Apatlaco-Tembembe. La distancia promedio del las líneas de borde del buffer a los cauces es de 5km.

El siguiente paso fue el de determinar la distancia de los avistamientos al cauce más cercano, para ello se utilizó la rutina del vecino más cercano mediante lenguaje R y la rutina `spDistsN1` (ecuación 2.8); la cual calcula las distancias bajo un modelo euclidiano y que consiste en determinar la distancia de un punto definido por sus coordenadas, a otra serie de puntos buscando la distancia más corta entre los puntos (Ver anexo), con estos resultados se elaboró una gráfica de distribución de frecuencias (Figura 3.4) y se obtuvo un cuadro de la relación de las frecuencias de avistamientos relacionadas a los cauces de la cuenca Apatlaco-Tembembe (Cuadro 3.1).

Distancia del Avistamiento al cauce más cercano

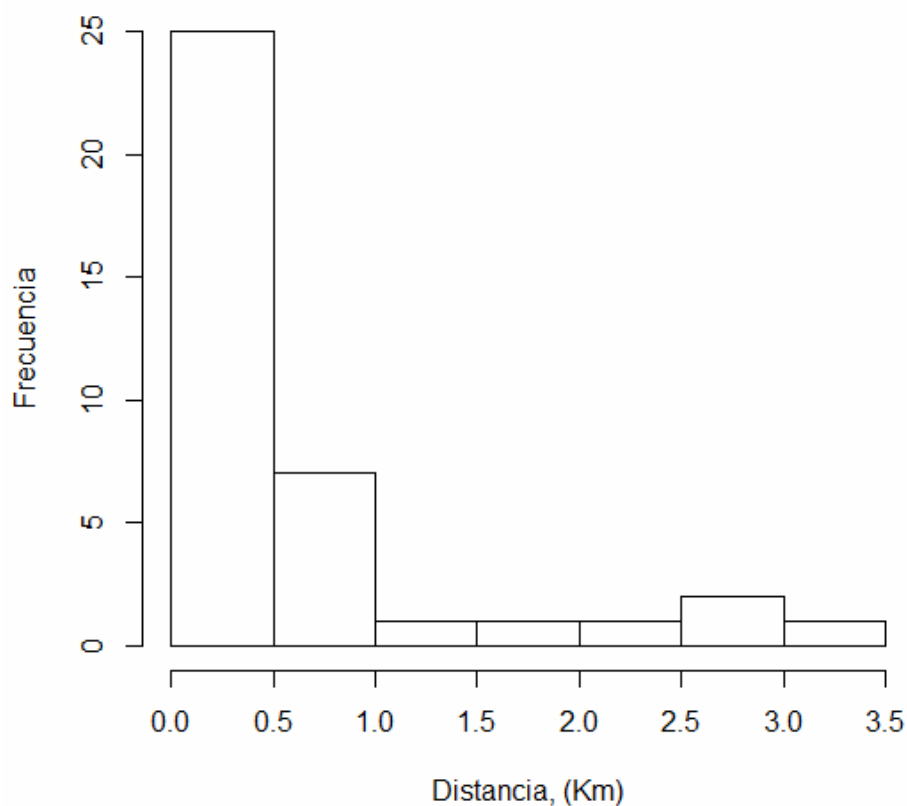


Figura 3.4. Distribución de frecuencias de las distancias de los 38 avistamientos de nutrias de río a los cauces en la cuenca Apatlaco-Tembembe. Se observa que en la mayoría de los avistamientos se registró una distancia menor a 0.5 km; y que después de un kilómetro los registros son prácticamente únicos.

Cauce	2	6	8	9	10	12	15	20	21	22	24	25	26	50	60	71
Frecuencia	2	1	1	9	3	1	1	1	2	1	1	1	1	2	1	10

Cuadro 3.1. Relación de frecuencias de avistamientos de nutrias relacionadas a los cauces de la cuenca Apatlaco-Tembembe. Se puede observar que el cauce 71 es el que tiene mayor registro de avistamientos, seguido del cauce 9.

El polígono buffer se dividió en cuadrantes de 10 x 10 Km (Figura 3.5) y en cada poligonito se determinó la densidad de nutrias de acuerdo a la ecuación 2.1 (Cuadro 3.2) y la función K de Ripley para verificar la aleatoriedad de los datos. Como resultado que los datos no se encontraban de forma aleatoria se procedió a recalcular la densidad mediante la función Kernel. Para ello a cada cuadrante que no contuviera registro de avistamiento y que estuviera dentro del área del polígono se le calculó el valor del centroide (Ver anexo).

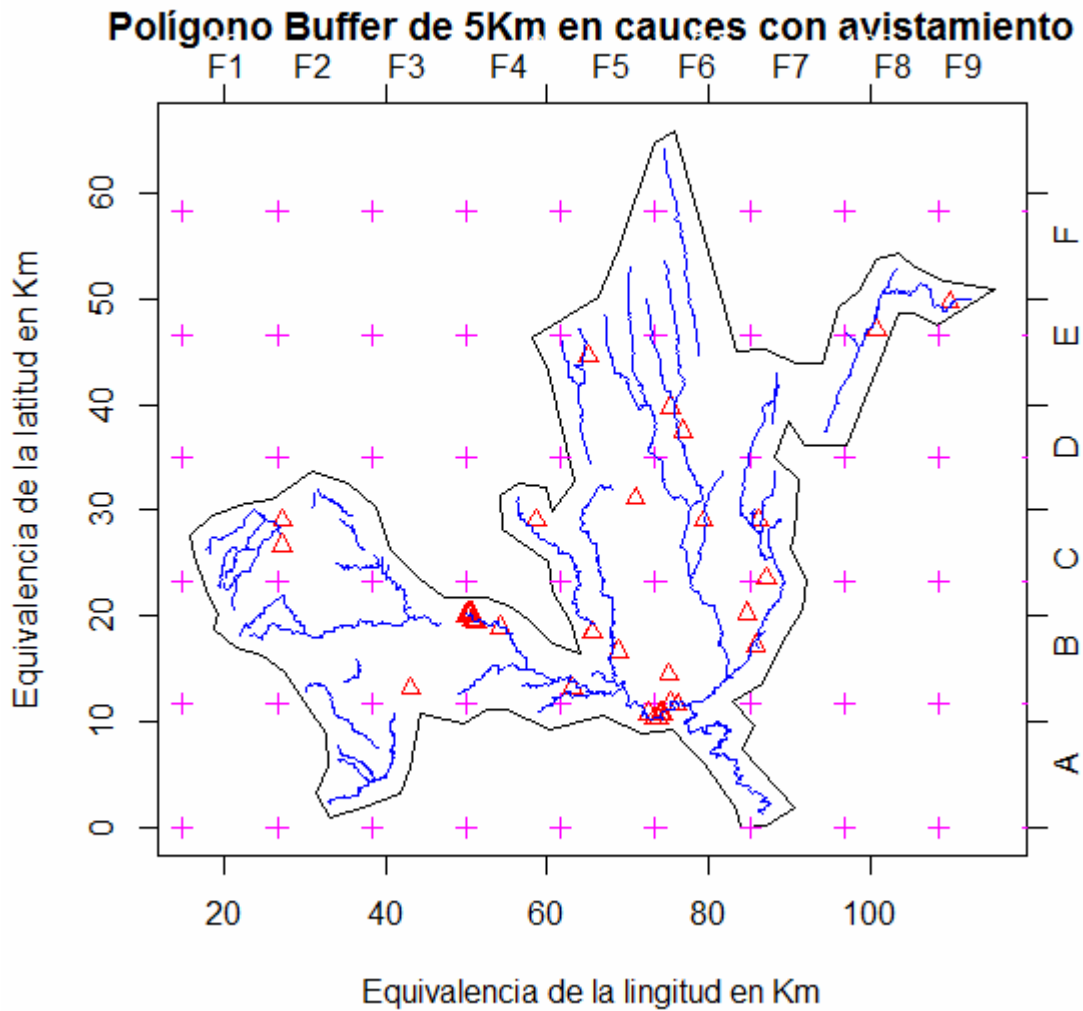


Figura 3.5. División de cuadrantes (10 x 10 Km) del polígono donde se avistaron las nutrias (Δ). El nivel más inferior de izquierda a derecha es el A1 hasta A9, donde los siguientes niveles se identifican por las letras B, C, D, E y F; en los cuales se mantuvo la sucesión numérica independientemente del renglón, hasta llegar al último en la parte superior derecha con el cuadrante F54. Este mapa se utilizó como disposición de las coordenadas y tamaños de cuadrantes para generar la función Kernel.

Cuadrante	Km ²	# de nutrias por cuadrante	λ
B13	96.8643	11	0.11356093
A5	23.32231	2	0.0857548
E45	14.35314	1	0.06967117
A6	69.81203	3	0.04297254
B16	47.41546	2	0.04218034
B14	128.307	4	0.03117523
C25	64.5926	2	0.0309633
C22	44.43195	1	0.02250633
B15	135.0349	3	0.02221648
C20	105.465	2	0.01896364
E44	61.56818	1	0.01624216
D33	131.6513	2	0.01519165
C23	126.5189	1	0.00790396
D32	130.1476	1	0.00768358
B12	132.9832	1	0.00751975
C24	136.6386	1	0.00731858
A2	63.66827	0	0
A3	53.19587	0	0
A4	15.71915	0	0
A7	19.90596	0	0
B10	51.78617	0	0
B11	127.5928	0	0
C19	68.99485	0	0
C21	18.03101	0	0
D31	14.59242	0	0
D34	101.279	0	0
D35	34.51058	0	0
E40	2.918481	0	0
E41	79.34672	0	0
E42	85.76261	0	0
E43	2.631529	0	0
F50	10.4529	0	0
F51	27.86355	0	0

Cuadro 3.2. Valores obtenidos de las densidades (λ) dentro del polígono buffer para cada cuadrante. La identificación de la ubicación del cuadrante en el mapa es de la parte inferior izquierda hacia la derecha iniciando con el A1 hasta A9 para el primer nivel, siendo cada nivel designado por las letras B, C, D, E y F. También se puede ver los valores del área en km² para cada cuadrante.

Asimismo se realizó el análisis de regresión por el método lineal generalizado, para determinar que variable es de mayor influencia para la distribución de las nutrias; en este caso se considero la densidad como el resultado de las variables, las cuales fueron: distancia al cauce, uso de suelo y cuadrante en el que se registra el avistamiento. En el caso de la variable uso de suelo se estableció un contraste, dado que los valores de la variable son de tipo categórico (Ver anexo).

Capítulo 4: RESULTADOS Y ANÁLISIS

La distribución de frecuencias de la distancia entre los avistamientos y el cauce más cercano se presenta en la figura 3.4. En donde se puede apreciar que los organismos se encuentran en mayor proporción a una distancia menor de 0.5 Km

Los resultados del cuadro 3.2 nos indican que la densidad no es homogénea para todos los cuadrantes; lo cual se hace evidente si observamos el valor de densidad en el cuadrante B13, con 11 nutrias por área.

La determinación de la función K de Ripley, Figura 4.1, muestra que el patrón de distribución es de tipo agregado, esto se nota ya que, al graficar los valores observados ($^{\circ}$), éstos se encuentran muy por arriba de los valores calculados de los límites del intervalo teórico (en rojo y azul), lo cual nos confirma que la densidad no es de forma homogénea, por lo tanto se tuvo que recalcular la densidad tomando en cuenta que hay un patrón agregado para los datos. Esto se logró mediante la función Kernel, para lo cual se utilizó el mapa con cuadrantes de 10 x 10 metros, Figura 3.5. El resultado del nuevo cálculo de la densidad se observa en la Figura 4.2, donde la interpretación de las densidades se tiene de forma colorimétrica.

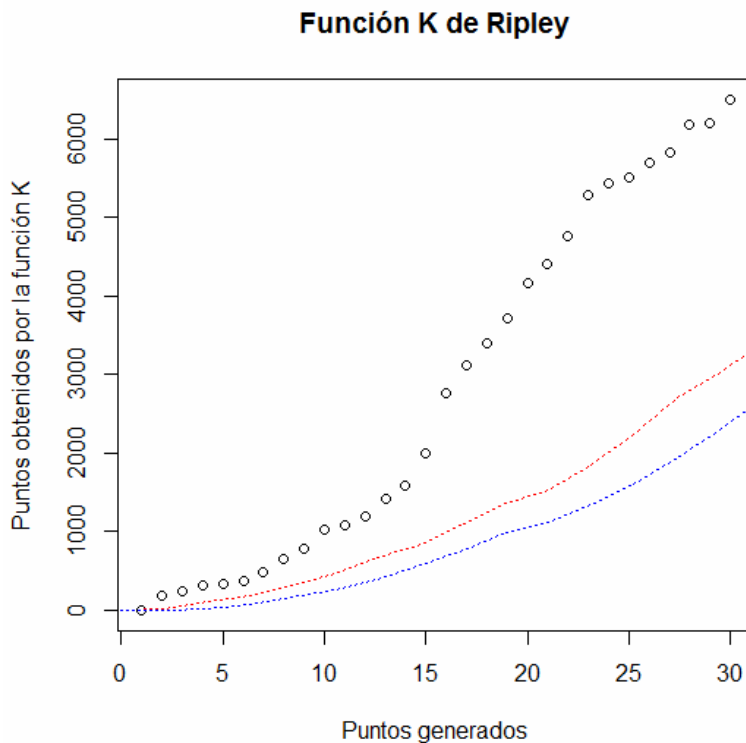


Figura 4.1. Grafica de la función de la K de Ripley, en la cual se puede apreciar que los datos no se distribuyen de forma aleatoria, más bien resaltan que se presentan en una distribución de tipo agregados. Esto debido a que los límites superior (.....) e inferior (.....) del intervalo se encuentran por debajo del valor de los valores generados por la función K de Ripley respecto a los datos observados en los avistamientos de las nutrias.

Estimación colorimétrica de la densidad mediante la función kernel

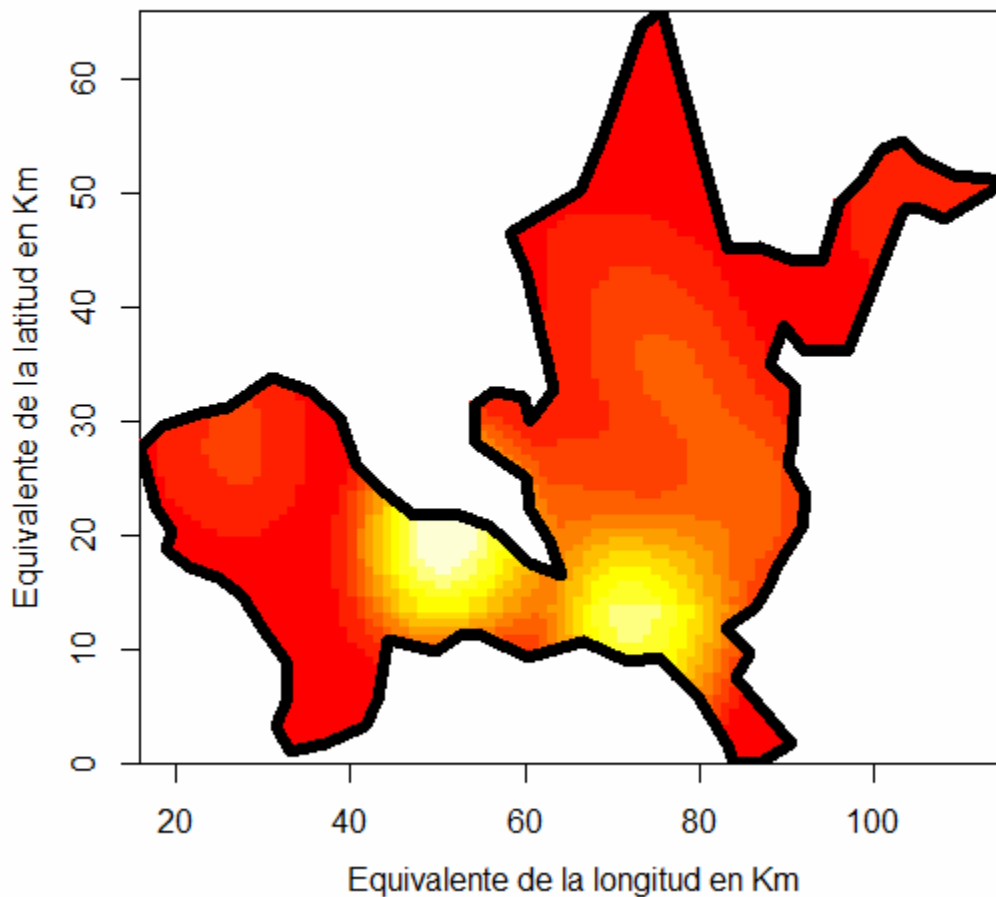


Figura 4.2. Grafica de estimación de la densidad de nutrias mediante la función Kernel, en el polígono de la porción de estudio de la cuenca Apatlaco-Tembembe . Se puede apreciar que las zonas con color claro son las de mayor densidad. Lo que nos indica que la distribución espacial de las nutrias de río no es homogénea y que es mayor en dos zonas, que son las correspondientes al los cauces 71 y 9.

El resultado de la generación de la función Kernel, figura 4.2, muestra que hay dos zonas de alta densidad de nutrias, ambas zonas coinciden con la alta densidad de la determinación de densidad anterior, que es donde hay mayor número de registros de avistamientos, ambas zonas se encuentran en tipo de suelo de selva caducifolia. Otra característica que comparten es que en ambos casos las zonas se encuentran limitadas con bosque de encino, áreas sin vegetación y zona de manejo agrícola.

En la figura 4.3 se observa que hay una subdivisión en los valores de las densidades obtenidas a partir del modelo ajustado para los tipos de suelo que reportaron mayor registro de avistamientos. Los agrupamientos responden a la cercanía física del registro del avistamiento

respecto a un cauce; es la proyección del agrupamiento de los avistamientos por zonas; particularmente en el tipo de suelo 2 el punto de mayor densidad ajustado corresponde al avistamiento número 6, el cual también tiene la menor distancia al cauce número 71, además de que se encuentra en la zona en donde se han registrado el mayor número de avistamientos, lo cual resalta que la distancia es relevante para el registro de nutrias.

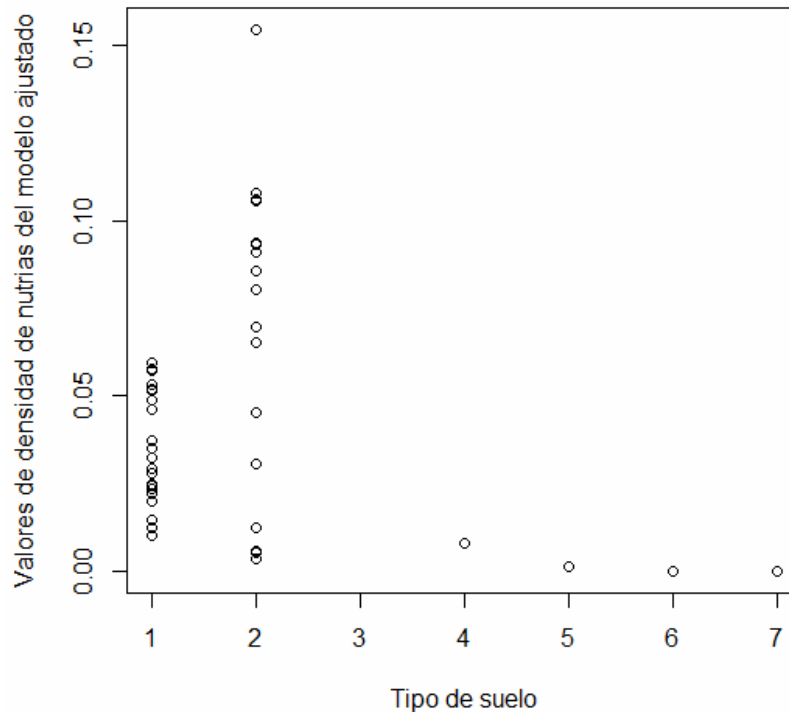


Figura 4.3. Se observa que la mayor densidad de nutrias se tiene los tipos de suelo de manejo agrícola y de selva caducifolia, además de resaltar que la disposición de los organismos es en grupos dentro de las zonas con ese tipo de suelo; cabe considerar la posibilidad de incluir algunas otras variables para poder tener una mejor explicación del comportamiento agregado de las nutrias.

En el análisis de regresión se clasificaron siete tipos de suelo asociados a los avistamientos y a los centroides; los cuales fueron: 1) Manejo agrícola, pecuario y forestal, 2) Selva caducifolia, 3) Bosque de coníferas distinto de pino, 4) Laguna, 5) Bosque de encino, 6) Área sin vegetación, 7) Bosque mesófilo de montaña. En el cuadro 4.1 se puede observar las salidas de los resultados con los dos análisis que se hicieron, en el primero, *a*, se consideraron todas las variables, para el segundo, *b*, se juntaron los tipos dos tipos de bosque, bosque de coníferas distinto de pino y bosque de encino.

En la realización de los dos modelos se utilizó el tipo de suelo de manejo agrícola se como contraste. En el caso *a* se observa que la distancia es el más significativo y en segundo grado los tipos de suelo 5, 6 (Bosque de encino y Área sin vegetación, respectivamente), el intercepto, y por último el tipo de suelo 2 (selva caducifolia). En el caso *b* del análisis los significativos son el intercepto, la distancia, bosque de encino y área sin vegetación, es claro que el grado de significancia es menor, sin embargo haciendo un análisis general de todo lo que se obtiene en la

salida nos quedaremos con el análisis *b*; esto debido a que en el en ambos hay ausencia de normalidad (figura 4.4 *a.2* y *b.2*), el análisis de los residuales con los ajustados muestra que hay un dato influyente y que no hay una distribución de tipo normal (figura 4.4 *a.1* y *b.1*); así mismo en la grafica de los residuales contra el apalancamiento (figura 4.4 *a.3* y *b.3*) se observa que se tiene menor efecto de datos que puedan ser influyentes en el caso *b*.

(a) Sin hacer conversiones del tipo de suelo

Summary(modelofinal)

Call:

```
glm(formula = (datlutre$Den + 1e-07) ~ datlutre$Distancia +
as.factor(datlutre$Suelo),
family = Gamma(link = "inverse"), x = T)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-4.6277	-0.3502	0.0000	0.1364	1.5696

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.541e+01	4.446e+00	3.466	0.001138 **
datlutre\$Distancia	3.300e+01	7.971e+00	4.141	0.000143 ***
as.factor(datlutre\$Suelo)2	-9.505e+00	4.516e+00	-2.105	0.040671 *
as.factor(datlutre\$Suelo)3	1.700e+02	1.323e+02	1.285	0.205134
as.factor(datlutre\$Suelo)4	1.466e+01	9.122e+01	0.161	0.872994
as.factor(datlutre\$Suelo)5	1.000e+07	3.132e+06	3.192	0.002517 **
as.factor(datlutre\$Suelo)6	1.000e+07	2.859e+06	3.497	0.001039 **
as.factor(datlutre\$Suelo)7	1.000e+07	7.004e+06	1.428	0.159984

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.4905803)

Null deviance: 435.04 on 54 degrees of freedom
Residual deviance: 113.98 on 47 degrees of freedom

(b) Con las conversiones del tipo de suelo tipo 3 → 5

summary(modelofinal)

Call:

```
glm(formula = (datlutre$Den + 1e-07) ~ datlutre$Distancia +
as.factor(datlutre$Suelo),
family = Gamma(link = "inverse"), x = T)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-4.6279	-0.8801	0.0000	0.1364	2.9091

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.542e+01	7.394e+00	2.085	0.0424 *
datlutre\$Distancia	3.297e+01	1.325e+01	2.488	0.0164 *
as.factor(datlutre\$Suelo)2	-9.508e+00	7.511e+00	-1.266	0.2116
as.factor(datlutre\$Suelo)4	1.474e+01	1.517e+02	0.097	0.9230
as.factor(datlutre\$Suelo)5	8.359e+02	4.103e+02	2.037	0.0472 *
as.factor(datlutre\$Suelo)6	1.000e+07	4.755e+06	2.103	0.0407 *
as.factor(datlutre\$Suelo)7	1.000e+07	1.165e+07	0.858	0.3949

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 1.356827)

Null deviance: 435.04 on 54 degrees of freedom
Residual deviance: 201.81 on 48 degrees of freedom

Cuadro 4.1 Salidas de resultados de los Modelos Lineales Generalizados para el análisis de distribución de nutrias en la cuenca Apatlaco-Tembembe , Donde el modelo *a* tiene los siete tipos de uso de suelo y el modelo *b* se utilizaron los dos tipos de suelo (3 y 5) como un solo tipo.

En ambos casos el análisis de residuos muestra que es una función decreciente de acuerdo a los datos y que en el caso *b* es más marcada la intensidad de la variación decreciente, cabe mencionar que el análisis es de una función inversa, así que en realidad lo que indica es que la función es creciente, lo que nos indica que a menor distancia se tendría una esperanza de mayor densidad de nutrias, lo que es correspondiente ya que, el punto que induce éste efecto es el dato 6, que corresponde al que tiene un registro de mayor densidad a una menor distancia de los cauces de acuerdo al registro de las nutrias.

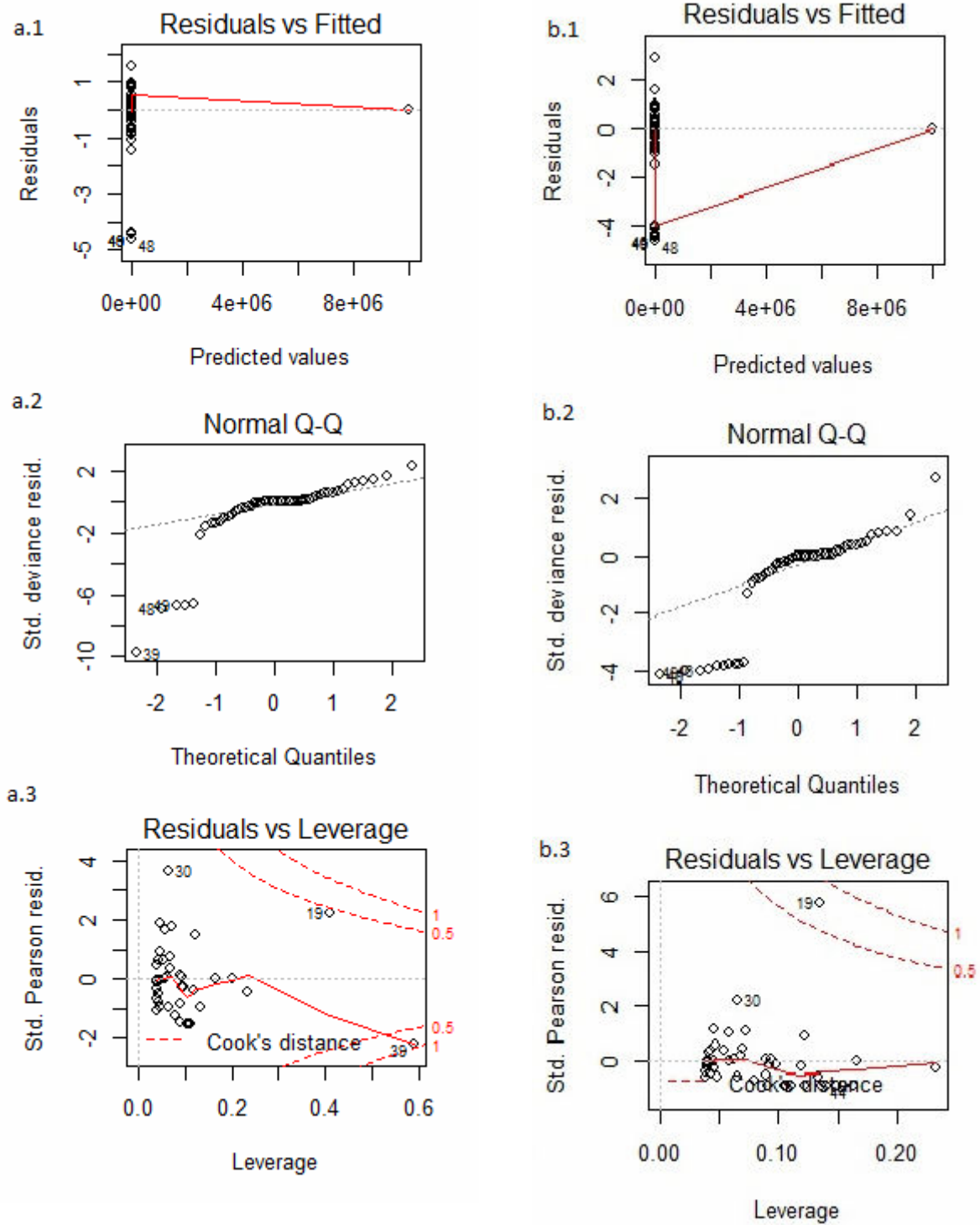


Figura 4.4. Izquierda a) análisis de residuos con las siete variables de tipo de suelo, derecha b) análisis de residuos con la variable 3 y 5 fusionadas y tratadas como tipo 5.

En un análisis individual de los residuos, figura 4.5, se identificaron algunos de los datos de los registros, en los cuales se observa que en ambos casos los datos 19 (parte superior izquierda) y el 39 (parte inferior izquierda), considerando que se debe interpretar como una función inversa, inducen un decremento en la función y si tenemos en cuenta que, el dato 19 es el avistamiento que se encuentra a mayor distancia de los cauces (3.2 km), y el segundo en menor densidad reportado (0.00751975), corroboramos la expectativa de que en medida que se aleje uno de los cauces será menos probable el avistamiento de las nutrias. Por otro lado se ve como la mayoría de los datos que tiene reportado valores de densidad cero de nutrias se encuentran en la parte inferior izquierda, lo que nos indica una menor probabilidad de avistamiento. Ahora bien retomando las graficas *a.3* y *a.4* de la figura 4.4, en la primera se puede observar que el dato 39 se encuentre cercano a los límites de los datos que podrían considerarse como influyentes, así que se realizaron graficas individuales del análisis de la distancia de Cook para ver cuales son los datos que se pueden considerar influyentes para cada caso, figura 4.6, en dicha figura se observa como los datos 19 y 39 son los que generan mayor grado de influencia cuando se tienen los siete tipos de usos de suelo, y que cuando se fusionan las variables bosques el único que hace apalancamiento es el dato 6, que ya mencionamos es el de mayor densidad y menor distancia. Otros de los datos que podemos analizar de las salidas de los casos *a* y *b* en los modelos lineales generalizados, son los resultados de la devianza, que de acuerdo a la ecuación 2.9, la devianza, en general, se espera sea de menor valor, para el caso *a* = 8.59 y para el caso *b* = 7.32, siendo el menor el valor obtenido para el análisis *b*, lo que indica que tiene menor variabilidad no explicada. Además que los resultados para el análisis de devianza, cuadro 4.2, podemos observar que los datos de la devianza calculados para las variables tipo de suelo es menor para el caso *b* que para el caso *a* lo cual nos indica que explica mejor la variabilidad de los datos el modelo con las variables bosque fusionadas, por tal motivo, la explicación del modelo de aquí en adelante se hace con base a los resultados del caso *b*.

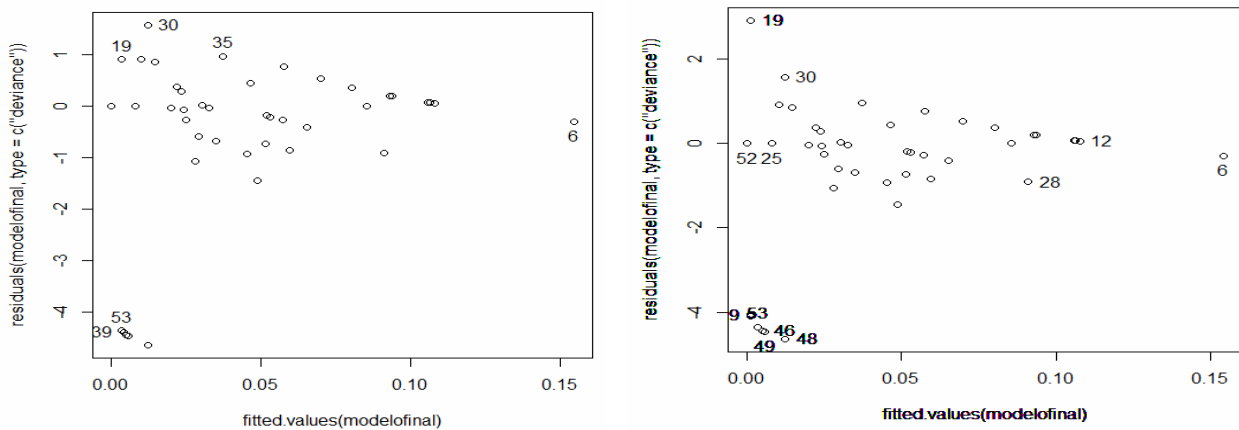


Figura 4.5 Análisis de residuos por el método de “deviance” para los datos con los siete usos de suelo (izquierda) y con los usos de suelo fusionados (derecha). Se observa en ambos casos que el dato 6 es el que un efecto de incremento de la función; recordando que es una función inversa por ser de la familia de la gamma con función vinculo inversa.

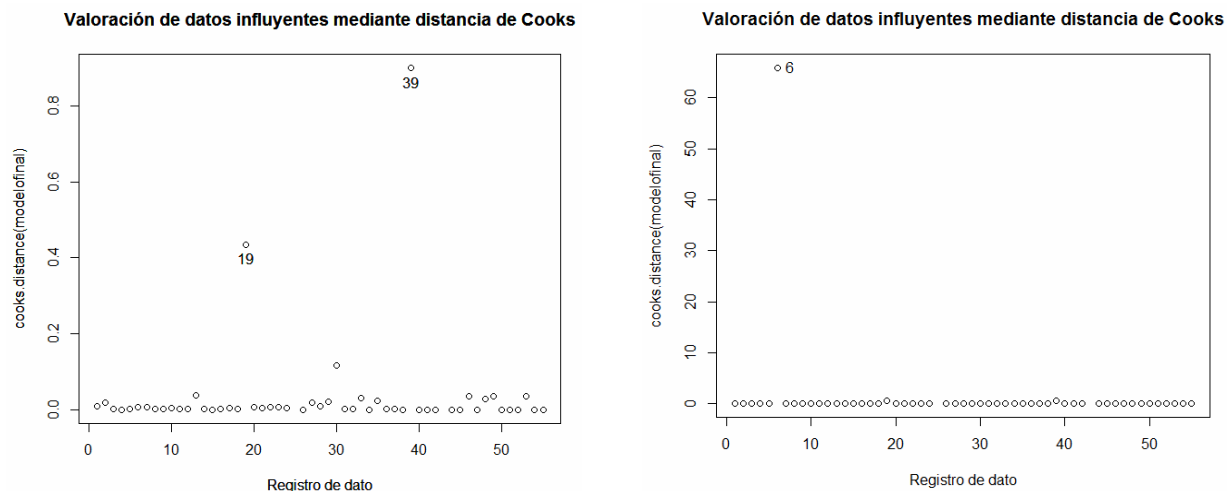


Figura 4.6. Análisis gráfico del grado de influencia mediante las distancias de Cook, para los avistamientos de las nutrias en la cuenca Apatlaco- Tembembe. Izquierda caso *a*, derecha caso *b*.

anova(modelofinal) caso *a*
Analysis of Deviance Table

Model: Gamma, link: inverse

Response: (datlutre\$Den + 1e-07)

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid.
Dev				
NULL			54	435.04
datlutre\$Distancia	1	77.659	53	357.38
as.factor(datlutre\$Suelo)	6	243.399	47	113.98

anova(modelofinal) caso *b*
Analysis of Deviance Table

Model: Gamma, link: inverse

Response: (datlutre\$Den + 1e-07)

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid.
Dev				
NULL			54	435.04
datlutre\$Distancia	1	77.659	53	357.38
as.factor(datlutre\$Suelo)	5	155.573	48	201.81

Cuadro 4.2 análisis de devianza para los dos casos del modelo lineal generalizado, caso *a* con siete tipos de uso de suelo, caso *b* con usos de suelo de bosque fusionados.

La interpretación de los resultados del análisis del modelo lineal generalizado, nos indica que los tipos de suelo: Bosque de encino y área sin vegetación, son significativos, cuadro 4.3; lo que expresan estos valores es grado de influencia para la determinación de la presencia de nutrias en la zona, en éste caso la presencia de estos dos tipos de suelo hace que la densidad sea mayor en los tipos de suelo como son el de manejo agrícola y la selva caducifolia, ya que en los registros originales de avistamientos los suelos de Bosque de encino y área sin vegetación no reportan registros de avistamientos asociados en la cuenca Apatlaco- Tembembe, lo cual nos indica que son un factor limitante para la ubicación en zonas aledañas.

Coefficiente	Valor	Error estándar	Valor de t	P-value
Intercepto (β_0)	1.542e+01	7.394e+00	2.085	0.0424 *
Distancia (β_1)	3.297e+01	1.325e+01	2.488	0.0164 *
Bosque de encino (β_2)	8.359e+02	4.103e+02	2.103	0.0472 *
Área sin vegetación (β_3)	1.000e+07	4.755e+06	2.103	0.0407 *
Selva caducifolia (β_4)	-9.508e+00	7.511e+00	-1.266	0.2116
Laguna (β_5)	1.474 e+01	1.517e+02	0.097	0.9230
Bosque mesofilo de montaña (β_6)	1.000 e+07	1.165e+07	0.858	0.3949

Cuadro 4.3. Valores de los coeficientes de regresión para el modelo de densidad de nutrias considerando distancia al cauce como (β_1), bosque de encino (β_2) y área sin vegetación (β_3), que fueron los únicos tipos de suelo en el área de estudio que se reportan como factores significativos 0.01 ‘*’.

Dado que no presentan ninguna significancia las variables área sin vegetación, Selva caducifolia, laguna, y bosque mesofilo de montaña, no se deben de considerar dentro del modelo final que nos permite calcular la densidad de nutrias; quedando el modelo con todas las variables.

$$\ln(\lambda) = \frac{1}{15.42 + 32.97x_1 + 835.9x_2 + 10000000x_3 - 9.508x_4 + 14.74x_5 + 10000000x_6}$$

donde λ es la densidad de nutrias, x_1 = Distancia, x_2 = Bosque de encino y x_3 = área sin vegetación, x_4 = Selva caducifolia, x_5 = Laguna, x_6 =Bosque mesófilo de montaña; de estos regresores los significativos son de el intercepto y x_1 , x_2 y x_3 , los otros no son significativos.

Sin embargo, al hacer la transformación de manera exponencial y considerando que es el inverso de la función, dado que se utilizo un MLG con función vinculo inversa, entonces los valores de las betas que eran significativas en el modelo, como son las de área sin vegetación y bosque de encino, pierden sentido ya que se hacen cero, y por otro lado el valor que representa a la selva caducifolia se incrementa adquiriendo sentido.

$$\lambda = \exp\{-15.42 - 32.97x_1 - 835.9x_2 + 9.508x_4 - 14.74x_5\}$$

donde λ es la densidad de nutrias, -15.45 es el intercepto, -32.97 el efecto de la distancia a los cuerpos de agua, -835.9 el efecto del bosque de encino, 9.508 el efecto de el tipo de suelo de selva caducifolia y -14.74 es el efecto de la laguna . De lo anterior se puede interpretar como que, si la distancia es mínima a los cuerpos de agua hay un mayor efecto para la densidad de las nutrias, figura 4.7. Por otro lado, dos valores que se consideran significativos en el modelo logarítmico al hacer la transformación al exponencial se hacen muy pequeños, uno de ellos prácticamente se hace cero (área sin vegetación), sin embargo, el tipo de suelo selva caducifolia, que no resulta significativo nos indica que su efecto se ve incrementado en medida en que el área que representa ese tipo de suelo se hace mayor, esto se ve favorecido por el carácter restrictivo del tipo de suelo bosque de encino y área sin vegetación, ya que su presencia hace cero la probabilidad de avistamiento dentro de cada una.

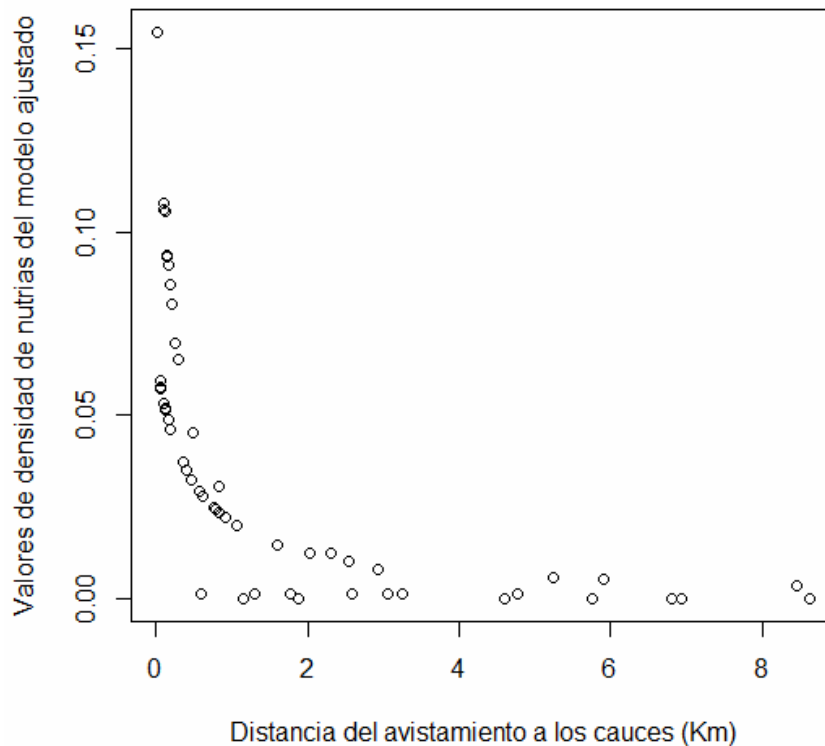


Figura 4.7. Gráfica de la proyección de densidad de nutrias a partir del modelo de regresión lineal generalizado. Se observa que a menor distancia hay una mayor densidad de nutrias, es decir entre más cerca de los cauces hay mayor densidad de nutrias, entre más se aleje de los cauces hay menor densidad de nutrias. Los datos que se ven en la parte inferior y que tienen una densidad de cero se deben a que se toma como valor de referencia a los centroides de los polígonos donde no hubo registro de nutrias.

Finalmente la figura 4.8, muestra los valores calculados para la densidad a partir del modelo de regresión obtenido, mostrando que la densidad es variable y que las densidades son bajas para prácticamente toda el área de estudio. Las zonas que se marcan con mayor densidad coinciden con las que se consideraron en el análisis de la densidad con la función Kernel.

Proyección colorimétrica y de niveles de la estimación de densidad
por el modelo lineal generalizado

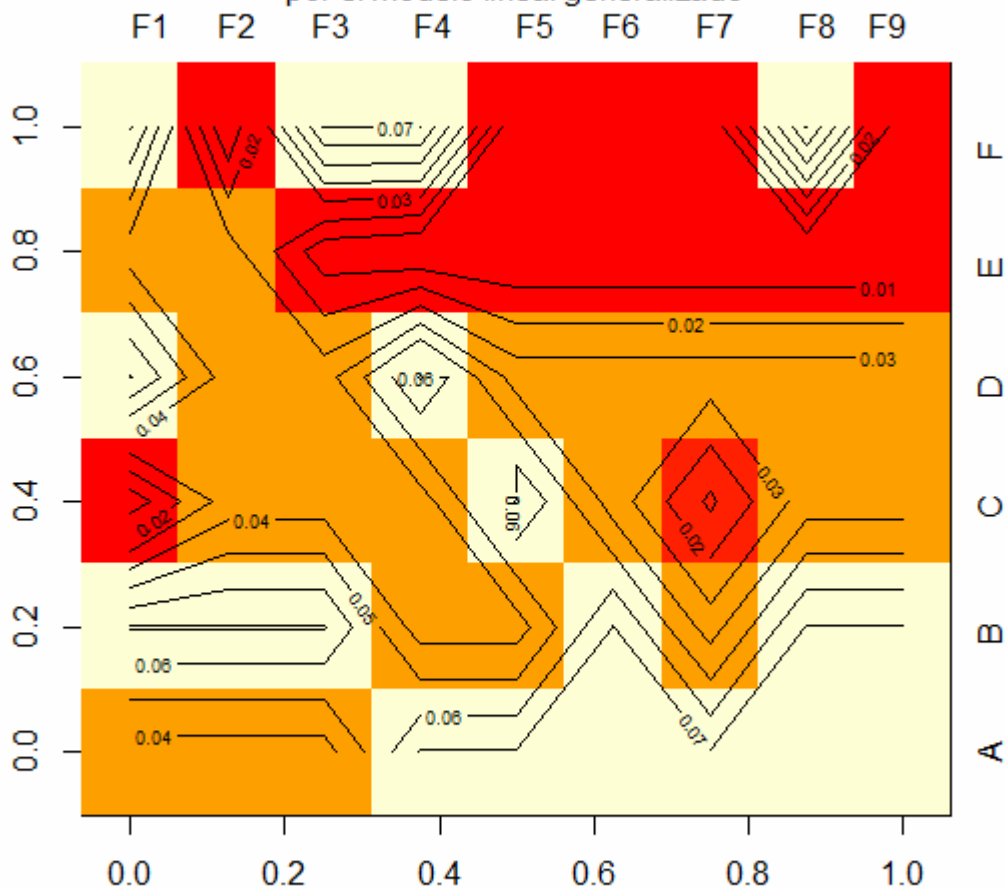


Figura 4.8. Gráfica de los valores ajustados, calculados a partir del modelo de regresión. Obsérvese que los valores de densidad son variables. Los valores van de 0.01 a 0.07 lo que nos indica que el encontrar nutrias bajo las condiciones del modelo es baja. Las zonas de mayor densidad son las que se encuentran en color claro, siendo las zonas de mayor densidad las de color crema, con valores de 0.06 y 0.07.

Capítulo 5: DISCUSIÓN Y CONCLUSIÓN

Los datos de la muestra de avistamientos de nutrias en la cuenca Apatlaco-Tembembe se obtuvieron de forma retrospectiva a partir de otros trabajos, que van desde 1986 hasta 2004 y se reportaron en el trabajo de Cirrelli Villanova, 2005. Por lo que el análisis que se realizó en este trabajo permite plantear una visión general del comportamiento de las nutrias respecto a su distribución espacial dentro la cuenca en diferentes momentos.

Haciendo un análisis de los resultados obtenidos en las figuras 3.1 y 3.2 así como el cuadro 3.1 se observa que la distribución no es uniforme, esto se hace evidente al observar que la ubicación de las nutrias teniendo una mayor presencia en los cauces 71 y 9 (10 y 9 avistamientos, respectivamente). Esto a su vez se ve respaldado por la gráfica de la función K de Ripley, 4.1, la cual indica que los datos de avistamiento no se distribuyen al azar, sino de manera agregada, esto no quiere decir necesariamente que los organismos habitan en grupos, sino que hay condiciones en la zona que favorece su presencia, más aun considerando que los registros de los avistamientos son atemporales, pero si se podría pensar que la asociación tiene más que ver con las características de la zona en la que se ubican los organismos, lo cuál se ve respaldado con el modelo lineal generalizado, que muestra gran significancia por la distancia a los cauces y los diferentes tipos de suelo.

La densidad mayor de avistamientos de nutrias, se encuentra en el cuadrante B13, en el cual también se encuentra el cauce número 71; esto también se respalda por los resultados de la función Kernel, que muestra que la zona con mayor intensidad (de color más claro, figura 4.2) y se ve caracterizado por un suelo de selva caducifolia.

La segunda zona con mayor densidad es el cuadrante A5, figura 3.5, que también se ve en la imagen de la función Kernel, figura 4.2 parte clara abajo a la derecha, pero aquí el suelo mayormente es de tipo manejo agropecuario y secundariamente selva caducifolia.

Particularmente el cauce 71 se encuentra ubicado en un tipo de suelo de selva caducifolia y de manera adjunta se encuentran tipos de suelo agrícola, bosque de encino, bosque de coníferas y bosque diferente de coníferas; en cuanto al cauce 9 se encuentra en la parte central, entre los cauces 14 y 2, hay mayor presencia de tipo de suelo de manejo agrícola y del lado derecho del cauce 2 sobre el cauce 9 hay mayor presencia de selva caducifolia, y en la parte sur del cauce 9 se encuentran zonas de tipo de suelo de bosque de encino y áreas sin vegetación, figura 3.2.

La distancia a los cauces de la mayoría de los avistamientos de nutria (65%) no es mayor a los 500 metros, figura 3.4. El cauce 71 que tiene el mayor número de avistamientos asociados y se encuentra dentro de esta distancia a los avistamientos, cuadro 3.1.

Por su parte el modelo lineal generalizado muestra que la distancia a los cauces y los tipos de suelo bosque de encino y área sin vegetación son significativos para la determinación de la densidad, cuadro 4.1, sin embargo, estos suelos se asocian con una densidad cero en el avistamiento de nutrias, de acuerdo al registro observado, cuadro 3.2; lo cual nos haría pensar que cuando hay este tipo de suelo la probabilidad de encontrar nutrias es nula, así que se asocia la presencia de este tipo de suelos como un factor limitante, es decir que, la presencia de estos tipos de suelo induce que las nutrias se concentren en suelos de otro tipo, por ejemplo suelos tipo selva caducifolia.

La gráfica de los valores ajustados a partir del modelo lineal generalizado, figura 4.8, muestra que la zona que se esperaría con mayor densidad de nutrias corresponde a los cuadrantes A5, A6, A7, B10, B15, C23, D31, F48 y F49 que si bien en los datos individuales reportados en el cuadro 3.2 los asociados a mayores densidades son los cuadros B13, B14, B15, A5 y A6 y que no todos coinciden con los observados, si hay tres de los cinco dentro de los esperados; aunque es

de resaltar que el de mayor densidad observada. B13, no aparece dentro de los esperados; por otro lado los cuadrantes aledaños y ellos mismos tienen asociados suelos de tipo de manejo agrícola y selva caducifolia; sin embargo estos dos tipos de suelo, no son significativos para el modelo, lo que podría estar pasando es que más bien los tipos de suelo bosque de encino y área sin vegetación son factores restrictivos en la distribución espacial, es decir favorecen la presencia de las nutrias en otros tipos de suelo. Lo que es congruente con el análisis presentado en el párrafo anterior.

De manera general los avistamientos se dieron mayormente en el tipo de suelo de selva caducifolia y de manejo agrícola. Sin embargo, los registros en el suelo de tipo agrícola pueden haberse registrado en mayor parte debido a que la presencia de humanos es más frecuente, ya que, pasan más tiempo ahí por cuestiones de trabajo, así que se puede decir que el esfuerzo fue mayor en el suelo de tipo de manejo agrícola.

Para realizar un muestreo sobre la preferencia de hábitat de la nutria de río, partiendo del análisis presentado, se debe de considerar que: la distribución espacial de las nutrias de río, la cual no es homogénea; por lo tanto en un muestreo se tendría que programar considerando estratos, los cuales deberán estar en razón del inverso de la densidad de avistamiento, si se quiere conservar la proporción en el esfuerzo. Por otro lado es conveniente considerar el tipo de suelo para hacer la planeación del muestreo, así que considerando que donde es suelo tipo bosque de encino o/y área sin vegetación la esperanza de encontrar nutrias es cero, estas zonas se deben de eliminar de los estratos, sin embargo, no se debe de perder la asociación de su presencia ya que es un factor limitante para que las nutrias se concentren en las zonas de selva caducifolia. He aquí la relevancia de la aplicación de este tipo de estudio, ya que permite la ubicación y asociación de diferentes variables considerando su dependencia y efecto entre ellas.

Por otra parte pese a que en la gráfica de los valores ajustados se esperaría encontrar una alta densidad de nutrias en el suelo de tipo manejo agrícola, este tipo de suelo es conveniente eliminarlo también de los posibles estratos, esto debido a que quizá los reportes obtenidos en esta zona, más bien responden a que hay mayor tiempo de la presencia de personas, favoreciendo con ello la probabilidad de avistamiento. Así que la recomendación es, darle mayor prioridad al tipo de suelo de selva caducifolia, además de considerar en todo momento la proximidad con los dos tipos de suelo: bosque de encino y área sin vegetación, de los cuales su significancia radica en ser factor limitante del área de selva caducifolia, sin los cuales posiblemente la distribución espacial de las nutrias sería no agregada; asimismo establecer una distancia de 500m a los cauces; ya que se pudo apreciar que a menor distancia de los cauces es mayor la densidad de nutrias encontradas, manteniéndose todas las demás condiciones constantes; de hecho se pueden utilizar los cauces como guías paralelas para la determinación de los transectos; además de hacer un esfuerzo de forma inversa a la densidad de nutrias avistadas para mantener la proporcionalidad dentro de los estratos.

Estas recomendaciones ponen en claro que el muestreo conociendo el como se distribuyen espacialmente los organismos, además de conocer algunos de los aspectos que contribuyen al establecimiento de la predilección del hábitat por los organismos, nos puede redituarse en una reducción de: esfuerzo, gasto de recursos, ya que no desperdiciaremos dinero, personal ni tiempo en zonas en las cuales no hay probabilidad de encontrar ningún organismo objetivo; además de minimizar los errores de subestimación o sobre estimación por desconocimiento de los efectos de asociación ecofisiológicas de las zonas de estudio.

En concreto un estudio de distribución espacial permite resaltar características relevantes para la planeación de un muestreo en un organismo vágil, bajo las condiciones de un muestreo estadístico, ya que al saber que la densidad no es homogénea, nos lleva a plantear una

estratificación de la zona la cual responderá al diferente grado de densidad de la zona, así se pueden tener los diferentes estratos y seleccionar sólo algunos y dentro de cada uno y escoger algunas partes de cada uno; además se puede planear el tiempo, número de personas y recursos necesarios para poder hacer proporcional el esfuerzo en los estratos respecto a la densidad esperada por estrato. Otro aspecto es el que se puede hacer la disposición de los transectos dentro de cada estrato considerando la distancia óptima a los cauces a la que se pueden instalar cada uno. Así el análisis de los datos puede ser confiable, permitiendo incluso hacer predicciones.

GLOSARIO

	Definición	Referencia
Sésil	Organismo que se mantiene en una posición, normalmente fijo a un sustrato	Parker, 1984.
Vágil	Organismo que se desplaza libremente.	Parker, 1984.
Merístico	Referente a, o dividido en segmentos, dicese de la determinación de una porción o parte de un organismo, anatómicamente o bien de las dimensiones de la parte.	Parker, 1984.
Efecto de Borde	El resultado de la interacción de ecosistemas adyacentes	Krebs, 1989
Cudrante	Área que está definida por cualquier forma geométrica, que se utiliza para un muestreo	Krebs, 1989
Transecto	Refiere a una línea de una extensión determinada sobre la cual se escogen puntos como referencia para realizar registros.	Krebs, 1989
Gregario	Se aplica al animal que vive en rebaño o en grupo	the free dictionary
Territorialidad	Conducta que hace referencia a cualquier área sociográfica, donde un animal de una particular especie, consistentemente defiende contra conspecíficos y, ocasionalmente, contra individuos de otras especies.	Begon, 1997
Generalistas	Refiere a aquellos organismos que pueden comer una amplia variedad de alimento.	Begon 1997
Competencia intraespecífica	Aquella que se produce entre organismos de la misma especie	Begon, 1997
Teodolito	Es un instrumento de medición mecánico-óptico que se utiliza para obtener ángulos verticales y, en el mayor de los casos, horizontales, ámbito en el cual tiene una precisión elevada. Con otras herramientas auxiliares puede medir distancias y desniveles.	http://es.wikipedia.org/wiki/Teodolito
Buffer	polígono que encierra el área de influencia resultante de dar una determinada distancia en torno a un punto	http://es.wikipedia.org/wiki/Buffer

ANEXO

Rutina vecino más cercano

#Para la rutina del vecino más cercano se cargaron los 38 avistamientos individualmente.

```
av1<-(as.matrix(read.table("C:\\MJHL\\Tesina\\Muestreo nutrias\\CaucesConv\\Av1.txt")))
```

```
.
```

```
.
```

```
.
```

```
av38<-(as.matrix(read.table("C:\\MJHL\\Tesina\\Muestreo nutrias\\CaucesConv\\Av38.txt")))
```

#Se cargaron los cauces

```
c1km<-(as.matrix(read.table("C:\\MJHL\\Tesina\\Muestreo nutrias\\CaucesConv\\c1km.txt")))
```

```
.
```

```
.
```

```
.
```

```
c71km<-(as.matrix(read.table("C:\\MJHL\\Tesina\\Muestreo nutrias\\CaucesConv\\c71km.txt")))
```

Se cargo la librería de splancs

```
library(splancs)
```

Se ordeno y calculo la distancia menor de cada avistamiento respecto a cada cauce

```
sort(spDistsN1(c2km,av1, longlat = FALSE))
```

```
...
```

Se realizó la distribución de frecuencias de las distancias más cercanas para los 38 avistamientos.

Calculo de centroides

Se tomaron los puntos medios de cada lado de los polígonos de 10 X10, considerando los puntos de los vértices.

CONTRASTE PARA MODELO DE REGRESIÓN

#Realización del modelo, el cual considera a la variable tipo de suelo como un factor para que pueda correr la rutina

```
glm(formula = (datlutre$Den + 1e-07) ~ datlutre$Distancia + as.factor(datlutre$Suelo),
```

#Una vez realizado el modelo se puede pedir ver el contraste.

```
contrasts(as.factor(datlutre$Suelo))
```

NUTRIA DE RÍO NEOTROPICAL

Son de color castaño claro a oscuro con una coloración más clara que va desde la parte baja de la nariz, boca y se expande hacia el cuello y pecho. Alcanza una talla de hasta 1.2 metros desde la punta de la nariz a la punta de la cola, y puede pesar hasta 12 kg, es usual que el macho sea más pesado que la hembra. Entre los dedos de las patas presenta membrana, que le facilita el nado. Se alientan principalmente de pescado, aunque incluyen también ranas, cangrejos, algunas aves y pequeños mamíferos. Usualmente tienen entre 1 y 4 crías, aunque lega a tener hasta 6 por camada, su periodo de gestación es entre 60 y 70 días; aunque en algunos casos pueden tardar hasta 12 meses, cuando presentan implantación retardada. Tienen cuidados parentales. Se distribuye desde Sonora en México, por la franja del Pacífico hasta zona sur del Perú, parte de la meseta central de México hasta llegar a la franja de las costas de Tamaulipas y de ahí a lo largo del Golfo de México atravesando Yucatán, Quintana Roo y luego hacia el sur del continente pasando por Colombia, Venezuela, Brasil, Paraguay, Uruguay y Argentina (Kruuk, 2006; MacDonald, 2006)



Phyllum: Vertebrata
Clase: Mammalia
Orden: Carnívora
Familia: Mustelidae
Subfamiia: Lutrinae
Genero: Lontra
Especie: Lontra longicaudis annectens
Nombre común: nutría de río.
Otros nombres: Tunsla (zona maya), perro de agua

<http://www.turismo-costa-rica.com/nutria.htm>



http://jesusenlared.blogspot.mx/2012_01_01_archive.html



Distribución de la nutria neotropical
http://es.wikipedia.org/wiki/Lontra_longicaudis



<http://yeimerarenas.blogspot.mx/2011/10/nutria-de-rio-fichatecnica-nombre.html>

BIBLIOGRAFIA.

- Azorín-Poch, F. (1972). *Curso de muestreo y aplicaciones*. Aguilar España. 4 – 9, 21 – 22.
- Azorín-Poch, F. y Sánchez-Crespo, J.L. (1986). *Métodos aplicados del muestreo*. Alianza. Madrid. 263 – 273.
- Baena, M.L.; Escobar, F. y Halffter, G. (2009). *Impacto de Omorgus suberosus (Fabricus) en nidos de la tortuga marina Lepidochelys olivacea: Efecto, distribución espacial y la densidad*. Memorias VIII reunión Latinoamericana de Escarabeidología, Xalapa, Ver. México, p 12.
- Bautista-Zúñiga, F.; Delfín-González, H; Palacio-Prieto, J. L. y Delgado-carranza. M. C. (2004). *Técnicas de muestreo para manejadores de recursos naturales*. UNAM, UAY, CONACYT, INE. México. 279 – 302
- Begon, M; Harper, J. & Townsend, C.R. (1997). *Ecología*. 2° reimpresión. Omega. Barcelona. 101, 206 – 221, 241 – 244.
- Buckland, S.T. (1985). Perpendicular distance model for line transect sampling. *Biometrics*. 41(1):29 – 42.
- Buckland, S.T., Anderson, D.R., Burnham, K.P. and Laake, J.L. (1993) *Distance sampling*. Chapman & Hall. UK. 104 – 172.
- Camarero, J.J. y Rozas, V. (2006). Técnicas de análisis espacial de patrones superficiales y detección de fronteras aplicadas en ecología forestal. *Invest Agrar: Sist Recurs For*. 15(1):66 – 87.
- Cayueta, L. (2010) Modelos lineales generalizados(GLM) Universidad de granada. <http://158.49.96.73:8080/documenta/bitstream/00000001/24/1/3-Modelos%20lineales%20generalizados.pdf>
- Cirelli Villanova, V. (2005). *Restauración ecológica en la cuenca Apatlaco-Tembembe. Estudio de caso: Modelado de la distribución de la nutria de río, Lontra longicaudis annectens*. Tesis de maestría en ciencias biológicas; orientación en reestructuración ecológica. Instituto de Biología. Universidad Nacional Autónoma de México.
- Clapham, A.R. (1932). The form of the observational unit in quantitative ecology. *The Journal of Ecology*. 20(19): 192 – 197.
- Diggle, P.J. (1983). *Statistical análisis of spatial point patterns*. Academic press. N.Y. 1-89
- Dixon, P.M. (2002). *Ripley's K function*. Encyclopedia of Environmetrics. Vol 3. 1796 – 1803. Anel H; & Piegorsch W.W. editors. Jonh Wiley & sons.
- Diccionario: <http://es.thefreedictionary.com/gregario>
- Giraldo-Henao, R. (2011). *Notas de clase*. Universidad Nacional de Colombia. <http://www.docentes.unal.edu.co/rgiraldoh/docs/NotasClase.pdf>
- Hayes, R.J. and Buckland. (1983). Radial-Distance model for the line-Transect method. *Biometrics*. 39(1):29 – 42.
- Hiby, L. and Krishna, M.B. (2001). Line Transect sampling from a curving path. *Biometrics*. 57(3): 727 – 731.
- Krebs, (1989). *Ecological methodology*. 64 – 124.
- Kruuk, H. (2006). Otters: ecology, behaviour and conservation. Oxford, U.K. 15 – 17 .
- Lohr, S.L. (2000). *Muestreo diseño y análisis*. Thomson. México. 480
- MacDonald, D. (2006). La gran enciclopedia de los mamíferos. Diana. Madris, Es. 136 – 141.
- Miñarro, A. (1998). *Estimación no paramétrica de la función de densidad*. 33-38. <http://www.ub.edu/stat/personal/minarro/documents/Nonpar.pdf>
- Murguía, M. (2005). *La estadística especial como herramienta de análisis de la biodiversidad. Sobre diversidad biológica: El significado de las diversidades α , β , γ* . Editores Halffter, G;

- Soberón, J; Koleff, P. y Meliá, A. Ed. S.E.A., CONABIO, CONACYT, UNIVERSITAS. En M3M. Vol 4. 63 – 72.
- Parker, P. S. (1984). *Diccionario de biología*. Tomo I, A-M. McGraw-Hill. Parker, P.S editor.
- Rozas, V. y Camarero, J.J. (2005). Técnicas de análisis espacial de patrones de puntos aplicadas en ecología forestal. *Invest Agrar: Sist Recurs For*. 14(1): 79 – 97.
- Seber, G.A. F. (1982). *The estimation of animal abundance*. 2nd. Charles Griffin & Co. Eng. 19 – 35.
- Sen, A.R; Touring, J. and Smith, G.E. (1974). On line transect sampling method. *Biometrics*. 30(2): 329 – 340.
- Servy, E; Cuesta, C. y Marí, G. (2005). *Estimación de funciones de densidad de datos correlacionados y/o con valores atípicos*. Décimas jornadas “Investigaciones en la Facultad” de ciencias económicas y estadísticas. Universidad Nacional de Rosario. Argentina. <http://www.fcecon.unr.edu.ar/fcecon.unr.edu.ar/sites/default/files/u16/Decimocuertas/Servy,%20mari,%20cuesta%20estimacion%20de%20funciones.pdf>
- Skellam, J.G. 1958. The mathematical foundations underlying the use of line transects in animal ecology. *Biometrics*. 3(14): 358 – 400. (sept)
- Smith, G.E.J. (1979) Some aspects of line transect sampling when the target populations moves. *Biométries*. 35(March): 323 – 239.
- Zar, J.H. (1999). *Bioestadistical análisis*. 4th ed. Prentice Hall. NJ. 575 – 578.