

CAPÍTULO 2- PROCEDIMIENTOS PARA EL ANÁLISIS DEL COMPORTAMIENTO ESPECTRAL DE LA COBERTURA VEGETAL

El procedimiento metodológico para determinar el comportamiento espectral de la cobertura vegetal aprovechando imágenes del satélite MODIS se ha concebido en correspondencia con los objetivos de la tesis y como resultado del análisis del estado de la temática en el ámbito internacional y en Cuba.

El contenido principal de la metodología, es la secuencia u orden de pasos en que se han dispuesto los diferentes procedimientos de trabajo para las imágenes satelitales y el mapa de cobertura vegetal, que considera las dependencias tecnológicas y organizativas entre ellos, para alcanzar los resultados (Ver figura 2.1) que se explica a continuación.

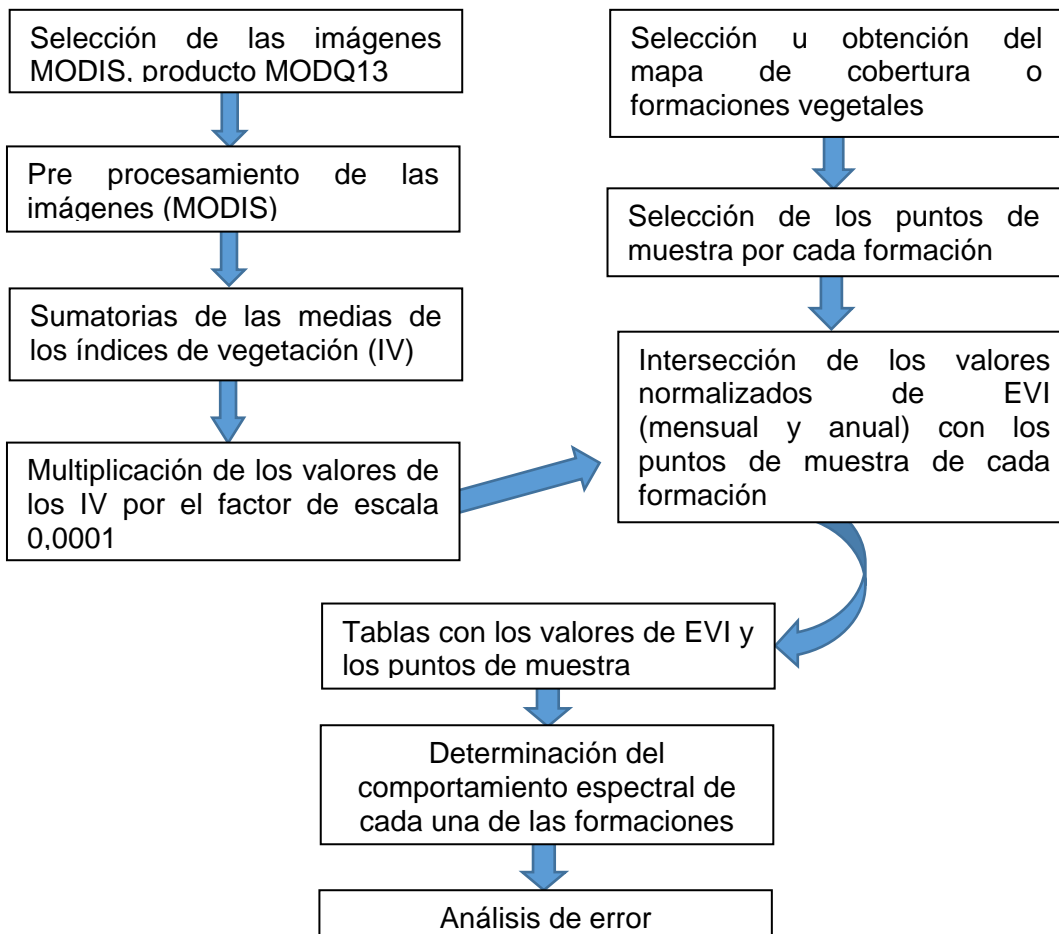


Figura 2.1 Esquema metodológico elaborado por el autor.

2.1- Selección de las imágenes MODIS, producto MODQ13

La selección de las imágenes se debe realizar, teniendo en cuenta los tipos de resoluciones de las imágenes satelitales que a continuación se mencionan:

1. Resolución espacial.
2. Resolución espectral.
3. Resolución temporal.

Resolución Espacial: Se refiere al objeto más pequeño que puede ser distinguido en una imagen (Chuvieco, 1995). En los sensores de barrido, se refiere al tamaño de la mínima unidad de información incluida en la imagen, y que se denomina pixel. Esta resolución depende fundamentalmente de la altura y la velocidad a la que órbita el satélite y del número de detectores. Los satélites dedicados a la observación terrestre portan sensores con una resolución espacial muy variable que va, desde los 60 cm del satélite Quickbird o 1 m en las imágenes Ikonos, hasta los 5 km de algunos satélites meteorológicos geostacionarios. Los valores más frecuentes están comprendidos entre los 15 y los 250 m (Aster, Landsat, Spot, MODIS).

Resolución espectral: Es la capacidad que tiene el sensor para separar diferentes longitudes de onda dependiendo del dispositivo de filtrado que separa la radiación recibida en bandas espectrales más o menos angostas (Chuvieco, 1995). La resolución espectral nos indica el número y anchura de bandas espectrales en las que un sensor registra la radiación; por ejemplo, el satélite Aster registra 14 bandas espectrales, Landsat 8, capta 11 bandas, el sensor HRS de Spot capta 5 bandas y el MODIS 36 bandas.

Resolución temporal: Es la periodicidad con la que el sensor adquiere información de una misma zona (Chuvieco, 1995). Varía en función de las características orbitales, del ángulo de observación y de la cobertura del sensor. Los satélites geostacionarios meteorológicos proporcionan datos cada 30 minutos aunque con una resolución espacial pobre. El sensor Aster pasa cada 16 días, Landsat repite imágenes cada dos semanas, Spot pasa por el mismo lugar aproximadamente entre 26 días y MODIS dos veces al día.

Teniendo en cuenta estas resoluciones de las imágenes satelitales, se determina el tipo de sensor e imagen a procesar. Para este tipo de trabajo se recomienda utilizar como imágenes de partida, en primer lugar aquellas que tengan alta resolución temporal, alta resolución espectral y alta resolución espacial. La selección se hace difícil debido a que la resolución temporal no va relacionada con la resolución espacial, por lo que se selecciona la imagen que responda mejor a nuestro objetivo, si el objetivo es analizar una cobertura X en el tiempo, lo conveniente es utilizar un sensor de alta resolución temporal todo lo contrario si el análisis se hace para el espacio.

Para nuestro estudio, se escogió imágenes de alta resolución temporal, como la del sensor MODIS, que permite un estudio detallado de la cobertura vegetal, en un gran período de tiempo y con gran frecuencia de datos para analizar.

2.2- Pre procesamiento de imágenes MODIS

Los índices de vegetación calculados a partir de la información espectral contenida en las imágenes MODIS están diseñados para ofrecer comparaciones espaciales y temporales de las condiciones de la vegetación. La utilización de este tipo de herramientas puede ofrecer, sin embargo, una serie de problemas técnicos en su procesamiento y manejo. En primer lugar las imágenes vienen en un formato especial (*.hdf) que no leen la mayoría de los SIG y software de teledetección. En segundo lugar, la información contenida en las distintas bandas de las imágenes MODIS viene en un tipo de proyección (sinusoidal) bastante atípico, por lo que suele ser necesaria la reproyección de esta información.

Lo anterior se resuelve en la etapa de pre-procesamiento donde se hace necesario llevar la información a un formato Geotiff (entendible por la mayoría de los SIG) con imágenes separadas de cada banda de NDVI, EVI y re-proyectar las imágenes a un sistema de coordenadas estándar para cada país que sea compatible con un software de GIS y de procesamiento de imágenes.

En el caso de imágenes MODIS para Cuba se debe considerar la proyección WGS-84 y la conformación de un mosaico de 3 “tiles” o cuadros que cubran todo el territorio nacional (Figura 2.2).

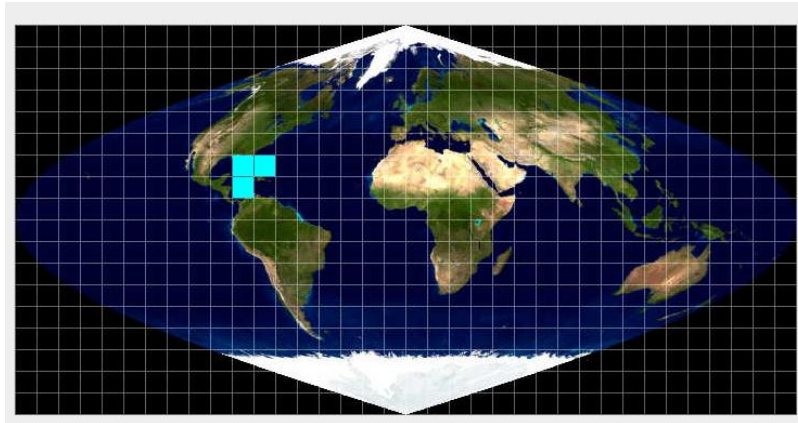


Figura 2.2- Cuadrículas con el cubrimiento del territorio de Cuba para imágenes MODIS.
Fuente: Martín (2012).

Para el pre-procesamiento de las Imágenes a un formato compatible con un software GIS, se hace necesario obtener los programas de instalación para el pre-procesamientos de los datos, los cuales se encuentran disponibles en la página del Servicio Geológico de los Estados Unidos de América (USGS) (<http://LDPAAC.usgs.gov>) de acuerdo al sistema operativo en que se desee trabajar.

Estos programas son:

- MODIS Reprojection Tool (MRT)
- MODIS Land Data Operation Product Evaluation (LDOPE)
- MODIS Swath Reprojection Tool (MRT Swath)
- MODIS Data Pool Extraction Tool (MODextract)

Estas herramientas pueden ser instaladas bajo Windows 2000, XP, Linux, SGI IRIX 6.5 y Solaris 2.7.

Cuando se descargan los datos MODIS estos están almacenados en un formato comprimido con extensión HDF. El siguiente paso es definir qué información almacenada en cada archivo hdf es útil para nuestro estudio, por lo que antes de comenzar el pre-procesamiento se deben tener en los siguientes requerimientos:

- Índices de vegetación (EVI y NDVI)
- Resolución temporal suficiente para entender el progreso de las respuestas de las formaciones vegetales.
- Resolución espacial: 250 metros
- Período: Año 2004
- Proyección: WGS-84
- Formato de salida: Geotiff

Para realizar el pre-procesamiento (mosaico y registro de las escenas), recomendamos emplear el programa “MODIS Reprojection Tool MTR”, el mismo permite re proyectar datos de niveles 2G, 3 y 4, en las proyecciones geográficas de: *Albers Equal Area*, *Geographic*, *Hammer*, *Integerized Sinusoidal*, *Interrupted Goode Homolosine*, *Lambert Azimuthal*, *Lambert Conformal Conic*, *Mollweide*, *Polar Stereographic*, *Sinusoidal*, *Transverse Mercator*, *Universal Transverse Mercator*; teniendo como referencia los Datums de: NAD27, NAD83, WGS66, WGS72 y WGS84.

Una vez instalado el programa “MODIS Reprojection Tool” este se abrirá y se le introducirán los datos correspondientes a la ubicación de las escenas (H10-V6; H10-V7 y H11-V6), la proyección en las que se quiere obtener (WGS-84), la fecha (día, mes y año) de la imagen que se requieren (síntesis cada 16 días), las escenas del MODQ13 (EVI), el tipo de remuestreo que se va a realizar (vecino más cercano) y por último el lugar en que se quiere guardar con extensión Geotiff. (Ver figura 2.3).

Una vez concluida la operación, de realizar el mosaico de las escenas, se procedió a recortar el mismo con ayuda de un vector que enmarcara al territorio cubano. Como resultado de este trabajo se obtuvo un total de 288 imágenes de Cuba por meses (desde 2000-2012).

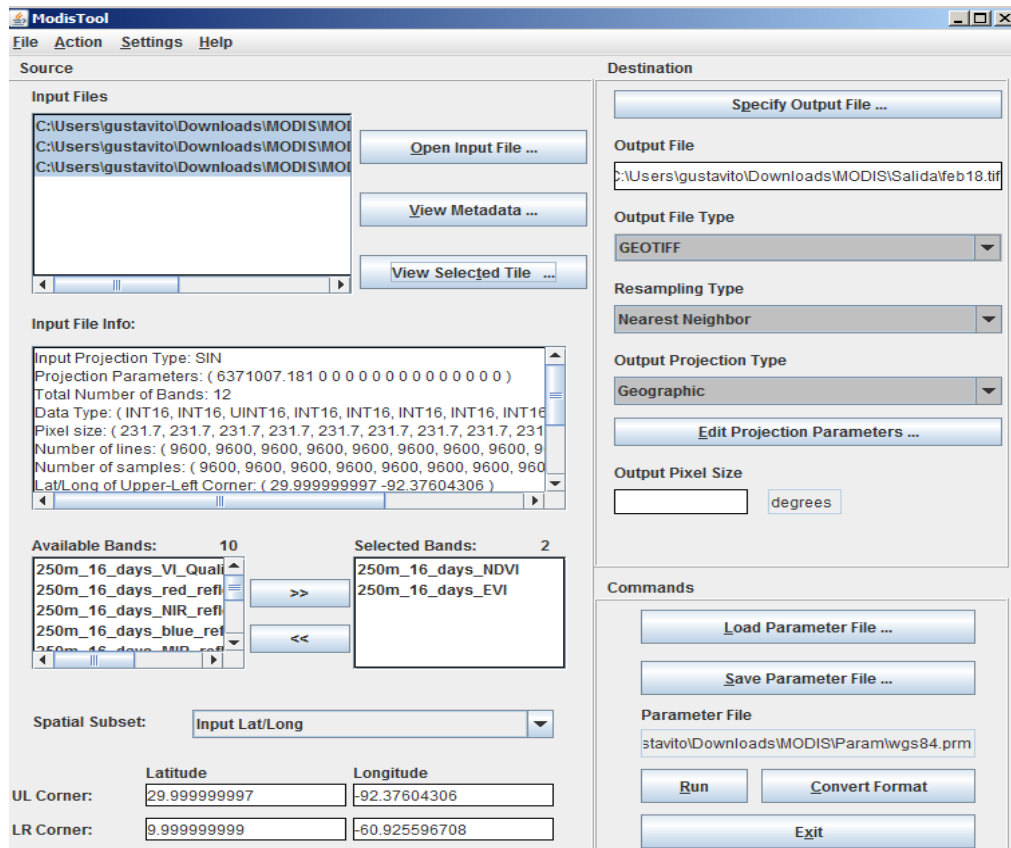


Figura 2.3- Imagen en display del programa MRT.
Fuente: Elaborado por el autor.

2.3- Sumatorias de las medias y multiplicación de los valores por el factor de escala

El producto MODQ13 de MODIS presenta dos índices de vegetación (IV) los cuales presentan una síntesis de 16 días, esta presenta los mejores valores de las 64 observaciones realizadas por el sensor. Después de tener estas síntesis se debe realiza una sumatoria de los dos valores medios que nos proporciona el satélite para obtener la media de cada mes. Esta acción se realiza empleando la herramienta álgebra de mapas de cualquier Sistema de Información Geográfica (SIG).

Estos valores de media mensual se encuentra en los rangos de los miles, siendo valores adimensionales, los valores de respuesta espectral de los índices de vegetación se normalizan entre -1 y 1 mostrando los valores pegados a 1 con alta actividad de fotosíntesis, aumento de la biomasa y buen estado de la vegetación. Para realizar esta normalización se ha establecido un factor de escala de 0,0001 que permite la estandarización de estos valores.

2.4- Selección y obtención de los datos de la cobertura vegetal

Con el objetivo de asignar los datos de los tipos de formaciones vegetales a las imágenes satélites se hace necesario contar con un mapa de la cobertura vegetal del área de estudio. La elección de este mapa se hace a partir de estudios existentes o realizando una clasificación con ayuda de la teledetección.

De no existir este mapa, es necesario realizar un mapa con la clasificación de la cobertura, el cual se puede realizar por varias vías. Una de las formas pudiera ser la digitalización de la cubierta a partir del mapa topográfico del territorio a la escala seleccionada para el estudio, posteriormente se le incorporaría la información del catastro con la ocupación del suelo existente (estas acciones se realizan con ayuda de los Sistemas de Información Geográfica). Teniendo como resultado el mapa de cobertura vegetal preliminar se debe ir con este al campo y realizar la comprobación del mismo delimitando correctamente las coberturas (Ponvert-Delisle, 2002).

La otra manera de realizar el mapa es utilizando las imágenes satelitales de alta resolución espacial, realizando una clasificación supervisada en caso de conocer el territorio, de lo contrario la clasificación se haría no supervisada y posteriormente en ambos casos, se comprobaría en el terreno este mapa preliminar, haciendo los ajustes a los polígonos con la información recogida en el campo.

De existir estudios previos, estos se elegirán teniendo en cuenta varios elementos principales: la escala, la leyenda empleada (clasificación de los tipos de coberturas), la fecha de confección y la calidad del dato.

La escala seleccionada estará directamente relacionada con la extensión superficial de nuestra área de estudio, así como a la escala de trabajo determinada previamente para la investigación. A su vez dará como resultado una leyenda que puede estar o no generalizada.

En el caso de la leyenda, como se parte de la escala, se pueden determinar varias generalizaciones de la clasificación como las determinadas por Ponvert-Delisle (2003).

Un primer nivel, muy general y abarcador, que representará todas las categorías del territorio. El nivel 2, también se inserta en un contexto generalizador propio para reflejar las categorías en ellos representados a escalas de publicación muy pequeñas (1: 250 000 a 1: 1 000 000 y aún menores). La descripción del territorio comienza en el tercer nivel, dicho por los especialistas y adoptado por varios países, como el de la generación de la cartografía a escala 1:100 000, mientras que el nivel 4, es más descriptivo aún y se inscribe en el contexto de la escala 1:50 000. Por último, el nivel 5 responde a un esquema muy detallado y es propio para la producción de cartografía a escala 1:25 000 y mayores (ver tabla 2.1).

Por la fecha de confección se tendrá en cuenta que el mapa seleccionado o a seleccionar se haya confeccionado dentro del período que se estudia en la investigación o de no estar dentro del mismo, al menos haber sido realizado cercano a la fecha de inicio del análisis.

Para la calidad del dato se deberá analizar la metodología empleada en la realización de la clasificación anterior, analizando cómo se obtuvieron los datos, es decir, si estos datos son obtenidos y evaluados en campo; si solamente se trabaja con los datos obtenidos por revisión bibliográfica; o si los datos son recopilados de imágenes aerocósmicas.

Siguiendo esta guía para la obtención de un mapa de cobertura vegetal nuestra investigación seleccionó el nuevo mapa de cobertura vegetal natural y seminatural de Cuba realizado por (Estrada et al., 2013). El mapa fue seleccionado por presentar una adecuada calidad de los datos aportados ya que este presenta una buena confiabilidad, siendo comprobada en campo todas las clasificaciones realizadas para su elaboración, además de tener información que responde al período de estudio y por tener una escala acorde a las características espaciales de nuestras imágenes satelitales.

Tabla 2.1 Leyenda de ocupación del suelo.

NIVEL 1	NIVEL 2	NIVEL 3	NIVEL 4	NIVEL 5	
1. VEGETACION NATURAL	1.1 BOSQUES	1.1.1 LATIFOLIAS	1.1.1.1 PLUVIAL	1.1.1.1.1 NUBLADO 1.1.1.1.2 SUB-MONTANO 1.1.1.1.3 MONTANO	
			1.1.1.2 SIEMPREVERDE	1.1.1.2.1 MICROFILO 1.1.1.2.2 MESOFILO 1.1.1.2.3 MANGLAR 1.1.1.2.4 BOSQUE DE CIENAGA	
			1.1.3 SEMIDECIDUO	1.1.1.3.1 MICROFILO 1.1.1.3.2 MESOFILO	
		1.1.2 PINARES	1.1.2.1 PINARES		
		1.2 VEGETACION ARBUSTIVA	1.2.1 MANIGUA COSTERA	1.2.1 MANIGUA COSTERA	1.2.1.1.1 MANIGUA COSTERA
			1.2.2 CUABAL	1.2.2 CUABAL	1.2.2.1.1 CUABAL
	1.2.3 CHARRASCAL		1.2.3 CHARRASCAL	1.2.3.1.1 CHARRASCAL	
	1.3 VEGETACION HERBACEA	1.3.1 HERBAZAL DE CIENAGA	1.3.2 HERBAZAL DE CIENAGA	1.3.2.1.1 HERBAZAL DE CIENAGA	
		1.3.2 SABANAS	1.3.2 SABANAS	1.3.2.1.1 SABANAS	
		1.3.3 VEGETACIÓN ACUÁTICA	1.3.4 VEGETACIÓN ACUÁTICA	1.3.4.1.1 VEGETACIÓN ACUÁTICA	
	1.4 COMPLEJO DE VEGETACION	1.4.1 VEGETACIÓN DE MOGOTE	1.4.2 VEGETACIÓN DE MOGOTE	1.4.2.1.1 VEGETACIÓN DE MOGOTE	
		1.4.2 VEGETACIÓN DE COSTA ROCOSA	1.4.3 VEGETACIÓN DE COSTA ROCOSA	1.4.3.1.1 VEGETACIÓN DE COSTA ROCOSA	
		1.4.4 VEGETACIÓN DE COSTA ARENOSA	1.4.5 VEGETACIÓN DE COSTA ARENOSA	1.4.5.1.1 VEGETACIÓN DE COSTA ARENOSA	
	2. VEGETACION SEMI-NATURAL	2.1 VEGETACIÓN SECUNDARIA	2.1.1 BOSQUE SECUNDARIO	2.1.1 BOSQUE SECUNDARIO	2.1.1.1.1 BOSQUE SECUNDARIO
			2.1.2 MATORRAL SECUNDARIO	2.1.3 MATORRAL SECUNDARIO	2.1.2.1.1 MATORRAL SECUNDARIO
			2.1.3 COMUNIDADES HERBÁCEAS SECUNDARIAS	2.1.3 COMUNIDADES HERBÁCEAS SECUNDARIAS	2.1.3.1.1 COMUNIDADES HERBÁCEAS SECUNDARIAS
3. PLANTACIONES	3.1 FORESTAL	MIXTAS	3.1.1.1 SIEMPREVERDES	3.1.1.1.1 SIEMPREVERDES	
		3.1.1 LATIFOLIAS	3.1.1.2 SEMIDECIDUOS	3.1.1.2.1 SEMIDECIDUOS	
			3.1.1.3 DECIDUOS	3.1.1.3.1 DECIDUOS	
	3.1.2 CONÍFERAS	3.1.2.1 PINARES	3.1.2.1.1 PINARES		
	3.2 NO FORESTAL	3.2.1 FRUTALES	3.2.1.1 ARBÓREOS	3.2.1.1.1 ARBÓREOS	
		3.2.2 NO FRUTAL	3.2.2.1 ARBUSTIVOS	3.2.2.1.1 ARBUSTIVOS	
3.2.2.1 NO FORESTAL HOMOGÉNEO			3.2.2.1.1 NO FORESTAL HOMOGÉNEO		

Fuente: Ponvert-Delisle (2002).

2.5- Selección del tipo de muestreo y tamaños de la muestra

La teoría del muestreo es un estudio de las relaciones existentes entre una población y las muestras extraídas de la misma. Tiene gran interés en muchos aspectos de la estadística. Permite por ejemplo, estimar cantidades desconocidas de la población (tales como la media poblacional, la varianza, etc.), frecuentemente llamadas parámetros poblacionales o brevemente parámetros, a partir del conocimiento de las correspondientes cantidades muestrales (tales como la media muestral, la varianza, etc.), a menudo llamadas estadísticos muestrales o brevemente estadísticos (Spiegel, 1978).

Durante el uso y empleo de la teoría del muestreo se presentan las siguientes dificultades:

- Elección de la muestra.
- Tipo de muestreo a desarrollar.

La identificación del tamaño de la muestra, es una tarea que se realiza durante el diseño de un estudio, ya que durante el mismo se estima la población a estudiar y por lo tanto se determinan la cantidad de individuos o elementos a participar en la misma.

El espectro de criterios para realizar el cálculo del tamaño muestral es amplio y variado, la elección concreta depende, entre otras cosas, de la información de que se disponga y de las características del procedimiento estadístico que se va a utilizar, para analizar los datos. Los parámetros que se utilizan en estas fórmulas están relacionados con los errores del mapa y la exactitud del mismo.

Una fórmula muy extendida que orienta sobre el cálculo del tamaño de la muestra, cuando se conoce el tamaño de la población a muestrear es la siguiente:

$$\frac{k^2 N p q}{e^2 (N - 1) + k^2 p q}$$

N: es el tamaño de la población o universo.

k: es una constante que depende del nivel de confianza que asignemos. El nivel de confianza indica la probabilidad de que los resultados de nuestra investigación sean ciertos. Los valores de k se obtienen de la tabla de la distribución normal estándar.

e: es el error muestral deseado.

p: proporción de individuos que poseen en la población la característica de estudio.

q: proporción de individuos que no poseen esa característica.

Cuando no se conoce el tamaño de la población a muestrear la fórmula varía de la siguiente manera:

$$n = \frac{Z_a^2 \times p \times q}{d^2}$$

Donde:

Z = nivel de confianza,

p = probabilidad de éxito o proporción esperada,

q = probabilidad de fracaso,

d = precisión (error máximo admisible en términos de proporción).

Estos cálculos a veces presentan limitantes, por lo que algunos autores han establecido unos valores para el tamaño muestral relacionados, con la extensión espacial del territorio y el tipo de muestreo a realizar.

Para que el muestreo sea representativo y para que los datos tengan una distribución normal, lo ideal sería realizar el mayor número de muestreos. A pesar que existen algunos métodos matemáticos para determinar el número de unidades muestrales, generalmente existen limitaciones de tiempo para realizar el número adecuado de muestreos. En estudios sobre ecología se debe muestrear el mayor número de unidades muestrales. Los criterios que generalmente se utilizan para determinar el tamaño de la muestra puede ser: la relación entre la superficie a muestrear y la superficie total, y la homogeneidad espacial de la variable o población a estudiarse.

Dadas las limitaciones de las fórmulas estadísticas, se recomienda obtener al menos 50 muestras (píxeles) por tipo de cobertura o de uso de suelo. Cuando el área sea muy grande (superior a las 350 000 ha) o el número de categorías sea superior a 12, el tamaño de muestra por categoría debe incrementarse a 75 o 100 muestras (Congalton, 1991).

Siguiendo esta afirmación realizada por varios especialistas que usan las técnicas de muestreo, en nuestra investigación, para el tamaño muestral realizamos la correspondencia de las características espaciales y cantidad de estratos de la cobertura vegetal (17 formaciones vegetales), en nuestra área, decidimos establecer el máximo de 100 muestras por formación, aunque en algunas formaciones por su extensión se establecieron 50 muestras o menos, haciendo la particularidad, de calcular el porcentaje de ocupación de cada formación dentro del área y atendiendo a los resultados alcanzados.

2.6- Intersección de los valores normalizados IV (por meses y anual) con los puntos de muestra de cada formación

Al tener en su poder las imágenes de satélite seleccionadas de su territorio, el mapa de coberturas, el muestreo de la cobertura, se lleva a cabo la intersección de estas informaciones. Esta intersección se realiza usando la herramienta intersectar de los SIG, donde se unen información areal y puntual, obteniendo una tabla con los datos de las capas intersectadas.

2.7- Exportación y determinación del comportamiento espectral de cada una de las formaciones

Esta tabla con los datos intersectados, se puede trabajar y realizar los análisis cualitativos dentro de un SIG, pero se ha comprobado que los gestores de base de datos son los idóneos para realizar estas operaciones.

Para la exportación de esta tabla, de un SIG a un gestor de base de datos, se abre la misma y se utiliza un comando de exportación de todos los valores, los cuales son guardados como *.DBF, los cuales si pueden ser propiamente abiertos por cualquier gestor.

Por último se abre un programa de gestión de base de datos y se comienzan a utilizar los comandos de análisis estadísticos y construcción de gráficos, permitiendo realizar un estudio relacionados con el comportamiento temporal de la respuesta espectral de la cobertura vegetal.

En el análisis geoestadístico se obtendrán los descriptores de la variable como el mínimo y el máximo valor de la respuesta, así como medidas de tendencia central (moda y media aritmética). Estos datos se obtienen usando el programa estadístico SPSS versión 22.

A partir de estos datos se grafican las series temporales de cada formación, a partir de las cuales se obtendrá el funcionamiento (anual) y evolución (durante el período de estudio) de este último gráfico se determina tendencia evolutiva de cada formación.

Después se determinará la correlación entre las variables de respuesta espectral y la precipitación.

La relación o correlación se refiere al grado de variación conjunta entre dos o más variables. Existen diferentes coeficientes de correlaciones que se utilizan para analizar las relaciones entre dos o más variables.

Entre estos coeficientes se encuentra el coeficiente de correlación de Pearson, pensado para variables cuantitativas (escala mínima de intervalo), es un índice que mide el grado de relación entre distintas variables relacionadas linealmente.

El coeficiente de correlación de Pearson es un índice de fácil ejecución e, igualmente, de fácil interpretación, oscila entre -1 y 1 . La magnitud de la relación viene dada por el valor numérico del coeficiente, reflejando el signo la dirección de tal valor. Tan fuerte es una relación de 1 como de -1 , en el primer caso la relación es perfecta positiva y en el segundo perfecta negativa (Ver figura 2.4).

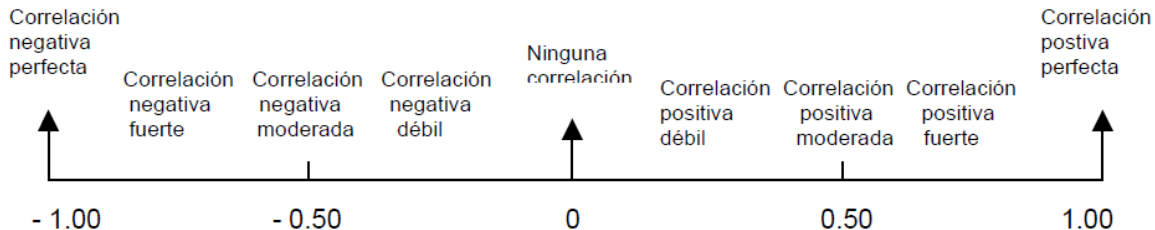


Figura 2.4 Análisis del coeficiente de correlación entre dos variables.

Fuente: (Vila et al., 2010)

2.8- Análisis de error

En este análisis se realiza una correlación entre la desviación típica de cada pixel y las distintas formaciones del territorio según el mapa empleado. La exactitud se realizará tomando cada punto de muestreo y comprobándolo contra la verdad terreno donde se emplearan imágenes de alta resolución espacial.

Para este análisis se emplearan imágenes de alta resolución las cuales servirán como la verdad terreno de cada formación utilizada para nuestro estudio.

Capítulo 2

Primero, se construye una tabla cruzada en un gestor de bases de datos con varias columnas, donde aparecen las formaciones del área de estudio, la cantidad de muestras y la verificación de las muestras con respecto a las imágenes MODIS y al mapa utilizado.

Posteriormente se realiza una comprobación de todos los puntos de muestra por formación, contra la respuesta dada y contra el mapa utilizado en la investigación.