

# ***ESTRUCTURA ESPACIAL DE LA SEQUÍA EN PASTOS Y SUS APLICACIONES EN EL SEGURO AGRARIO INDEXADO***

**AUTOR**

D. Juan José Martín Sotoca

**TUTORES**

Dña. Ana María Tarquis Alfonso

D. Antonio Saa Requejo

**TRABAJO FIN DE MÁSTER EN FÍSICA DE SISTEMAS COMPLEJOS**

**CURSO 2013/14**

**ESCUELA TÉCNICA SUPERIOR DE INGENIEROS AGRÓNOMOS (U.P.M.)**

*A mis padres y hermano*

## **AGRADECIMIENTOS**

Quiero expresar mi más sincero agradecimiento a los tutores de este Trabajo Fin de Master: Ana María Tarquis Alfonso y Antonio Saa Requejo, por todo el tiempo que han invertido en aconsejarme y guiarme por el tortuoso sendero de los multifractales y la inferencia estadística. Mi agradecimiento también a Javier Borondo, por el tiempo que ha empleado en proporcionarme los datos numéricos de MODIS.

En estos momentos no puedo olvidarme de mostrar mi agradecimiento a la persona que me animó a emprender el camino de la investigación: Juan Bautista Grau Olivé, guiándome en los pasos que debía dar, siendo la realización de este máster el primero que debía acometer.

Aprovecho también para dar las gracias a todos los profesores del Máster en Física de Sistemas Complejos por su excelente e incansable labor docente e investigadora.

También quiero dar las gracias al Centro de Estudios e Investigación para la Gestión de Riegos Agrarios y Medioambientales (CEIGRAM), por haber recibido la financiación que otorga en su convocatoria de ayudas a jóvenes investigadores de 2014.

# ÍNDICE

RESUMEN .....	5
1. INTRODUCCIÓN .....	6
1.1 Sequía y pastos .....	6
1.2 Índices satelitales de vegetación .....	7
1.3 El seguro de compensación por pérdida de pastos .....	8
1.3.1 Objeto .....	8
1.3.2 Definiciones .....	8
1.3.3 Aplicación del seguro .....	9
1.4 Objetivos .....	9
2. MATERIALES Y MÉTODOS .....	10
2.1 Índice de vegetación NDVI .....	10
2.2 Origen de los datos .....	10
2.3 Selección del área de estudio .....	11
2.4 Análisis espacial del índice NDVI .....	13
2.4.1 El espectro multifractal .....	14
2.4.2 El mapa de singularidades .....	14
2.5 Análisis estadístico de las series temporales del índice NDVI en pastos .....	15
2.5.1 El modelo de color HSL .....	15
2.5.2 Ajuste por máxima verosimilitud .....	17
2.5.3 Bondad de ajuste. Test de $\chi^2$ (Chi o Ji cuadrado) .....	19
3. RESULTADOS .....	21
3.1 Análisis espacial del índice NDVI .....	21
3.1.1 El espectro multifractal .....	21
3.1.2 El mapa de singularidades .....	23
3.2 Análisis estadístico del índice NDVI en pastos .....	27
4. CONCLUSIONES .....	34
5. REFERENCIAS .....	35
APÉNDICE A .....	36
APÉNDICE B .....	40

## RESUMEN

Períodos de sequía en zonas de pastos tienen graves repercusiones económicas para los ganaderos ya que se ven forzados a obtener alimento suplementario para sus animales. Para mitigar estos daños existe el seguro de compensación por pérdida de pastos de AgroSeguro. Este seguro utiliza el índice de actividad fotosintética NDVI como indicador de sequedad en los pastos. Uno de los puntos más importantes del seguro es decidir cuando hay que compensar a los ganaderos, es decir, definir cuando se produce el evento sequía.

Este Trabajo Fin de Máster trata dos aspectos importantes de la definición del evento sequía: cómo clasificar las zonas agrícolas con estadísticos similares del índice NDVI y cómo establecer los umbrales de decisión del evento sequía.

Para la clasificación de zonas realizamos un análisis espacial del índice NDVI de una zona rectangular que incluye a la Comunidad de Madrid. Del cálculo del espectro multifractal y el mapa de singularidades obtenemos información interesante a la hora de clasificar los píxeles. Nos interesarán los píxeles cuyo exponente de Lipschitz-Hölder sea aproximadamente 2.

Para el establecimiento de umbrales realizamos un análisis estadístico del índice NDVI de una zona de pastos situada en los municipios de Soto del Real y Colmenar Viejo (Comunidad de Madrid). Previamente realizamos un filtrado de la serie temporal del índice basado en el modelo de color HSL. A continuación realizamos un ajuste por máxima verosimilitud de varias distribuciones teóricas candidatas y obtenemos que la distribución de Gumbel de valores extremos mínimos y la Generalizada de Valores Extremos (GEV) son las distribuciones que mejor explican los valores observados del índice NDVI. Como consecuencia de estos resultados proponemos una modificación de los umbrales del evento sequía basada en el cálculo de percentiles.

**Palabras clave:** sequía, NDVI, Seguro Agrario, multifractal, exponentes de Lipschitz-Hölder, modelo HSL, distribución de Gumbel y distribución GEV.

# 1. INTRODUCCIÓN

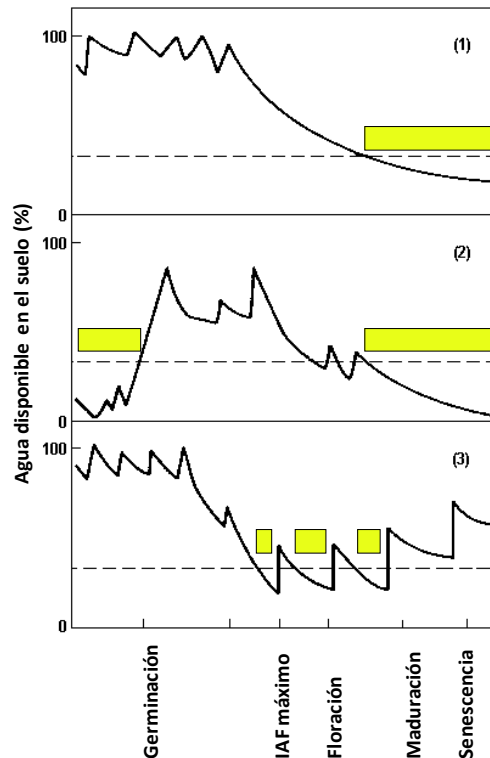
## 1.1 Sequía y pastos

De manera general la sequía se puede definir como la escasez temporal de agua, en relación con el suministro normal, en un período de tiempo prolongado. Esta definición podría calificarse de conceptual ya que no establece momento de inicio, finalización e intensidad de la sequía, ni un conocimiento más detallado de otras variables climáticas. Una definición más precisa de sequía, también llamada operacional, nos lleva a establecer distintos tipos de sequías dependiendo de la disciplina científica desde donde sea abordada. La clasificación con más aceptación establece cuatro tipos de definiciones de sequía:

- **Sequía meteorológica:** Se basa en las precipitaciones y mide sus desviaciones respecto a la media durante un período determinado de tiempo.
- **Sequía agrícola:** Se basa en la cantidad de humedad que necesita un determinado cultivo para desarrollarse. Por tanto no sólo depende de las precipitaciones que se produzcan sino que también depende de la capacidad de retención de agua del suelo, tipo de cultivo, hábitos de enraizamiento y estado de desarrollo (Hayes, 2004).
- **Sequía hidrológica:** Se basa en la cantidad de agua disponible en superficie (ríos, lagos, embalses) y subterránea.
- **Sequía socioeconómica:** Se basa en los daños económicos provocados por una disminución en los recursos hídricos. En este caso es determinante la manera en que la economía y la sociedad desarrollan sus actividades. Determinados sistemas socioeconómicos son más vulnerables a la escasez de agua.

En pascicultura se considera pasto todo tipo de vegetación que sirve de alimento al ganado, bien en pastoreo o bien como forraje. Hay diversas maneras de clasificar los pastos según el aspecto que se considere. Así, por su fisonomía se pueden distinguir los pastos arbóreos, arbustivos y herbáceos. Por su origen se hallan los pastos artificiales o de origen agrícola y los pastos naturales. Los pastos de origen agrícola son derivados de alguna actividad agrícola y comprenden los cultivos forrajeros (pastos sembrados en una rotación), los rastrojos (residuo de cosecha), los barbechos y el erial a pastos (antiguo terreno agrícola donde, por abandono del cultivo, crece vegetación espontánea).

Una característica del pasto es su fuerte estacionalidad: la producción se concentra en primavera y otoño, secándose en verano (Díaz-Ambrona, 1998a y 1998b). Centrándonos en pastos anuales, éstos están adaptados a los patrones usuales de sequía (ver figura 1). La diversidad de especies con ciclos de crecimiento diferentes hace que la incidencia de la sequía sea menor, aunque la falta de agua siempre supone una reducción de la producción.



**Figura 1.** Patrones de disponibilidad de agua en el suelo, identificando tres tipos de sequía: (1) terminal, (2) inicial y (3) variable. La línea rayada representa el límite inferior de disponibilidad hídrica en el suelo que soportarán para la evapotranspiración máxima del pasto y las barras los episodios de sequía. IAF: Índice de área foliar. Adaptado de Jordan y Miller (1980).

La mayor parte del pastoralismo se practica bajo condiciones de secano con un suministro de agua variable dependiente directamente de la lluvia, sólo en algunos casos incluye además los suministros locales por flujo superficial y por infiltración. Por lo tanto, las estrategias relacionadas con la gestión del agua en pastoreo en España están principalmente relacionadas con la falta de agua (San Miguel Ayanz, 2001).

## 1.2 Índices satelitales de vegetación.

En las últimas décadas se han puesto en órbita varios satélites (POES, Terra, Aqua, Deimos-1 y Deimos-2, etc) capaces de obtener imágenes de la superficie de la Tierra a intervalos de tiempo regulares y con una resolución aceptable. Estos satélites llevan sensores radiométricos (AVHRR, MODIS, etc) que han ido aumentando la cantidad de bandas de frecuencia que pueden detectar así como su resolución. Combinando estas bandas de frecuencia se han obtenidos los llamados índices de vegetación (IVs).

La justificación del uso de los IVs se fundamenta en la relación entre la fotosíntesis y la radiación absorbida/reflejada por la vegetación. La clorofila absorbe mayoritariamente la luz visible (de 0.4 a 0.7  $\mu\text{m}$ ) para que la fotosíntesis se produzca. Por otra parte, la estructura celular de las hojas refleja mayoritariamente la banda del infrarrojo cercano (de 0.7 a 1.1  $\mu\text{m}$ ). Por tanto, en la mayoría de las definiciones de los IVs se incorporará este principio. Existen varias definiciones de IVs donde se tienen en cuenta otros factores como el efecto del suelo o las influencias atmosféricas. En todos ellos, al final la definición será una combinación aritmética de distintas bandas de frecuencia. Se definen a continuación tres de los índices más utilizados:

- **Índice de Vegetación de Diferencia Normalizada** (Normalized Difference Vegetation Index, NDVI). Se basa fundamentalmente en el principio de que la vegetación sana absorbe mayoritariamente la luz visible y refleja el infrarrojo cercano. Intenta mitigar el efecto de las nubes y la nieve restando la banda visible del rojo.
- **Índice de Vegetación Ajustado al Suelo** (Soil Adjusted Vegetation Index, SAVI). Se trata de un índice muy adecuado para trabajos en zonas semiáridas, donde la contribución del suelo es muy importante. Así, cuando se realice un estudio sobre una zona de estas características, este índice resultará más consistente que el NDVI, gracias a esa mayor distinción entre el suelo y la vegetación.
- **Índice de Vegetación Mejorado** (Enhanced Vegetation Index, EVI). Este índice fue diseñado para tener en cuenta las zonas con alta densidad de biomasa, las variaciones estructurales de la cubierta vegetal y la influencia atmosférica.

### 1.3 El seguro de compensación por pérdida de pastos

El seguro de compensación por pérdida de pastos del Sistema Español de Seguros Agrarios Combinados está aplicando nuevas tecnologías, como la teledetección con índices de vegetación (Flynn, 2006), con el fin de solucionar los problemas de peritación que surgen cuando se tiene que determinar la cantidad y calidad del pasto en distintas zonas del territorio nacional. El seguro de compensación de los daños por sequía en pastos es actualmente el principal instrumento para hacer frente al gasto que supone la necesidad de suplemento de alimentación del ganado reproductor debido a la sequía.

A continuación damos una breve descripción del seguro de compensación por pérdida de pastos comercializado por Agroseguro (Agroseguro, 2014). Este tipo de seguro define el evento sequía en función de los valores que toma el índice satelital NDVI en un periodo determinado de tiempo, por lo que también recibe el nombre de seguro indexado.

Los datos del índice NDVI se obtienen del sensor MODIS a bordo del satélite Terra (EOS AM-1). La resolución de píxel utilizada por el Seguro es de 250 m x 250 m.

#### 1.3.1 Objeto

El seguro tiene como objeto indemnizar a los ganaderos cuando se produce un evento sequía en pastos y tienen que proporcionar un alimento suplementario al ganado reproductor.

#### 1.3.2 Definiciones

Las definiciones más importantes a tener en cuenta son las siguientes:

**Zona homogénea:** Cada Comarca será una zona homogénea.

**NDVI medio:** Es el índice NDVI decenal (cada diez días), calculado para cada zona homogénea de pastoreo, como media aritmética de los NDVIs de la serie 2000 al 2011 de MODIS.



**NDVI actual:** Es el índice NDVI decenal (cada diez días) del periodo de garantía, calculado para cada zona homogénea de pastoreo, a partir del Máximo Valor Compuesto Decenal (MVCD).

**NDVI garantizado** (umbrales): Se distinguen dos estratos de garantía:

1. **Estrato 1:** NDVI medio decenal, calculado para cada zona homogénea de pastoreo, menos 0.7 veces la desviación típica del mismo, para la misma decena.
2. **Estrato 2:** NDVI medio decenal, calculado para cada zona homogénea de pastoreo, menos 1.5 veces la desviación típica del mismo, para la misma decena.

Tanto el NDVI medio como su desviación típica se multiplicarán por el factor 0.99 a efectos del cálculo del NDVI garantizado.

**Máximo Valor Compuesto Decenal (MVCD):** Es el valor máximo de las lecturas diarias de cada decena de días para cada píxel. Representa el indicador de la actividad de los pastos en cada una de las decenas naturales del año, eliminando los efectos de las nubes, las discrepancias producidas por la diferente iluminación en cada uno de los diez días y demás efectos perturbadores.

### 1.3.3 Aplicación del seguro

El seguro establece cinco opciones diferentes de contratación. Salvo el caso especial de opción de franquiciado, todas las opciones comparten la siguiente definición de siniestro mínimo indemnizable: aquel en el que el NDVI actual sea inferior al NDVI garantizado en más de tres decenas a lo largo del periodo de garantías.

Implícitamente se está dando una definición del evento sequía como el siniestro mínimo indemnizable.

## 1.4 Objetivos

Nos hemos encontrado con dos ámbitos interesantes de reflexión en la definición que se hace del concepto sequía en el seguro de compensación por pérdida de pastos. Por un lado está el establecimiento de zonas homogéneas de pastos, es decir, zonas con similares características agrícolas (precipitaciones, tipo de suelo, tipo de cultivo, etc.), donde se puedan aplicar los mismos criterios de sequía. Por otro lado está el estudio de la estadística asociada al evento sequía en pastos, definición de inicio y fin de la sequía así como su intensidad. Todos los análisis realizados en este Trabajo Fin de Máster (TFM) se centrarán en la Comunidad de Madrid.

Por tanto, estos dos ámbitos más la aplicación de estos conceptos al seguro de compensación de Agroseguro constituirán los objetivos de este TFM. Esquemáticamente los objetivos serán:

- Análisis espacial del índice NDVI: explorar la estructura espacial del índice de vegetación NDVI para el establecimiento de zonas espacialmente homogéneas en la Comunidad de Madrid.
- Análisis estadístico de las series temporales del índice NDVI en pastos: profundizar en la definición del evento sequía en una zona de pastos seleccionada de la Comunidad de Madrid.
- Aplicar los resultados obtenidos al modelo de sequía en pastos de Agroseguro.

## 2. MATERIALES Y MÉTODOS

### 2.1. Índice de vegetación NDVI.

Como ya hemos comentado, el índice NDVI muestra la actividad fotosintética de la vegetación (Flynn, 2006). A su vez la actividad fotosintética está relacionada con la humedad que necesita un cultivo para desarrollarse correctamente, a mayor humedad disponible, mayor actividad fotosintética. Por tanto resulta razonable estudiar el evento sequía a través del índice NDVI. Ya hemos visto que existen otros índices más precisos que incorporan los efectos del suelo y las influencias atmosféricas, sin embargo por la sencillez de cálculo y su fácil interpretación utilizaremos el índice NDVI para todos los estudios realizados en este TFM. Por tanto, el evento sequía vendrá definido al establecer umbrales en el índice NDVI y su duración dependerá del tiempo que el índice permanezca por debajo de ese umbral.

Para el cálculo del índice NDVI utilizamos la expresión matemática:

$$NDVI = \frac{IRc - R}{IRc + R}$$

Donde  $IRc$  y  $R$  son los valores de la reflectancia en las bandas de frecuencia del Infrarrojo Cercano y del Rojo respectivamente. El rango de valores del índice NDVI es de -1 a +1. Valores inferiores a cero indican actividad fotosintética nula y es característico de zonas con gran acumulación de agua como ríos, lagos, embalses, etc. Cuanto más alto sea el valor de NDVI, mayor será la actividad fotosintética.

### 2.2. Origen de los datos

Para el cálculo del índice NDVI en las zonas de estudio se ha optado por la información que proporciona el satélite de investigación científica Terra (EOS AM-1) puesto en órbita por la NASA el 18 de diciembre de 1999. De todos los sensores que incorpora este satélite, el que nos va a proporcionar la información de bandas de frecuencia es el sensor MODIS (Moderate Resolution Imaging Spectroradiometer).

La mayoría de los sistemas de procesamiento de datos satelitales reconocen 5 niveles distintos de procesamiento de la información recogida. Los datos de nivel 0 son los datos proporcionados por los sensores en bruto, sin ningún tipo de corrección. Los datos de nivel 1 han sido sometidos únicamente a una calibración radiométrica. Los datos de nivel 2 son datos de nivel 1 que, con una corrección por la influencia de la atmosfera, proporcionan datos de reflectancia en la superficie terrestre (a nivel de suelo). Los datos de nivel 3 son datos de nivel 2 georeferenciados, normalmente también son datos que han sido promediados o reconstruidos de otros. Los datos de nivel 4 son datos de niveles inferiores sometidos a un procesamiento adicional.

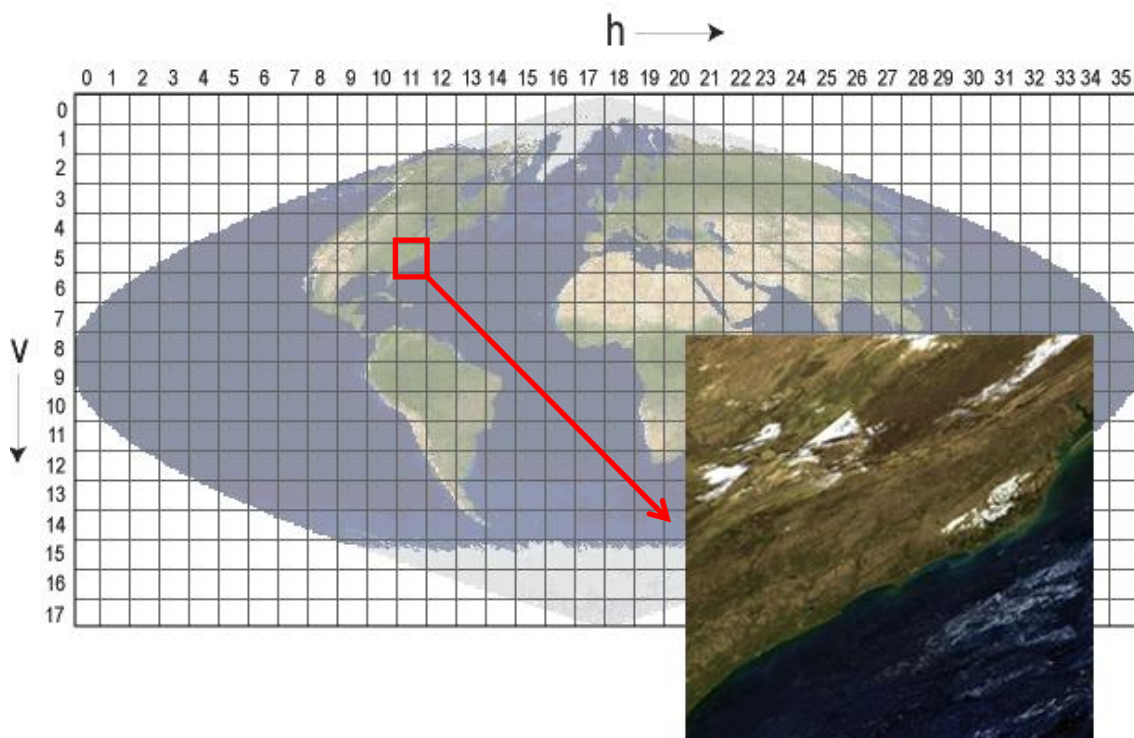
La información proporcionada por MODIS viene organizada por "productos". El producto que hemos utilizado es el MOD09A1 (LP DAAC, 2014).

MOD09A1 es el producto de MODIS que proporciona información de nivel 3 de reflectancia en superficie. Incorpora siete bandas de frecuencia: banda 1 (620-670 nm), banda 2 (841-876 nm), banda 3 (459-479 nm), banda 4 (545-565 nm), banda 5 (1230-1250 nm), banda 6 (1628-1652 nm) y banda 7 (2105-2155 nm). Al ser un producto de nivel 3, la reflectancia en superficie ha sido estimada como si hubiera sido medida a nivel del suelo, sin los efectos de la absorción y

scattering atmosférico. También se han corregido los efectos provocados por gases y aerosoles. MOD09A1 proporciona imágenes georeferenciadas con una resolución de píxel de 500 m x 500 m. Este producto hace una composición de las mejores mediciones de reflectancia de cada píxel en períodos de 8 días. Por tanto disponemos de una secuencia de 46 valores para cada píxel y año. Para este TFM se van a utilizar los datos del 2001 al 2012, ambos inclusive.

Las bandas utilizadas para el cálculo del índice NDVI son: banda 1 para el rojo y banda 2 para el infrarrojo cercano.

MODIS utiliza un sistema de proyección cartográfica sinusoidal donde la superficie terrestre se proyecta en una rejilla de 35 x 17 teselas rectangulares (ver figura 2). Las 4 teselas que contienen a España serían la h17v04, h17v05, h18v04 y h18v05.



**Figura 2.** Rejilla de 35 x 17 teselas rectangulares utilizada por MODIS para la proyección cartográfica sinusoidal. Se muestra como ejemplo la tesela h11v05 en formato RGB. Imagen procedente de la guía de usuario del producto MOD09A1 (LP DAAC, 2014).

MODIS permite cambiar el sistema de proyección cartográfica. En este TFM hemos decidido trabajar con el sistema de proyección UTM30 y modelo de geode WGS 84 por ser uno de los más implantados. El formato de fichero de imagen que hemos utilizado es GeoTIFF con información georeferenciada.

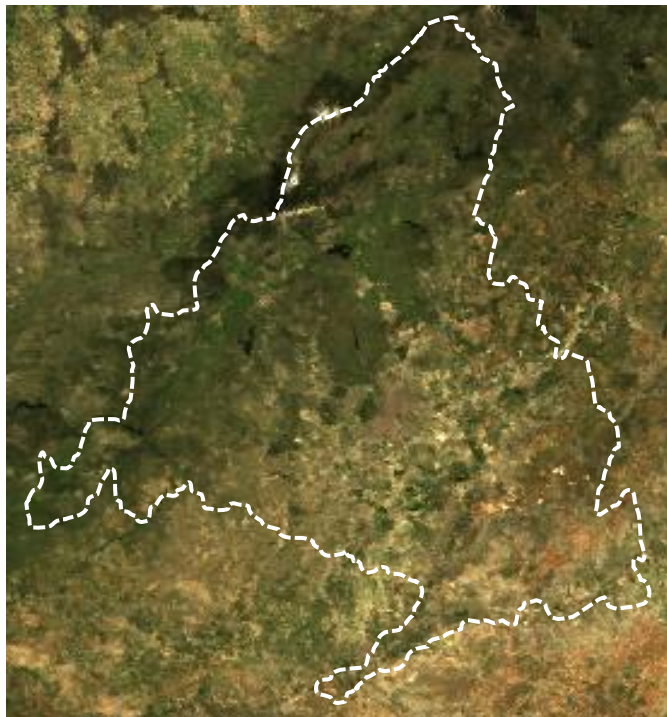
### 2.3. Selección del área de estudio

Debido a que se van a hacer dos tipos de análisis, uno espacial y otro estadístico, se van a considerar dos zonas de escalas diferentes.

Para el estudio espacial se va a considerar una zona rectangular que incluye a toda la Comunidad de Madrid. Esta zona está constituida por 300 x 280 píxeles de 500 m x 500 m (ver figura 3). Las coordenadas UTM30 de los límites de esta zona son:

- (1) Esquina superior izquierda: 360162 m E, 4560040 m N.
- (2) Esquina superior derecha: 499662 m E, 4560040 m N.
- (3) Esquina inferior izquierda: 360162 m E, 4410540 m N.
- (4) Esquina inferior derecha: 499662 m E, 4410540 m N.

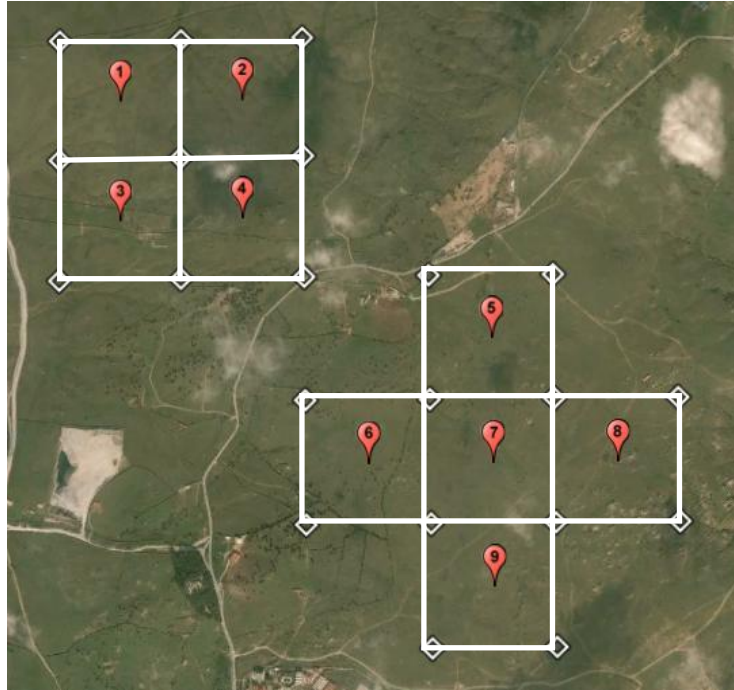
Donde la coordenada corresponde al centro del píxel de 500 m x 500 m.



**Figura 3.** Imagen RGB de la zona rectangular de 300 x 280 píxeles utilizada para el estudio espacial (fecha de captura por MODIS: 15/04/2011).

Para el estudio estadístico se ha seleccionado una zona de pastos situada en los municipios de Soto del Real y Colmenar Viejo (Comunidad de Madrid). Esta zona está constituida por 9 píxeles de 500 m x 500 m (ver figura 4). Las coordenadas UTM30 de los centros de los píxeles son:

- (1) 434662 m E, 4508040 m N.
- (2) 435162 m E, 4508040 m N.
- (3) 434662 m E, 4507540 m N.
- (4) 435162 m E, 4507540 m N.
- (5) 436162 m E, 4507040 m N.
- (6) 435662 m E, 4506540 m N.
- (7) 436162 m E, 4506540 m N.
- (8) 436662 m E, 4506540 m N.
- (9) 436162 m E, 4506040 m N.



**Figura 4.** Detalle de los 9 píxeles utilizados para el análisis estadístico. Cada cuadrado representa un píxel de 500 m x 500 m (imagen procedente de Google Earth). El sistema de proyección cartográfica utilizado es UTM30.

Gracias a que los ficheros GeoTiff están georeferenciados con sistema de proyección cartográfico UTM30 podemos seleccionar los píxeles que corresponden a estas áreas de estudio.

## 2.4. Análisis espacial del índice NDVI

Para el estudio espacial del índice NDVI utilizaremos dos herramientas de Teoría de la Medida, a saber: el cálculo del espectro multifractal y el cálculo del mapa de singularidades. Aplicaciones de estos conceptos se pueden encontrar en otros trabajos, como por ejemplo en los artículos de Lovejoy, Tarquis, Gaonac'h y Schertzer (2008) y Milne, Castellanos, Cartagena, Tarquis y Lark (2010).

En teoría de la medida (Falconer, 2003) una distribución de masa o medida  $\mu$  puede adoptar formas muy irregulares. De hecho, el conjunto de puntos donde la concentración local de masa o probabilidad local obedece a una ley de potencias de índice  $\alpha$ , es decir  $\mu(B(x,r)) \cong r^\alpha$  cuando  $r$  tiende a cero, puede formar diferentes conjuntos fractales para diferentes valores de  $\alpha$ . A los valores de  $\alpha$  se les denomina exponentes de Lipschitz-Hölder o de singularidad. Así, un gran abanico de fractales puede surgir de una medida  $\mu$ . A este tipo de medidas se las denomina "medidas multifractales".

Se define el espectro multifractal o de singularidades  $f(\alpha)$  a la función que nos proporciona la dimensión fractal del conjunto de puntos que cumplen  $\mu(B(x,r)) \cong r^\alpha$  cuando  $r$  tiende a cero.

Se define el mapa de singularidades al lugar geométrico de los puntos  $x$  que tiene el mismo exponente de Lipschitz-Hölder, es decir:

$$\alpha = T(x)$$

Donde  $T(\cdot)$  es la función que nos proporciona el exponente de Lipschitz-Hölder o de singularidad de cada punto  $x$ .

Supondremos que la distribución espacial del índice NDVI corresponde a una medida o distribución de masa  $\mu$  en  $\mathbb{R}^2$ .

### 2.4.1 El espectro multifractal

Para el cálculo del espectro multifractal seguiremos a Feder (1989) y utilizaremos el denominado “método de los momentos” por “conteo de cajas”. Sea  $S$  un conjunto con  $N$  puntos, si hacemos una partición de  $S$  mediante celdas de lado  $\delta$ , tendremos  $N_i$  puntos en cada una de esas  $i$ -ésima celdas. Utilicemos la probabilidad  $\mu_i = \frac{N_i}{N}$  de la celda  $i$ -ésima para construir la siguiente medida del conjunto  $S$ :

$$M_d(q, \delta) = \sum_{i=1}^N \mu_i^q \delta^d = X(q, \delta) \delta^d \xrightarrow{\delta \rightarrow 0} \begin{cases} 0 & d > \tau(q) \\ \infty & d < \tau(q) \end{cases}$$

Esta medida tiene un “exponente de masa”  $d = \tau(q)$  para el que la medida ni se anula ni diverge a  $+\infty$  cuando  $\delta \rightarrow 0$ . El exponente de masa  $\tau(q)$  dependerá del momento  $q$  elegido. La medida estará caracterizada por una secuencia de exponentes  $\tau(q)$  que controlan cómo los momentos de las probabilidades  $\{\mu_i\}$  escalan con  $\delta$ .

Se define la función de momentos como:

$$X(q, \delta) = \sum_{i=1}^N \mu_i^q \sim \delta^{-\tau(q)}$$

Por lo que el exponente de masa vendrá dado por la expresión:

$$\tau(q) = -\lim_{\delta \rightarrow 0} \frac{\ln X(q, \delta)}{\ln \delta} = -\lim_{\delta \rightarrow 0} \frac{\ln \sum_{i=1}^N \mu_i^q}{\ln \delta}$$

Para el cálculo del exponente de Lipschitz-Hölder o de singularidad  $\alpha(q)$  y el espectro multifractal  $f(\alpha)$  utilizamos las siguientes expresiones:

$$\alpha(q) = -\frac{d}{dq} \tau(q)$$

$$f(\alpha(q)) = q\alpha(q) + \tau(q)$$

### 2.4.2 El mapa de singularidades

Para el cálculo del mapa de singularidades o exponentes de Lipschitz-Hölder seguiremos a Cheng (2001) y Falconer (2003). Para medidas multifractales los valores de la probabilidad  $\mu(x)$  calculados para varios tamaños de celda  $\delta$  centrados en el punto  $x$  siguen una ley de potencias con  $\delta$ :

$$\mu(x) \sim \delta^{\alpha(x)}$$

Esta ley de potencias se cumple en cierto rango de  $\delta$ . Las localizaciones  $\alpha(x) \neq E$ , donde  $E$  es la dimensión topológica del problema, se denominan localizaciones singulares. Es más, podemos diferenciar las singularidades positivas  $\alpha(x) < E$ , de las singularidades negativas  $\alpha(x) > E$ . El exponente de singularidad o de Lipschitz-Hölder  $\alpha(x)$  puede ser calculado mediante la expresión:

$$\alpha(x) = \lim_{\delta \rightarrow 0} \frac{\ln \mu(x)}{\ln \delta}$$

Cheng (2001) aplica estos conceptos a un mapa con geolocalización de concentración de un cierto elemento químico. Establece que las concentraciones  $Z(x)$  calculadas para varios tamaños de celda  $\delta$  centradas en el punto  $x$  siguen una ley de potencias con  $\delta$ :

$$Z(x) \sim \delta^{\alpha(x)-E}$$

donde  $E=2$  para problemas 2-D y  $E=1$  para problemas 1-D. Las singularidades positivas con  $\alpha(x) < 2$  se corresponden con valores elevados de concentración del elemento químico, mientras que las singularidades negativas con  $\alpha(x) > 2$  se corresponden con valores bajos de concentración. Por tanto, la estimación de singularidades en un mapa de concentración de un elemento químico puede ser utilizada para caracterizar los patrones de concentración del elemento proporcionando información muy útil a la hora de interpretar anomalías relacionadas con la mineralización o procesos locales.

Para el caso de estudio de este TFM podremos dar interpretaciones similares a las que se hace en Cheng (2001) respecto al índice NDVI. En lugar de hablar de concentración de un elemento químico podremos hablar de concentración de actividad fotosintética. Sin embargo, los lugares que más nos interesarán serán aquellos en los que  $\alpha(x) \cong 2$ , ya que la variabilidad espacial del índice en estas zonas será relativamente uniforme y podríamos considerarlas como zonas homogéneas.

## 2.5. Análisis estadístico de las series temporales del índice NDVI en pastos

Previamente al estudio estadístico realizamos un filtrado de las series temporales del índice NDVI utilizando para ello el modelo de color HSL. Una vez filtradas se hace un análisis estadístico por máxima verosimilitud para distintas funciones densidad de probabilidad buscando la mejor bondad de ajuste.

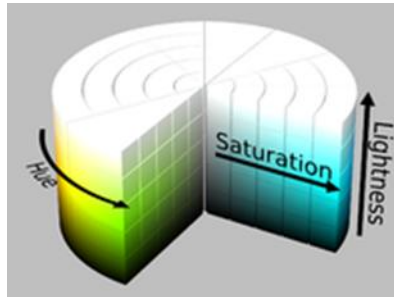
### 2.5.1 El modelo de color HSL

Una característica de los pastos es precisamente que su color evoluciona al mismo tiempo que lo hace su actividad fotosintética. Por tanto, los pastos pasarán por una gama de colores que va desde el verde (con actividad fotosintética) al amarillo pajizo (pasto seco).

Nos vamos a basar en esta característica de los pastos para descartar píxeles que se corresponden con colores "grisáceos". Un color "grisáceo" del píxel corresponderá normalmente con pastos cubiertos por nieve o por nubes espesas, por lo que el valor que tomará el índice NDVI será anómalo.

El modelo de color HSL nos proporciona una definición de lo que entendemos por color "grisáceo". También nos permite cuantificar este concepto, por lo que podremos establecer umbrales para descartar los valores que consideremos no deseables. A continuación pasamos a describir brevemente el modelo de color HSL.

El modelo de color HSL (Levkowitz, 1997) es una transformación geométrica del modelo RGB dando lugar a un sistema de coordenadas cilíndrico (ver figura 5).

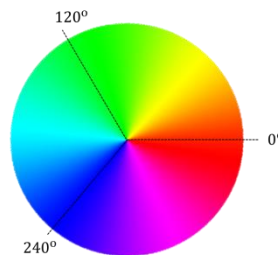


**Figura 5.** Sistema de coordenadas cilíndrico del modelo de color HSL (Imagen procedente de Wikimedia Commons).

Este nuevo sistema de coordenadas resulta más intuitivo que el sistema RGB cartesiano ya que está inspirado en la tradicional rueda de color.

Describimos brevemente las tres coordenadas que definen este modelo:

- **matiz** (Hue) es el ángulo alrededor del eje central. Empieza en el rojo (valor  $0^\circ$ ), pasa por el verde (valor  $120^\circ$ ), el azul ( $240^\circ$ ) y vuelve al rojo inicial ( $360^\circ$ ) (ver figura 6). Se corresponde con la longitud de onda dominante del color que vemos.



**Figura 6.** Rueda de color del modelo HSL.

- **saturación** (Saturation) es la distancia al eje central, toma valores de 0 a 1. Se corresponde con el grado de pureza de un color. Colores puros serán aquellos con saturación igual a 1. Colores con saturación muy baja, cercanos al eje central, son llamados colores neutros, acromáticos o grisáceos. Estos colores grisáceos van desde el negro puro, con valor de luminosidad 0, al blanco puro, con valor de luminosidad 1.
- **luminosidad** (Lightness) es la altura del cilindro, toma valores de 0 a 1. Se corresponde con la claridad u oscuridad de un color. Colores con luminosidad cercana a 1 serán casi blancos mientras que colores con luminosidad cercana a 0 serán casi negros.

En la circunferencia,  $L=1/2$  y  $S=1$ , estarán localizados los colores puros, mezclas lineales de los colores primarios y secundarios.

A continuación se muestran las transformaciones entre el modelo RGB y el modelo HSL:



$$H = \begin{cases} \text{no definido,} & \text{si } MAX = MIN \\ 60^\circ \times \frac{G-B}{MAX-MIN} + 0^\circ, & \text{si } MAX = R \\ & \text{y } G \geq B \\ 60^\circ \times \frac{G-B}{MAX-MIN} + 360^\circ, & \text{si } MAX = R \\ & \text{y } G < B \\ 60^\circ \times \frac{B-R}{MAX-MIN} + 120^\circ, & \text{si } MAX = G \\ 60^\circ \times \frac{R-G}{MAX-MIN} + 240^\circ, & \text{si } MAX = B \end{cases}$$

$$L = \frac{1}{2}(MAX + MIN)$$

$$S = \begin{cases} 0, & \text{if } MAX = MIN \\ \frac{MAX-MIN}{MAX+MIN} = \frac{MAX-MIN}{2L}, & \text{if } L \leq \frac{1}{2} \\ \frac{MAX-MIN}{2-(MAX+MIN)} = \frac{MAX-MIN}{2-2L}, & \text{if } L > \frac{1}{2} \end{cases}$$

Donde R, G y B son los valores de la reflectancia de la banda del rojo, verde y azul respectivamente, y donde:

$$MAX = \max(R, G, B)$$

$$MIN = \min(R, G, B)$$

Ya habíamos señalado que nuestro objetivo era descartar los píxeles que se corresponden con colores grisáceos. Por las definiciones dadas anteriormente estos colores serían los situados cerca del eje central, con valores de saturación cercanos a cero. Por tanto, ya estamos en disposición de definir un umbral en la coordenada de la Saturación que filtre nuestra secuencia de valores NDVI.

### 2.5.2 Ajuste por máxima verosimilitud

Las series temporales del índice NDVI, una vez filtradas según se indica en el apartado anterior, pasan a formar parte de la fuente de información disponible del análisis estadístico. Ya hemos señalado que nuestro objetivo es establecer un umbral en el índice NDVI por debajo del cual consideraremos sucede el evento sequía.

Definimos pues la variable aleatoria (v.a.)  $X$  que toma los valores del índice NDVI en la zona de estudio. En realidad tendremos 46 v.a.'s ( $X_i$  con  $i=1, \dots, 46$ ) ya que un año viene muestreado cada 8 días disponiendo por tanto de 46 valores por año y píxel.

Para la estimación de las funciones densidad de probabilidad (f.d.p.) de las v.a.'s  $X_i$  probaremos con las siguientes distribuciones conocidas:

- **Distribución normal:** Su f.d.p. viene definida como:

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Donde  $\mu$  es la media y  $\sigma$  es la desviación típica de la v.a.  $X$ .

- **Distribución gamma:** Su f.d.p. viene definida como:

$$f(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$$

Donde  $\Gamma(\cdot)$  es la función gamma,  $\alpha$  y  $\beta$  son los parámetros de la distribución.

- **Distribución beta:** Su f.d.p. viene definida como:

$$f(x; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}$$

$\Gamma(\cdot)$  es la función gamma,  $a$  y  $b$  son los parámetros de la distribución. La v.a.  $X$  está definida en  $0 \leq x \leq 1$ .

- **Distribución Gumbel de valores extremos mínimos:** su f.d.p. viene definida como:

$$f(x; \mu, \sigma) = \frac{1}{\sigma} t(x) e^{-t(x)} \quad \text{donde } t(x) = e^{(x-\mu)/\sigma}$$

$\mu \in \mathbb{R}$  es el parámetro de localización,  $\sigma > 0$  es el parámetro de escala.

- **Distribución Generalizada de Valores Extremos (GEV):** Su f.d.p. viene definida como:

$$f(x; \mu, \sigma, \xi) = \frac{1}{\sigma} t(x)^{\xi+1} e^{-t(x)} \quad \text{donde } t(x) = \begin{cases} \left(1 + \left(\frac{x-\mu}{\sigma}\right)\xi\right)^{-1/\xi} & \text{si } \xi \neq 0 \\ e^{-(x-\mu)/\sigma} & \text{si } \xi = 0 \end{cases}$$

$\mu \in \mathbb{R}$  es el parámetro de localización,  $\sigma > 0$  es el parámetro de escala y  $\xi \in \mathbb{R}$  es el parámetro de forma. El soporte de esta función viene dado por:

$$\begin{aligned} x &\in \left[\mu - \frac{\sigma}{\xi}, +\infty\right) \text{ cuando } \xi > 0, \\ x &\in (-\infty, +\infty) \text{ cuando } \xi = 0, \\ x &\in \left(-\infty, \mu - \frac{\sigma}{\xi}\right] \text{ cuando } \xi < 0. \end{aligned}$$

Para la estimación de los parámetros  $\theta = \{\alpha, \beta, \mu, \sigma, \dots\}$  de las citadas f.d.p.'s utilizaremos el método de máxima verosimilitud (MV).

El método de MV (Larson, 1982) se basa en la función de densidad conjunta de  $n$  v.a.'s  $X_1, \dots, X_n$  dependientes de los parámetros  $\{\alpha, \beta, \mu, \sigma, \dots\}$  sobre los cuales maximizamos la función de densidad conjunta para el caso de una realización  $x_1, \dots, x_n$ .

Definimos la función de verosimilitud de  $n$  v.a.'s  $X_1, \dots, X_n$  como la función densidad de probabilidad conjunta de las  $n$  variables, es decir  $f(x, \theta)$ , la cual es considerada como una función de  $\theta = \{\alpha, \beta, \mu, \sigma, \dots\}$ . En particular si  $x = (x_1, \dots, x_n)$  es una muestra aleatoria de f.d.p.  $f(x, \theta)$ , entonces la función de verosimilitud es:

$$L = f(x, \theta) = \prod_{i=1}^n f(x_i; \alpha, \beta, \mu, \sigma, \dots)$$

En la práctica se utiliza el logaritmo de esta función a la hora de maximizar la función.

El estimador de MV de  $\theta = (\alpha, \beta, \mu, \sigma, \dots)$  es el formado por los valores  $(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}, \dots)$  que maximizan la función de verosimilitud de la muestra  $(x_1, \dots, x_n)$  obtenida. Las principales propiedades de este estimador son:

- Invarianza.
- Consistencia.
- Insesgado asintóticamente.
- Normalidad asintótica.

El cálculo de los estimadores de MV de los parámetros de las f.d.p.'s planteadas al principio de este apartado lo resolvemos de la siguiente manera:

- Para el caso de la distribución normal los estimadores de los parámetros  $\mu$  y  $\sigma$  se pueden calcular utilizando expresiones exactas, que son la media muestral y la desviación típica muestral respectivamente.

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \hat{\sigma} = s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

- Para el caso de las distribuciones gamma, beta, Gumbel y GEV tenemos que recurrir a métodos numéricos para su resolución. En este TFM se han utilizado las funciones que MATLAB tiene disponibles para este fin.

### 2.5.3 Bondad del ajuste. Test de $\chi^2$ (Chi o Ji cuadrado)

Para determinar si el conjunto de observaciones del índice NDVI procede de una determinada distribución de probabilidad utilizaremos el test  $\chi^2$  (Larson, 1982). A continuación planteamos los conceptos de forma resumida.

Sea  $f(x, \theta)$  la f.d.p. teórica conocida de la v.a.  $X$  que depende del parámetro  $\theta = (\alpha, \beta, \mu, \sigma, \dots)$ , y sea  $x_1, \dots, x_n$  una muestra aleatoria de tamaño  $n$  de la v.a.  $X$  agrupada en  $k$  clases con  $n_i$  datos en la clase  $i$ .

Establecemos la siguiente hipótesis:

- $H_0$  = los datos se ajustan a la distribución teórica  $f(x, \theta)$ .
- $H_1$  = los datos no se ajustan a la distribución teórica  $f(x, \theta)$ .

Definimos el estadístico de prueba  $\chi_c^2$  como:

$$\chi_c^2 = \sum_{i=0}^k \frac{(n_i - e_i)^2}{e_i}$$

Donde  $n_i$  es el número de datos o frecuencia observada en la clase  $i$ , y  $e_i = n \cdot P(\text{clase } i)$  es la frecuencia esperada de la clase  $i$ .  $P(\text{clase } i)$  es la probabilidad de que la v.a.  $X$  tome valores dentro del intervalo definido por la clase  $i$ .

Establecemos también un nivel de significación como:

$$\alpha = P(\text{Rechazar } H_0 / H_0 \text{ es verdadero})$$

Por último establecemos la siguiente **regla de decisión**: "rechazar  $H_0$  al nivel de significación  $\alpha$  si:

$$\chi_c^2 = \sum_{i=0}^k \frac{(n_i - e_i)^2}{e_i} > \chi_{(k-m-1, 1-\alpha)}^2$$

Donde  $\chi_{(k-m-1, 1-\alpha)}^2$  es una distribución  $\chi^2$  con  $k - m - 1$  grados de libertad ( $m$  es el número de parámetros).

Por tanto, mediante esta prueba podremos decidir cual de las cinco f.d.p.'s candidatas se ajusta a la estadística anual del índice NDVI en la zona de pastos seleccionada.

### 3. RESULTADOS

#### 3.1 Análisis espacial del índice NDVI

Para el estudio de este TFM disponemos de imágenes de 300 x 280 píxeles que contienen a la Comunidad de Madrid. Para cada pixel hemos calculado el valor del índice NDVI, que como ya sabemos toma valores entre -1 y +1.

Los píxeles con índice NDVI menores que cero son normalmente zonas de acumulación de agua sin actividad fotosintética. En este estudio no interesa cómo se distribuye el índice en estas zonas por lo que los valores de estos píxeles se igualarán a cero

Sea  $A (a_{j,k})$  la matriz de 300 filas x 280 columnas con los valores del índice NDVI. Previamente al análisis espacial realizamos dos sencillas tareas:

1. Recortamos la matriz para dejarla con un tamaño de 256 x 256, ya que el algoritmo que calcula el espectro requiere tamaños de matriz potencia de 2 (aunque también se podría implementar con otras potencias).
2. Normalizamos la matriz, de tal forma que el valor de la probabilidad de la celda  $i$ -ésima de lado  $\delta$  será:

$$\mu_i = \frac{\sum_{j,k \in \text{celda } i} a_{j,k}}{\sum_{j=1}^{256} \sum_{k=1}^{256} a_{j,k}}$$

Para el análisis espacial vamos a comparar los resultados obtenidos en cuatro fechas diferentes del año 2011-2012. Se ha tomado una fecha representativa por cada estación del año. Las fechas seleccionadas fueron: 15 de abril de 2011 (primavera), 12 de julio de 2011 (verano), 16 de octubre de 2011 (otoño) y 1 de enero de 2012 (invierno).

##### 3.1.1 El espectro multifractal

Se han realizado los cálculos para  $q$  en el rango  $[-10,+10]$ . Las gráficas con los resultados obtenidos para la función de momentos  $X(q, \delta)$  y exponentes de masa  $\tau(q)$  de las cuatro fechas se muestran en el apéndice A. En esta sección compararemos los resultados que se obtuvieron del cálculo de los exponentes de Lipschitz-Hölder  $\alpha(q)$  y el espectro multifractal  $f(\alpha(q))$ .

Todos los cálculos se han realizado utilizando MATLAB. Para el cálculo de las pendientes de la función de momentos  $X(q, \delta)$  se ha utilizado la función de MATLAB "polyfit", la cual utiliza el método de mínimos cuadrados para obtener la recta de regresión. Se ha utilizado un tamaño mínimo de  $\delta$  de  $2^{-6}$ . Por debajo de estos valores de  $\delta$  el coeficiente de determinación  $R^2$  de las rectas de regresión generadas es demasiado bajo para considerar los resultados aceptables.

En la figura 7 se muestran los resultados para el exponente de Lipschitz-Hölder.

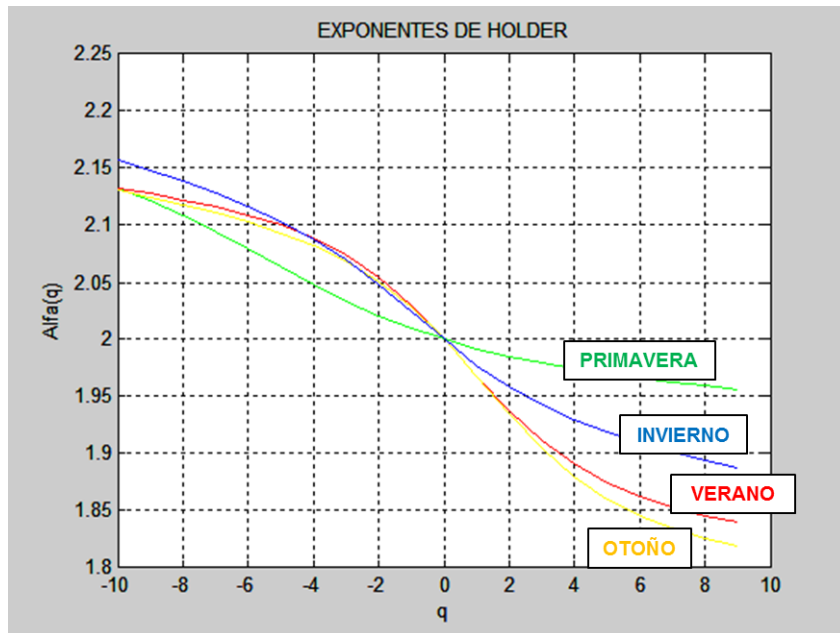


Figura 7. Exponentes de Lipschitz-Hölder para las cuatro fechas seleccionadas del año 2011-2012.

En la figura 8 se muestran los resultados para el espectro multifractal.

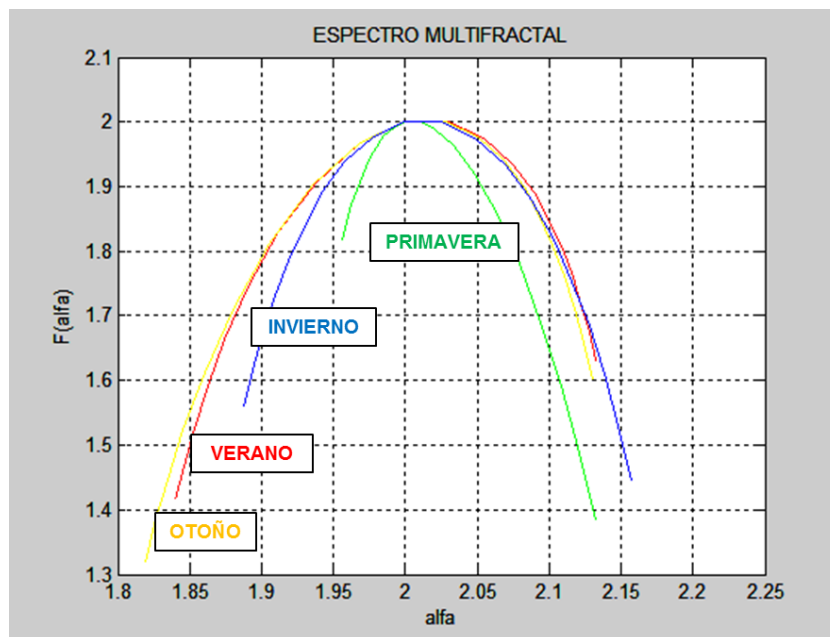


Figura 8. Espectro multifractal para las cuatro fechas seleccionadas del año 2011-2012.

Del gráfico del espectro multifractal vemos que el valor  $\alpha_{min}$  para cada fecha se va haciendo más pequeño según la secuencia primavera-invierno-verano-otoño, a la vez que su dimensión fractal va disminuyendo, es decir, valores de  $f(\alpha_{min})$  cada vez más bajos. El valor  $\alpha_{max}$  para cada fecha no varía mucho con el paso de las estaciones aunque en su dimensión fractal  $f(\alpha_{max})$  cabe distinguir dos valores, uno más alto para las estaciones de primavera e invierno y otro más bajo para las estaciones de verano y otoño.

### 3.1.2 El mapa de singularidades

En esta sección mostraremos los resultados del mapa de singularidades para un exponente de Lipschitz-Hölder continuo y para un exponente “discretizado”, es decir, para exponentes situados en un determinado rango de valores. Dicha discretización se realizó bajo las siguientes condiciones:

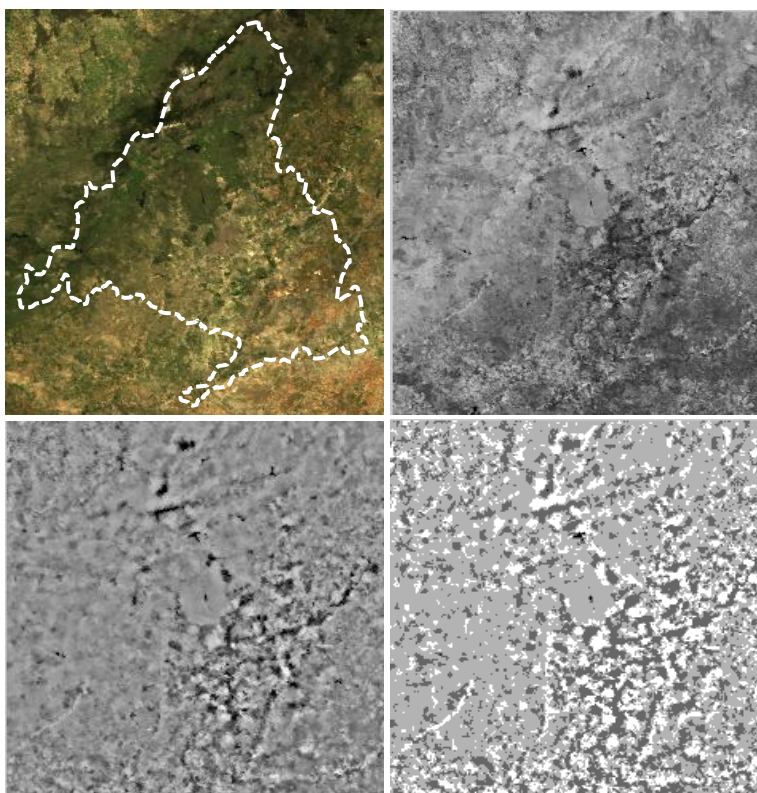
- $\alpha_A$  = exponentes  $\alpha < 1.95$ . Zona de aumento de la concentración del índice NDVI (color blanco).
- $\alpha_B$  = exponentes  $1.95 < \alpha < 2.05$ . Zona neutra o zona homogénea (color gris claro).
- $\alpha_C$  = exponentes  $\alpha > 2.05$ . Zona de reducción de la concentración del índice NDVI (color gris oscuro).
- Zona de agua (color negro).

Todos los cálculos se han realizado en MATLAB. Para esta sección también se ha utilizado la función de MATLAB “polyfit” para el cálculo de pendientes.

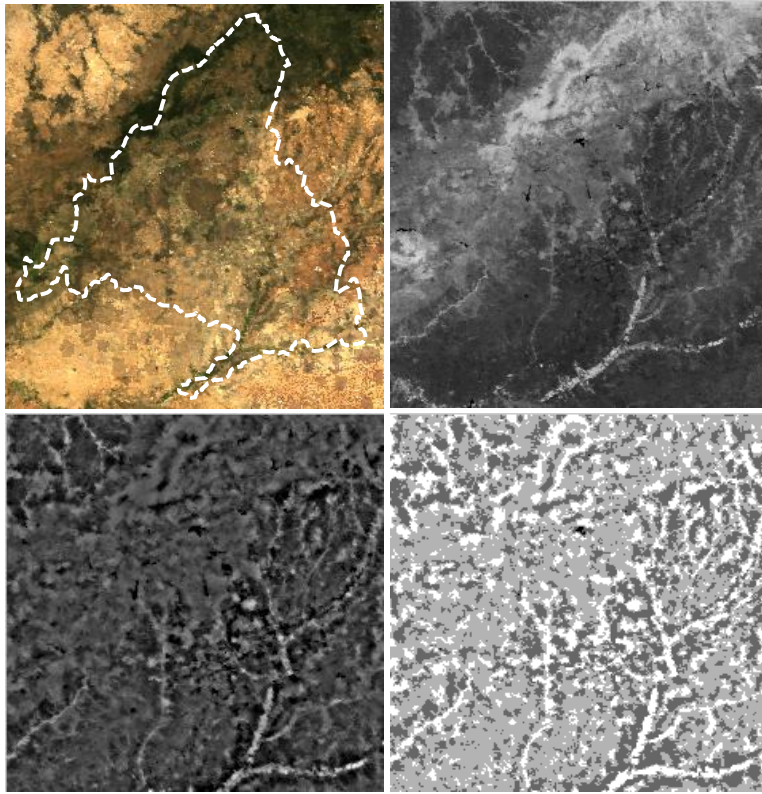
Para cada fecha se mostrarán cuatro imágenes: la imagen en color RGB, la imagen con el índice NDVI en escala de grises (el color blanco es NDVI = 1), el mapa de singularidades continuo en escala de grises (el color blanco es  $\alpha = \alpha_{\min}$ ) y el mapa de singularidades “discretizado” (con los valores señalados anteriormente).

En el Apéndice A se muestran los histogramas del exponente de Lipschitz-Hölder para cada una de las fechas analizadas, donde se pueden ver los valores máximo y mínimo del exponente.

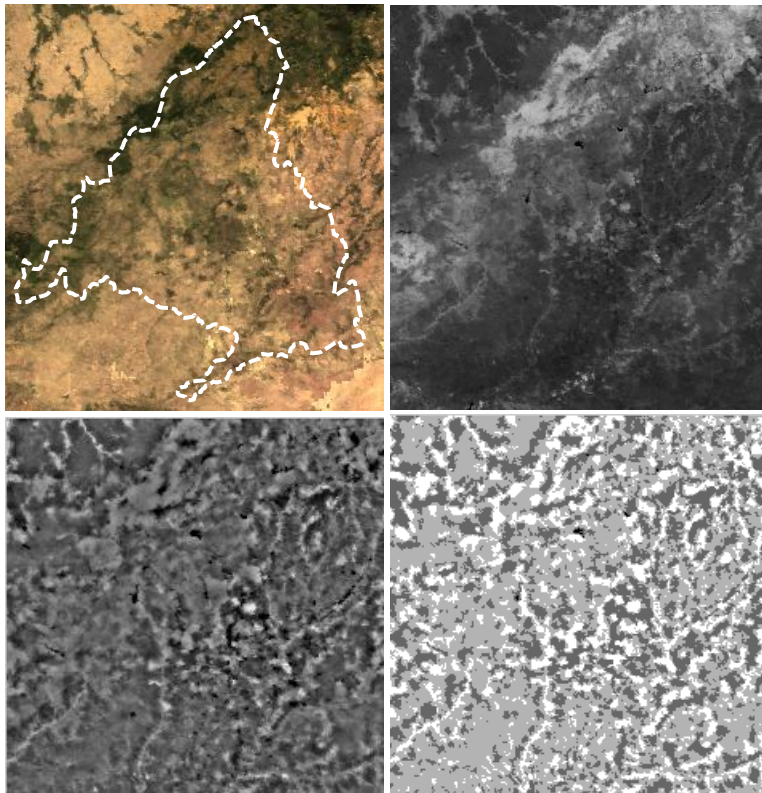
Se presentan a continuación los resultados en las figuras 9, 10, 11 y 12:



**Figura 9.** Imágenes de la zona de estudio con fecha 15/04/2011 (de izquierda a derecha y de arriba a abajo: RGB, NDVI, Hölder continuo, Hölder discretizado).

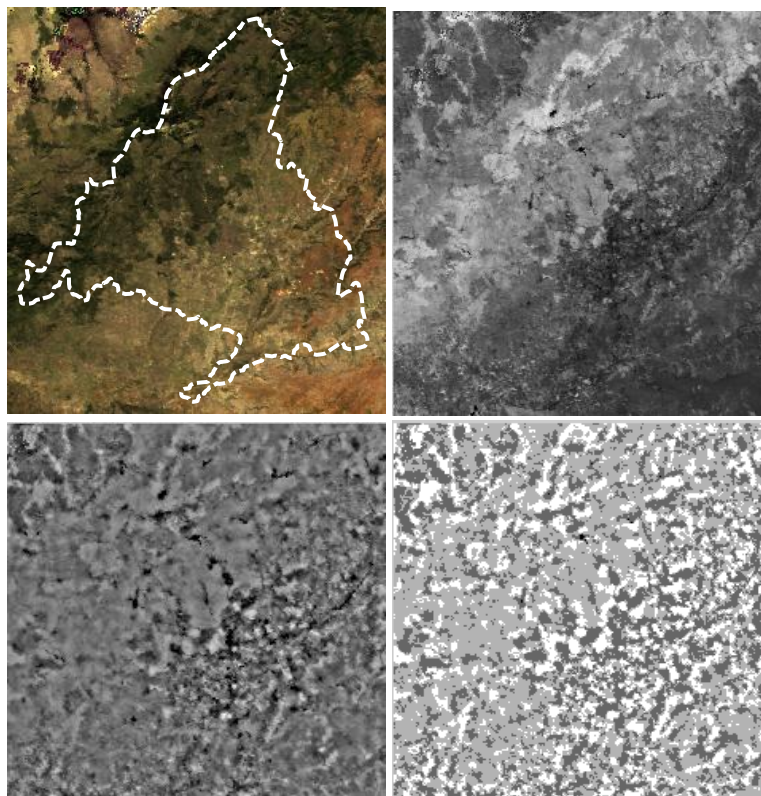


**Figura 10.** Imágenes de la zona de estudio con fecha 12/07/2011 (de izquierda a derecha y de arriba a abajo: RGB, NDVI, Hölder continuo, Hölder discretizado).



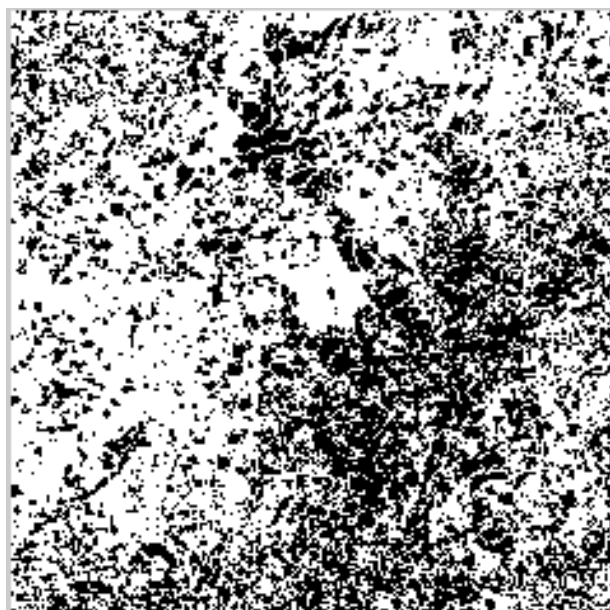
**Figura 11.** Imágenes de la zona de estudio con fecha 16/10/2011 (de izquierda a derecha y de arriba a abajo: RGB, NDVI, Hölder continuo, Hölder discretizado).



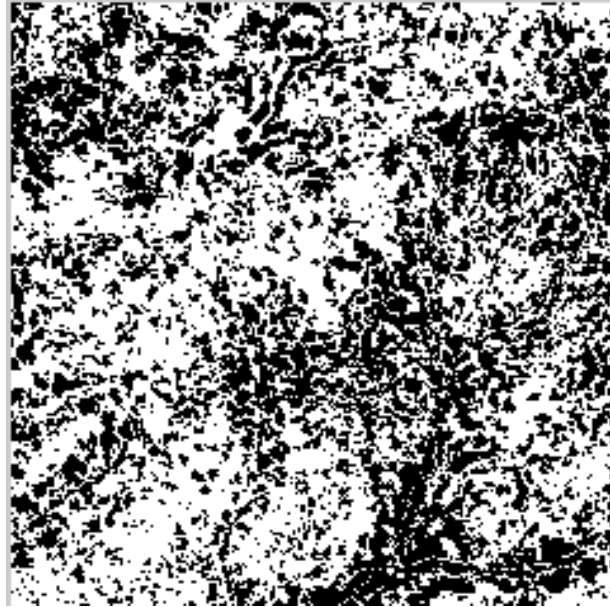


**Figura 12.** Imágenes de la zona de estudio con fecha 01/01/2012 (de izquierda a derecha y de arriba a abajo: RGB, NDVI, Hölder continuo, Hölder discretizado).

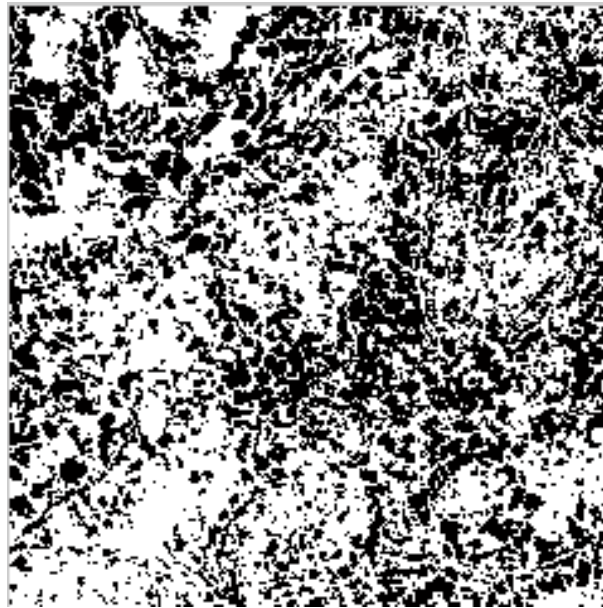
Si definimos una zona homogénea como aquellos píxeles con un exponente de Lipschitz-Hölder  $\alpha(x) \cong 2$  entonces obtenemos las figuras 13, 14, 15 y 16 para las cuatro fechas analizadas. Las zonas con píxeles en blanco constituyen los puntos para los que  $1.95 \leq \alpha(x) < 2.05$ . Los píxeles blancos conexos constituirán zonas homogéneas en la fecha analizada.



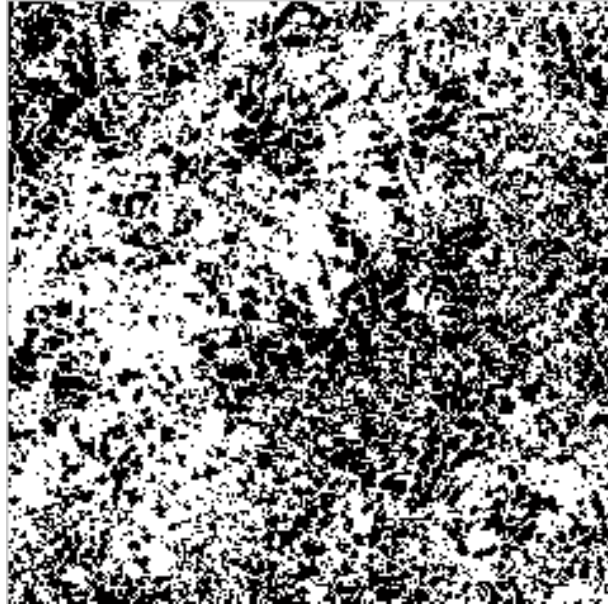
**Figura 13.** Zonas homogéneas mostradas como píxeles blancos ( $1.95 \leq \alpha(x) < 2.05$ ) de la zona de estudio con fecha 15/04/2011.



**Figura 14.** Zonas homogéneas mostradas como píxeles blancos ( $1.95 \leq \alpha(x) < 2.05$ ) de la zona de estudio con fecha 12/07/2011.

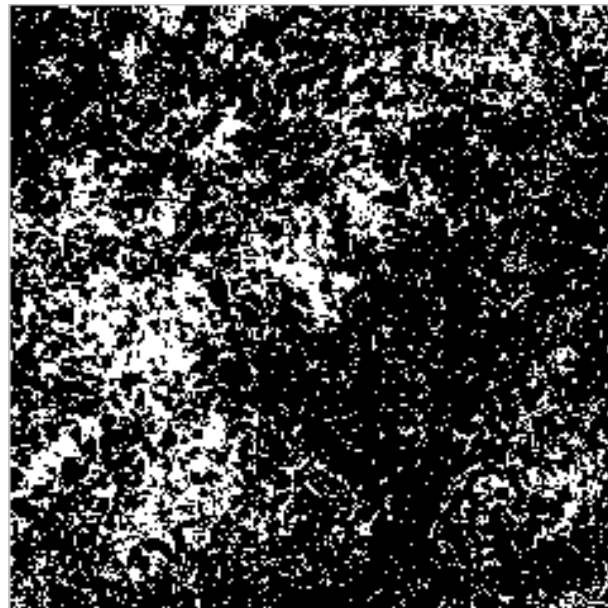


**Figura 15.** Zonas homogéneas mostradas como píxeles blancos ( $1.95 \leq \alpha(x) < 2.05$ ) de la zona de estudio con fecha 16/10/2011.



**Figura 16.** Zonas homogéneas mostradas como píxeles blancos ( $1.95 \leq \alpha(x) < 2.05$ ) de la zona de estudio con fecha 01/01/2012.

Si combinamos las cuatro imágenes obtenemos la figura 17, es decir, zonas que mantienen su homogeneidad a lo largo de un año (en rigor sería al menos en las cuatro fechas).



**Figura 17.** Zonas homogéneas mostradas como píxeles blancos ( $1.95 \leq \alpha(x) < 2.05$ ) resultante de la combinación de las cuatro fechas.

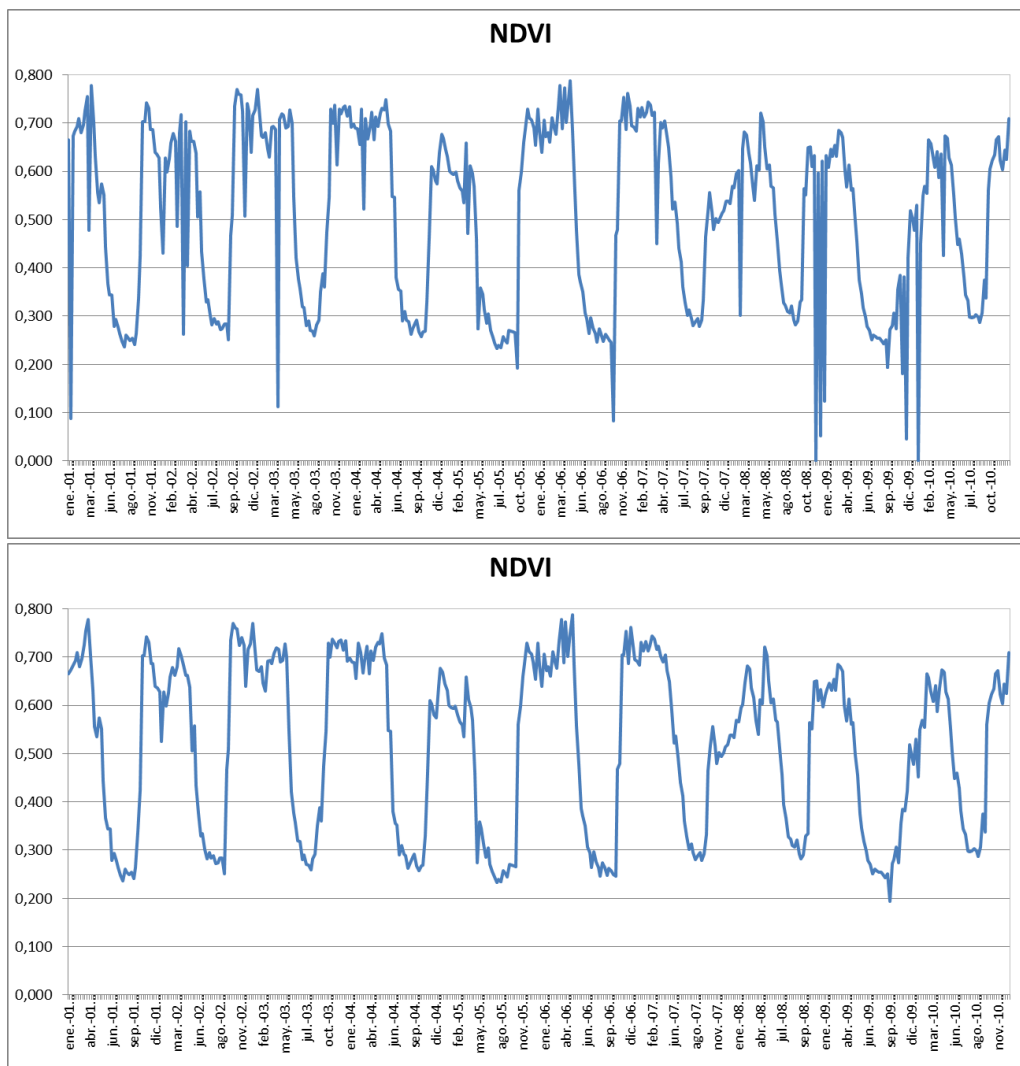
### 3.2 Análisis estadístico del índice NDVI en pastos

De los nueve píxeles que consta la zona de pastos de estudio se obtienen nueve series temporales del índice NDVI (de 2001 a 2010). Para la limpieza de estas series utilizamos el siguiente criterio: los píxeles válidos de la serie serán aquellos que tengan una saturación (modelo de color HSL) mayor o igual que 0,15. Píxeles por debajo de una saturación de 0,15

son píxeles con un color poco definido, tendiendo al gris, con una probabilidad muy alta de que el pasto este cubierto por nieve o nubes.

En la figura 18 se muestra un ejemplo en la serie temporal procedente del píxel nº 7. En la parte superior está la serie sin filtrar desde el 2001 al 2010. En la parte inferior nos hemos quedado con los valores de índice NDVI cuyo valor de saturación HSL sea mayor o igual que 0,15.

Una vez filtradas todas las series de la zona de estudio se agruparon todos los datos que corresponden al mismo intervalo. Recordemos que cada año dispone de 46 intervalos o “muestras cada 8 días”. Todos los valores de cada intervalo  $i$  serán realizaciones de las v.a.´s  $X_i$  que queremos modelar ( $i=1,\dots,46$ ).



**Figura 18.** Serie temporal del índice NDVI del año 2001 al 2010 sin filtrar (gráfica superior) y filtrada utilizando en criterio de saturación HSL (gráfica inferior) del píxel número 7 de la zona de estudio.

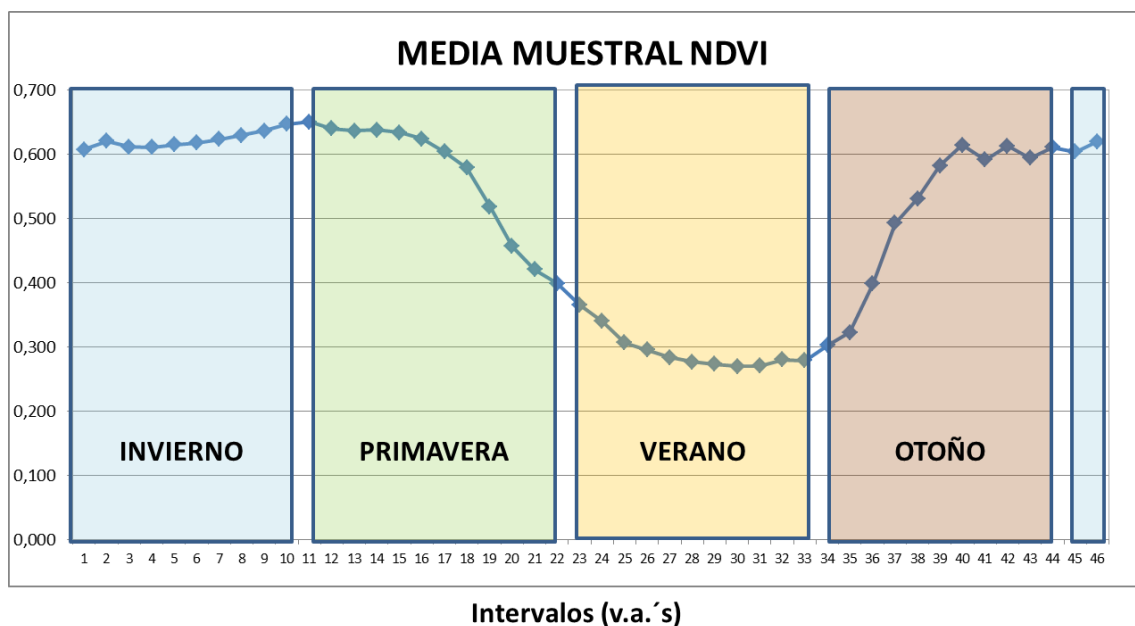
Para el estudio de las 46 v.a.´s hemos procedido a obtener por máxima verosimilitud (MV) los parámetros de las distribuciones candidatas: normal, gamma, beta, Gumbel de valores extremos mínimos y GEV. Para ello hemos utilizado las funciones MATLAB: normfit, gamfit, betafit, evfit y gevfit respectivamente. En el apéndice B se muestran todos los parámetros obtenidos por MV.

En la tabla 1 se muestra el inicio y el final del intervalo de cada v.a., así como el número de muestras disponibles para la obtención de los parámetros por MV.

INTERVALO (Nº V.A.)	INICIO	FINAL	Nº DE MUESTRAS
1	1-ene.	8-ene.	81
2	9-ene.	16-ene.	63
3	17-ene.	24-ene.	84
4	25-ene.	1-feb.	83
5	2-feb.	9-feb.	85
6	10-feb.	17-feb.	90
7	18-feb.	25-feb.	68
8	26-feb.	5-mar.	80
9	6-mar.	13-mar.	89
10	14-mar.	21-mar.	90
11	22-mar.	29-mar.	64
12	30-mar.	6-abr.	67
13	7-abr.	14-abr.	74
14	15-abr.	22-abr.	78
15	23-abr.	30-abr.	89
16	1-may.	8-may.	90
17	9-may.	16-may.	87
18	17-may.	24-may.	90
19	25-may.	1-jun.	90
20	2-jun.	9-jun.	90
21	10-jun.	17-jun.	88
22	18-jun.	25-jun.	90
23	26-jun.	3-jul.	90
24	4-jul.	11-jul.	90
25	12-jul.	19-jul.	90
26	20-jul.	27-jul.	90
27	28-jul.	4-ago.	90
28	5-ago.	12-ago.	90
29	13-ago.	20-ago.	90
30	21-ago.	28-ago.	90
31	29-ago.	5-sep.	90
32	6-sep.	13-sep.	90
33	14-sep.	21-sep.	88
34	22-sep.	29-sep.	90
35	30-sep.	7-oct.	90
36	8-oct.	15-oct.	87
37	16-oct.	23-oct.	80
38	24-oct.	31-oct.	90
39	1-nov.	8-nov.	83
40	9-nov.	16-nov.	77
41	17-nov.	24-nov.	84
42	25-nov.	2-dic.	81
43	3-dic.	10-dic.	86
44	11-dic.	18-dic.	74
45	19-dic.	26-dic.	90
46	27-dic.	31-dic.	74

**Tabla 1.** Descripción de cada uno de los intervalos en los que se divide un año. Se muestra el inicio y fin de los intervalos, así como el nº de muestras disponibles para cada uno.

En la figura 19 se muestran las medias muestrales del índice NDVI de las 46 v.a.'s así como una referencia del comienzo y final de cada estación astronómica del año.



**Figura 19.** Gráfica con las medias muestrales del índice NDVI para cada uno de los intervalos en los que se divide el año. Se muestra esquemáticamente el comienzo y final de las estaciones.

Para comprobar la bondad de ajuste de las distribuciones candidatas se ha utilizado el test Chi cuadrado para 9 clases y para 15 clases cumpliendo el requisito de que en cada clase haya más de cinco observaciones. En la tabla 2 aparecen de forma resumida los resultados del test (una X indica que ha pasado el test).

INTERVALO (Nº V.A.)	NORMAL	GAMMA	BETA	GUMBEL (min)	GEV
1	X	X	X	X	X
2	X	X	X	X	X
3					
4	X	X	X	X	X
5	X	X	X	X	X
6	X	X	X	X	X
7				X	X
8	X		X	X	X
9				X	
10			X	X	X
11	X	X	X	X	X
12			X	X	
13	X	X	X	X	X
14	X		X	X	
15				X	
16				X	X
17					
18				X	
19					
20		X			X
21					
22					
23					X

INTERVALO (Nº V.A.)	NORMAL	GAMMA	BETA	GUMBEL (min)	GEV
24	X	X	X		X
25					X
26	X	X	X		X
27					
28				X	
29					X
30					X
31	X	X	X	X	
32	X	X	X		
33					
34					
35					X
36	X	X	X		
37					
38	X		X	X	X
39	X	X	X	X	X
40	X	X	X	X	X
41				X	X
42	X		X	X	X
43				X	X
44					
45	X	X	X	X	X
46					X

Tabla 2. Resultado del test Chi cuadrado para cada intervalo y cada f.d.p. candidata (una x significa que pasa el test).

Para una mejor evaluación de la bondad de las distribuciones candidatas se han considerado 4 tramos en un año que no coinciden con las estaciones astronómicas (ver figura 20).

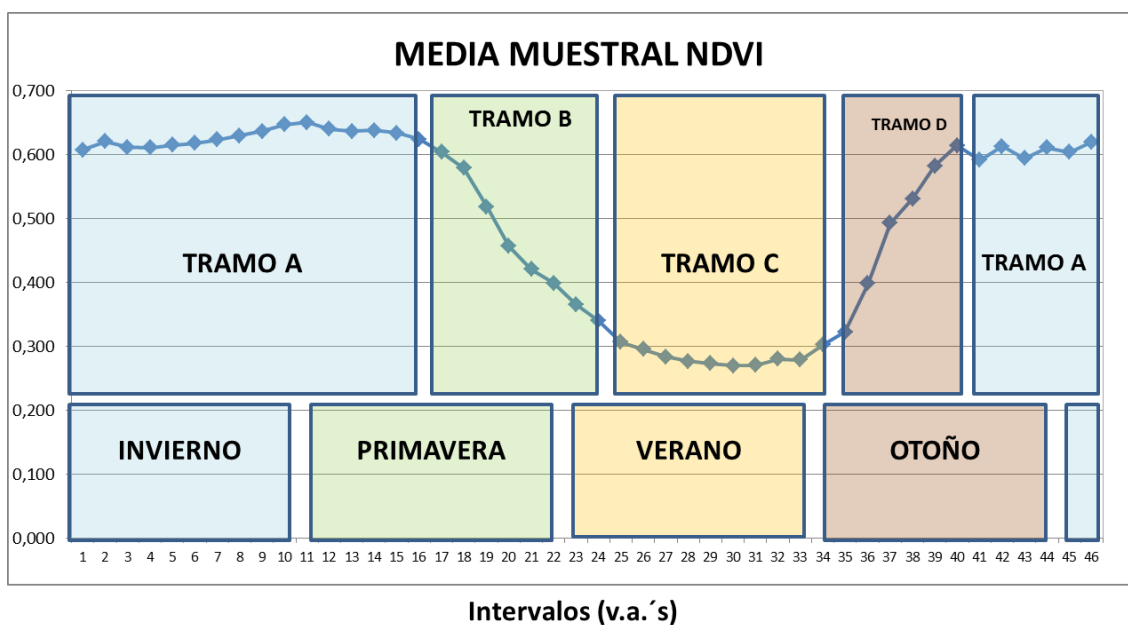


Figura 20. Gráfica con las medias muestrales del índice NDVI para cada uno de los intervalos en los que se divide el año. Se muestra esquemáticamente el comienzo y final de los cuatro tramos A, B, C y D utilizados para la evaluación de las f.d.p. s candidatas.

A continuación se describe cada tramo:

- **TRAMO A:** Desde el intervalo 1 al 16 (del 1-ene al 8-may) y del intervalo 41 al 46 (del 17-nov al 31-dic). En este tramo la media muestral del índice NDVI es alta y estable. Correspondería con todo el invierno, el inicio de la primavera y el final del otoño.
- **TRAMO B:** Desde el intervalo 17 al 24 (del 9-may al 11-jul). En este tramo la media muestral del índice NDVI baja de manera rápida. Correspondería con el final de la primavera e inicio del verano.
- **TRAMO C:** Desde el intervalo 25 al 34 (del 12-jul al 29-sep). En este tramo la media muestral del índice NDVI es baja y estable. Correspondería con gran parte del verano.
- **TRAMO D:** Desde al intervalo 35 al 40 (del 30-sep al 16-nov). En este tramo la media muestral del índice NDVI sube de manera rápida. Correspondería con el final del verano e inicio del otoño.

En la tabla 3 podemos ver el nº de intervalos en los que el test Chi cuadrado es positivo para cada uno de los tramos.

TRAMOS	INTERVALOS DEL TRAMO	NORMAL	GAMMA	BETA	GUMBEL (min)	GEV
TRAMO A	22	11	8	14	18	16
TRAMO B	8	1	2	1	1	3
TRAMO C	10	3	3	3	2	4
TRAMO D	6	4	3	4	3	4
<b>TOTAL</b>	<b>46</b>	<b>19</b>	<b>16</b>	<b>22</b>	<b>24</b>	<b>27</b>

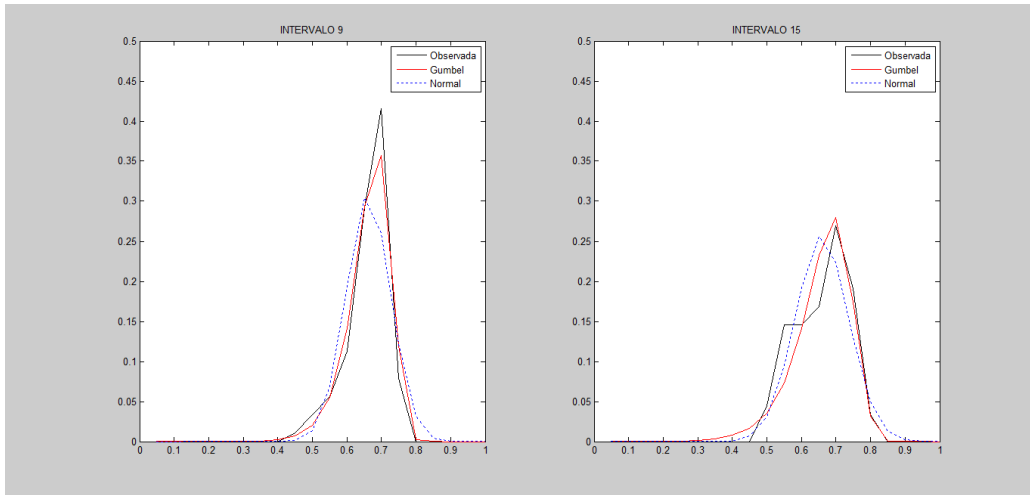
**Tabla 3.** Resultados del test Chi cuadrado para cada tramo y cada f.d.p candidata.

En el tramo A la distribución que más intervalos explica es la Gumbel de valores extremos mínimos (18 de 22 intervalos). En los tramos B y C sería la distribución GEV, aunque el % de intervalos explicados es muy bajo. En el tramo D no hay un candidato claro.

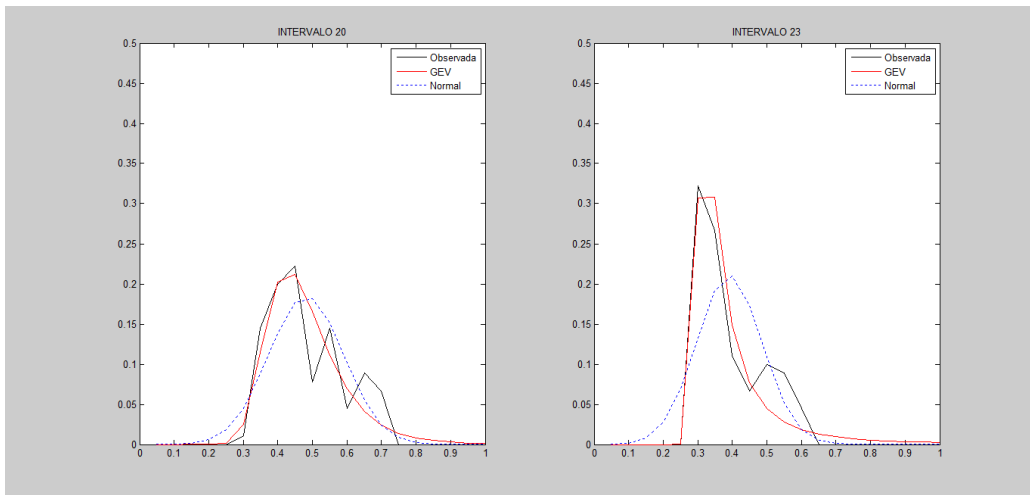
Si hacemos una valoración global, sin atender a la clasificación por tramos, vemos que la distribución GEV es la que más intervalos explica (27 de 46 intervalos).

En el apéndice B se muestra, para cada uno de los 46 intervalos, la distribución observada y la distribución “ganadora” del tramo al que pertenece. En todas ellas se muestra también la distribución normal. Las funciones MATLAB utilizadas para representar las distribuciones teóricas fueron: normcdf, gamcdf, betacdf, evcdf y gevcd.

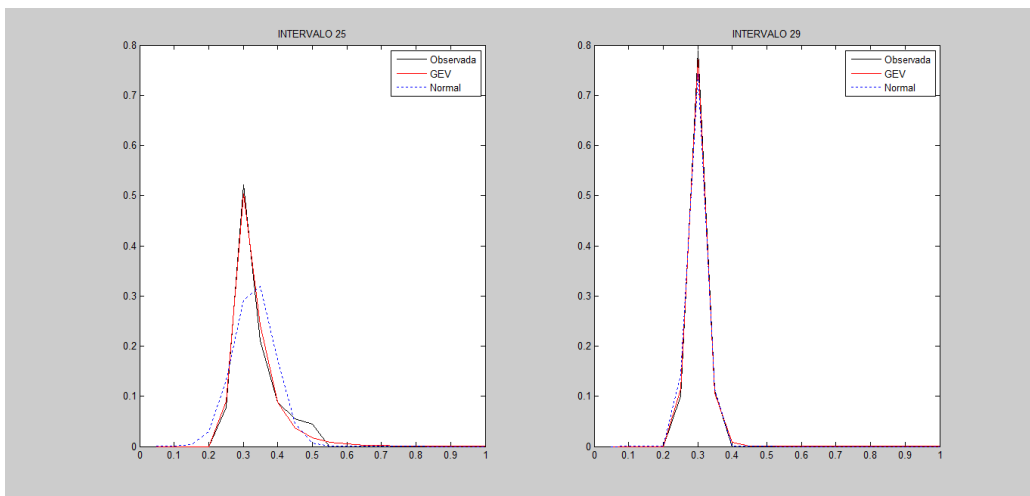
Presentamos en las figuras 21, 22, 23 y 24, como ejemplo, dos intervalos de cada tramo.



**Figura 21.** F.d.p. observada, Gumbel y normal para los intervalos 9 y 15 del Tramo A. La distribución normal no pasa el test Chi cuadrado en ninguno de los dos intervalos.

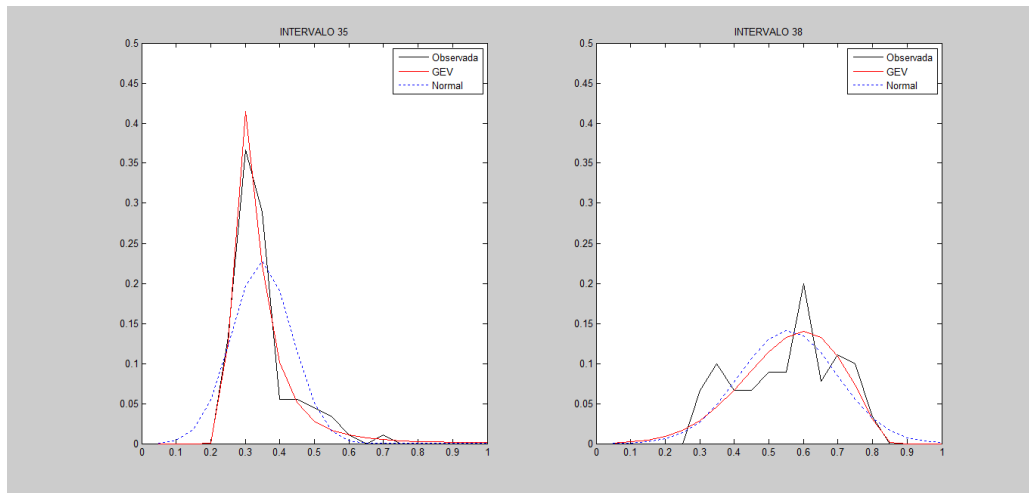


**Figura 22.** F.d.p. observada, GEV y normal para los intervalos 20 y 23 del Tramo B. La distribución normal no pasa el test Chi cuadrado en ninguno de los dos intervalos.



**Figura 23.** F.d.p. observada, GEV y normal para los intervalos 25 y 29 del Tramo C. La distribución normal no pasa el test Chi cuadrado en ninguno de los dos intervalos.





**Figura 24.** F.d.p. observada, GEV y normal para los intervalos 35 y 38 del Tramo D. La distribución normal no pasa el test Chi cuadrado en el intervalo 35 y si en el 38.

Recordemos que el seguro de compensación por pérdida de pastos de Agroseguro había definido el siniestro indemnizable o evento sequía como aquella situación en la que el índice NDVI actual fuera inferior al índice NDVI garantizado en más de tres decenas. El cálculo del NDVI garantizado se realizaba en función de la media del índice NDVI en una decena menos un  $k$  veces la desviación típica, donde  $k$  podía ser 0.7 ó 1.5 dependiendo del estrato contratado.

Atendiendo a los resultados obtenidos en la zona de estudio vemos que la f.d.p. normal no es la que mejor se ajusta a las observaciones realizadas, de forma muy clara en el Tramo A. Por tanto se propone modificar la definición de NDVI garantizado de tal forma que tenga en cuenta la f.d.p. teórica de la decena en la que se realizan las mediciones.

Para el cálculo del nuevo NDVI garantizado se debería recurrir al uso de percentiles ya que es más apropiado para el tipo de f.d.p. que hemos obtenido, es decir, Gumbel de valores extremos mínimos y GEV.

## 4. CONCLUSIONES

Del análisis espacial del índice NDVI de la zona de estudio podemos concluir que el cálculo del mapa de singularidades puede ser una herramienta útil a la hora de establecer zonas homogéneas para el Seguro Agrario, es decir, zonas en las que la definición del evento sequía se obtenga de la misma f.d.p. (mismos estadísticos). Para la determinación de estas zonas hemos calculado los píxeles cuyo exponente de Lipschitz-Hölder es aproximadamente igual a 2, es decir los píxeles que no son puntos singulares.

La intersección de los mapas de singularidades para los distintos intervalos de los que consta un año formará nuestro punto de partida para el establecimiento de dichas zonas homogéneas. Como el índice NDVI no nos da información sobre el tipo de vegetación que estamos midiendo habrá que determinar por otros medios si la zona homogénea que hemos calculado es de tipo herbáceo, arbustivo, arbóreo o mixto.

Hemos filtrado las series temporales del índice NDVI mediante un criterio de saturación del modelo de color HSL, eliminando los valores de las series que eran anómalos. Del análisis estadístico de las series temporales filtradas del índice NDVI en la zona de pastos hemos observado que las f.d.p. 's para cada uno de los intervalos en los que hemos dividido un año se ajustaban mejor a funciones tipo Gumbel de valores extremos mínimos y GEV. Por lo tanto, si adoptamos estas f.d.p. 's a la hora de definir el siniestro mínimo indemnizable o evento sequía, deberíamos recurrir al uso de percentiles en lugar de desviaciones típicas, ya que estas últimas son utilizadas generalmente cuando la f.d.p. es normal.

Además, hemos identificado cuatro tramos temporales A, B, C, y D que no se corresponden con las estaciones astronómicas del año y que pueden ser descritos por el mismo tipo de f.d.p. (sobre todo el tramo A).

Debido al tiempo limitado del que disponemos para la realización de este TFM hemos dejado para futuras investigaciones el estudio de la evolución temporal de las localizaciones de las singularidades positivas y negativas, y el análisis estadístico de otras zonas de la Comunidad de Madrid.

## 5. REFERENCIAS.

- Agroseguro (2014). Seguro de compensación por pérdida de pastos. *Sistema Español de Seguros Agrarios Combinados*. <http://agroseguro.es/productos/sectores/seguro-pecuario/133-perdida-de-pastos>
- Cheng, Q. (2001). Singularity Analysis For Image Processing And Anomaly Enhancement. *Annual Conference of the International Association for Mathematical Geology*. Cancún.
- Díaz-Ambrona, C. G. H. (1998a). La dehesa, un sistema agro-silvo-pastoral para el tercer milenio. *Agricultura*, 67 (suplemento 16, 789), 1, 3 - 4.
- Díaz-Ambrona, C. G. H. (1998b). *La dehesa, aprovechamiento sostenible de los recursos naturales*. Madrid: Fundación Premio Arce, Caja Madrid y Editorial Agrícola Española.
- Falconer, K. (2003). *Fractal Geometry. Mathematical foundations and applications (2nd edition)*. West Sussex: John Wiley & Sons.
- Feder, J. (1988). *Fractals (physics of solids and liquids)*. New York: Plenum Press.
- Flynn E. S. (2006). *Using NDVI as a pasture management tool*. Master Thesis. University of Kentucky.
- Hayes, M. (2004). Comparison of Major Drought Indices. *National Drought Mitigation Centre*. <http://drought.unl.edu/Planning/Monitoring/ComparisonofIndicesIntro.aspx>.
- Jordan, W. R., Miller, F. R. (1980). Genetic variability in sorghum root systems: implication for drought tolerance. In N. C. Turner and P. J. Kramer (Eds.). *Adaptation of Plants to Water and High Temperature Stresses* (pp. 383-399). New York: Wiley Interscience.
- Larson, H. J. (1982). *Introduction to Probability Theory and Statistical Inference (3rd edition)*. New York: John Wiley and Sons.
- Levkowitz, H. (1997). *Color Theory and Modelling for Computer Graphics, Visualization and Multimedia Applications*. Norwell (Massachusetts): Kluwer Academic Publishers.
- Lovejoy, S., A.M. Tarquis, H. Gaonac'h and D. Schertzer (2008). Single and Multiscale remote sensing techniques, multifractals and MODIS derived vegetation and soil moisture. *Vadose Zone Journal*, 7(2), 533-546.
- LP DAAC (Land Processes Distributed Active Archive Center) (2014). Surface Reflectance 8-Day L3 Global 500m. *NASA and USGS*. [https://lpdaac.usgs.gov/products/modis\\_products\\_table/mod09a1](https://lpdaac.usgs.gov/products/modis_products_table/mod09a1)
- Milne, A.E., M.T. Castellanos, M.C. Cartagena, A.M. Tarquis y R.M. Lark (2010). Investigating the effect of historical treatments on wheat yield over multiple spatial frequencies. *Biogeosciences*, 7, 2739-2747.
- San Miguel Ayanz, A. (2001). *Pastos naturales españoles*. Madrid: Fundación Conde del Valle de Salazar y Ediciones Mundi-Prensa.

# APÉNDICE A

## 1) RESULTADOS ADICIONALES ESPECTRO MULTIFRACTAL

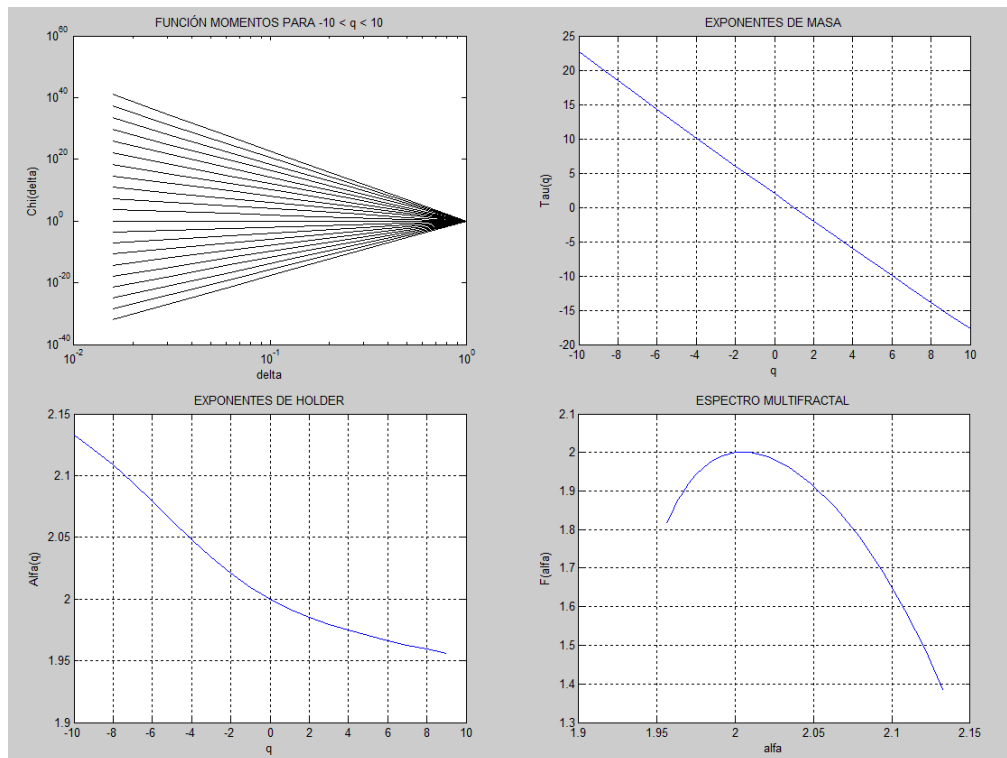


Figura 25. Resultados del análisis espacial: función momentos, exponentes de masa, exponentes de Hölder y espectro multifractal de la imagen de la zona de estudio con fecha 15/04/2011.

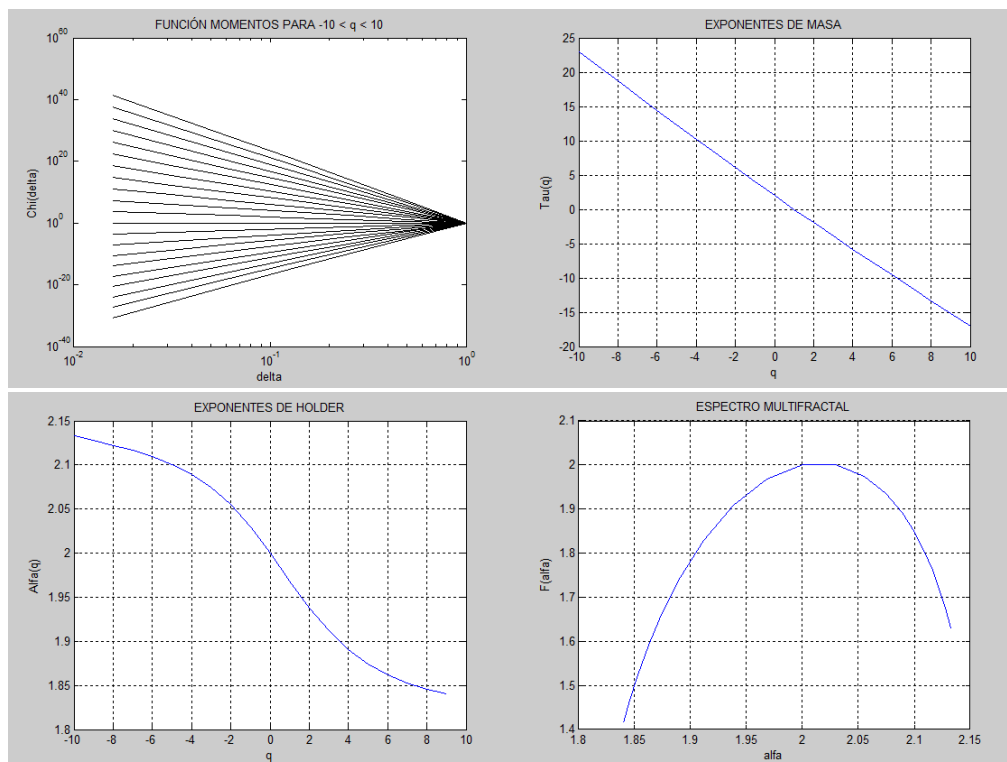
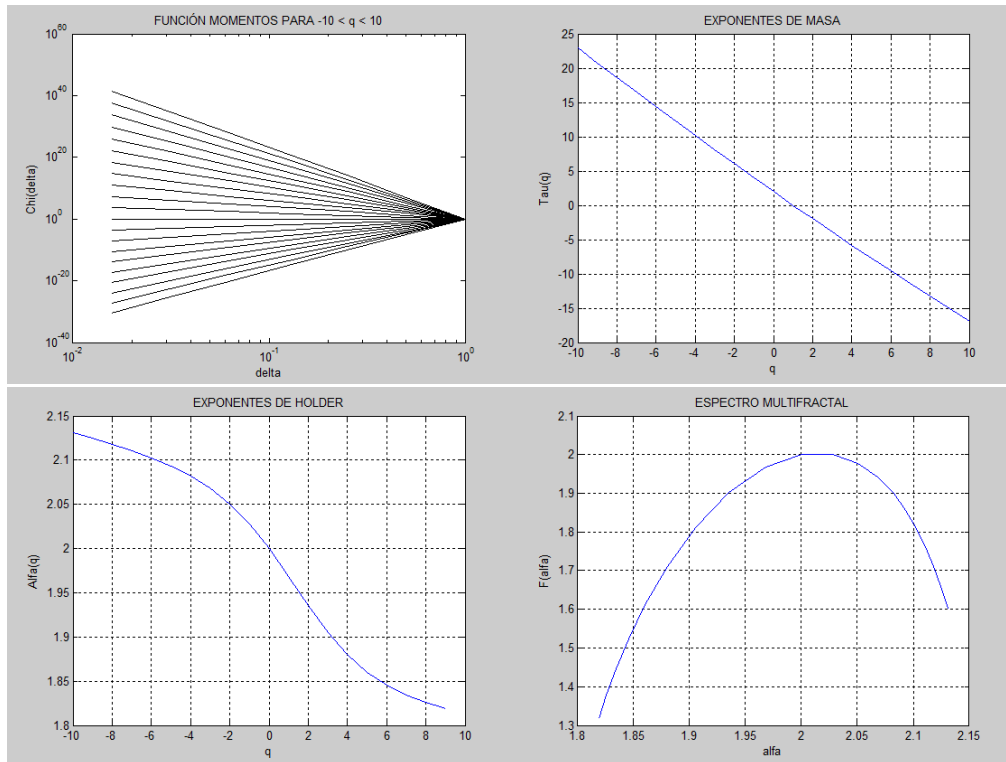
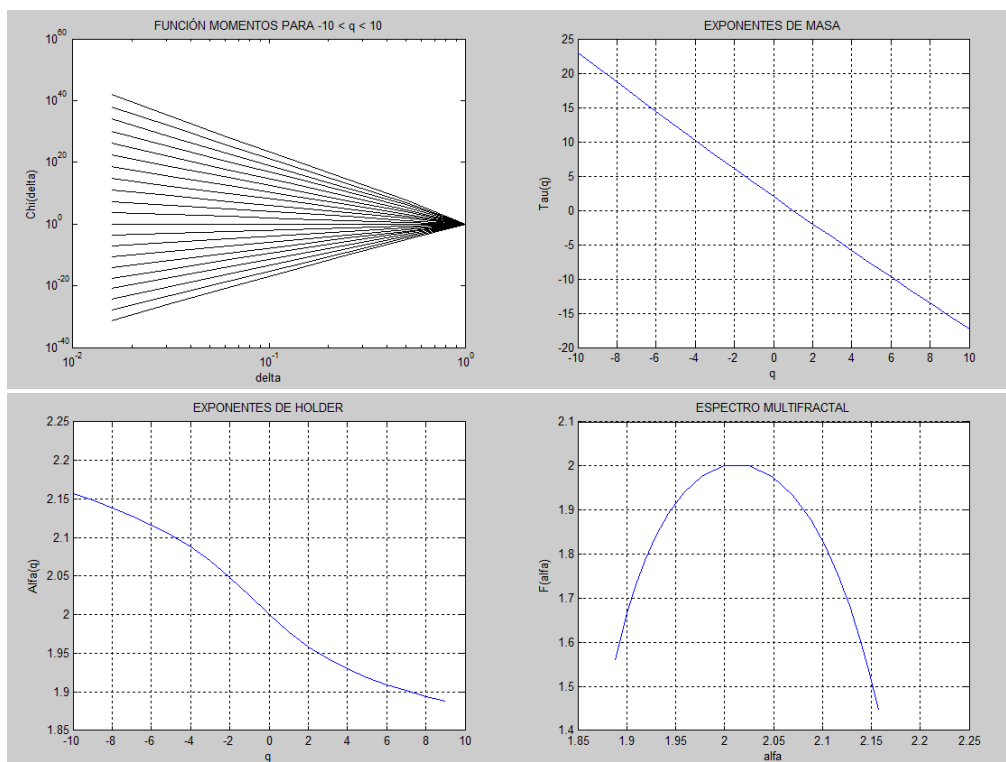


Figura 26. Resultados del análisis espacial: función momentos, exponentes de masa, exponentes de Hölder y espectro multifractal de la imagen de la zona de estudio con fecha 12/07/2011.

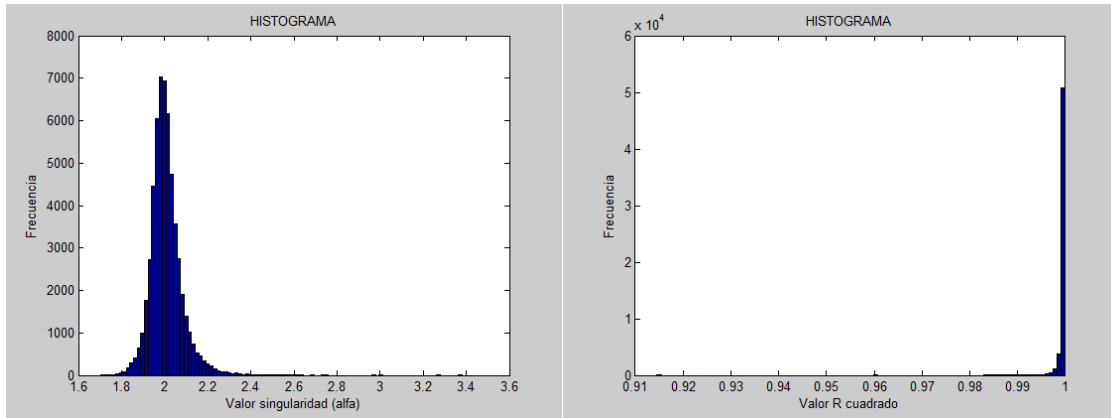


**Figura 27.** Resultados del análisis espacial: función momentos, exponentes de masa, exponentes de Hölder y espectro multifractal de la imagen de la zona de estudio con fecha 16/10/2011.

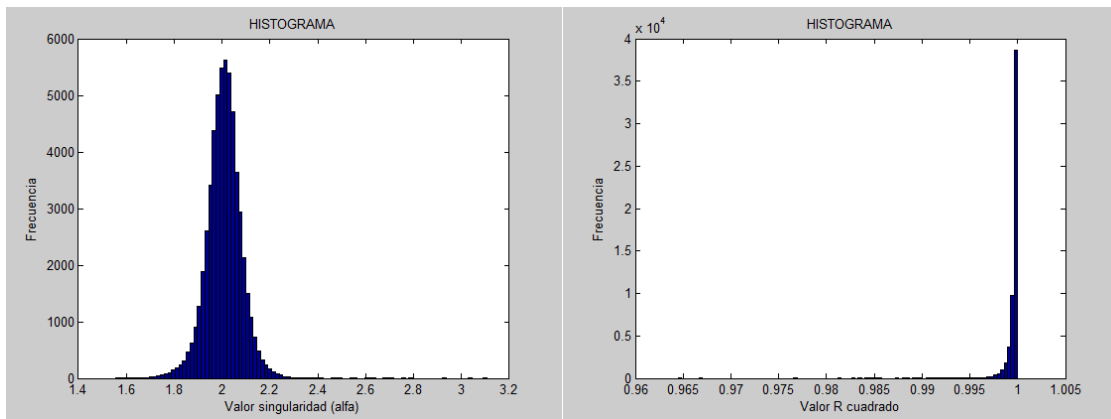


**Figura 28.** Resultados del análisis espacial: función momentos, exponentes de masa, exponentes de Hölder y espectro multifractal de la imagen de la zona de estudio con fecha 01/01/2012.

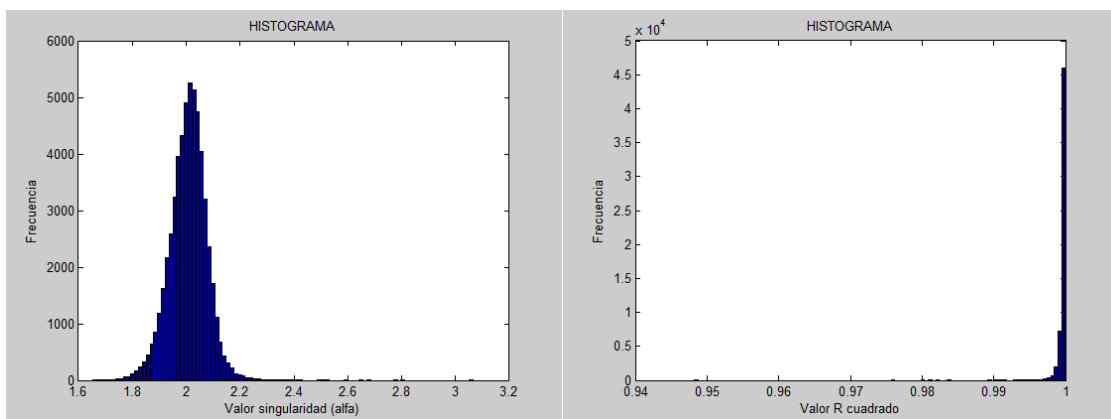
## 2) RESULTADOS ADICIONALES MAPA DE SINGULARIDADES



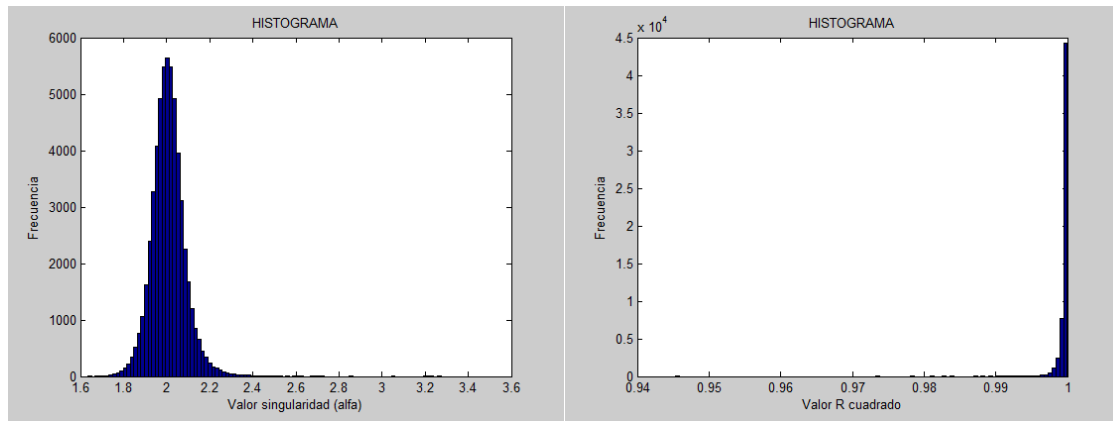
**Figura 29.** Resultados del análisis espacial: histograma de valores de singularidad y  $R^2$  de la imagen de la zona de estudio con fecha 15/04/2011.



**Figura 30.** Resultados del análisis espacial: histograma de valores de singularidad y  $R^2$  de la imagen de la zona de estudio con fecha 12/07/2011.



**Figura 31.** Resultados del análisis espacial: histograma de valores de singularidad y  $R^2$  de la imagen de la zona de estudio con fecha 16/10/2011.



**Figura 32.** Resultados del análisis espacial: histograma de valores de singularidad y  $R^2$  de la imagen de la zona de estudio con fecha 01/01/2012.

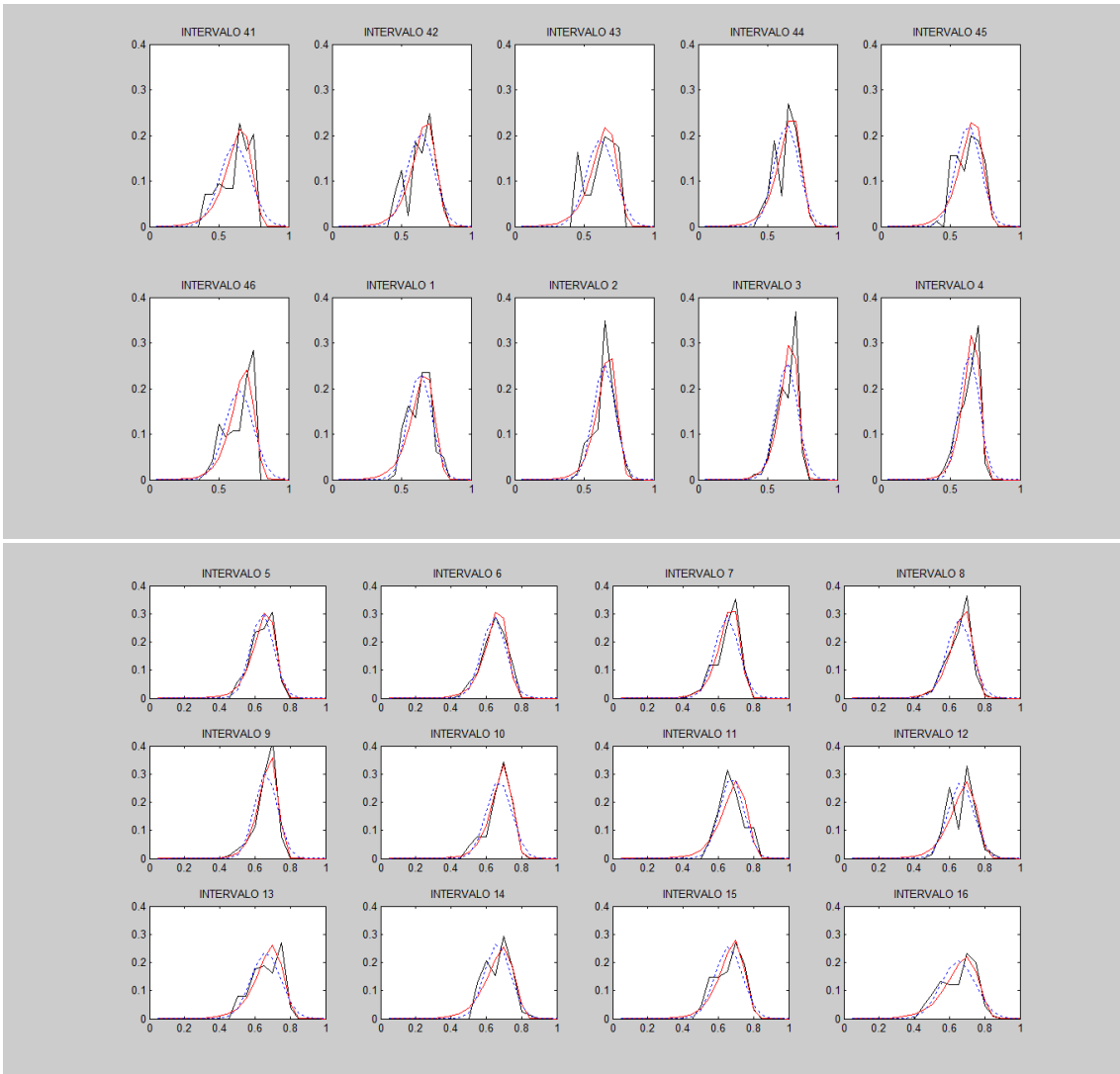
# APÉNDICE B

## 1) RESULTADOS ADICIONALES DEL ANÁLISIS ESTADÍSTICO

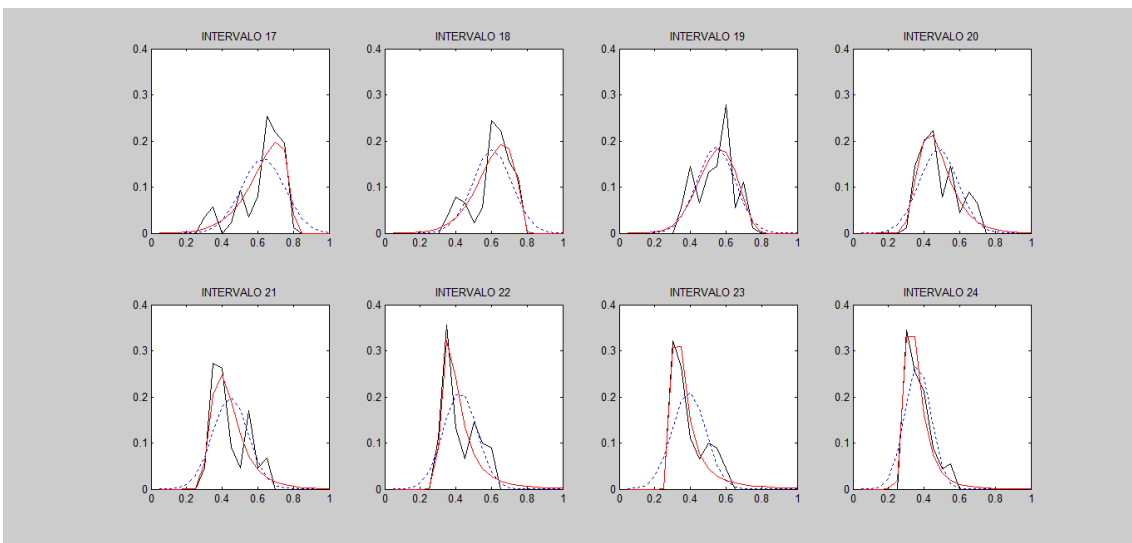
INTERVALO (Nº V.A.)	NORMAL		GAMMA		BETA		GUMBEL (min)		GEV		
	$\mu$	$\sigma$	$\alpha$	$\beta$	a	b	$\mu$	$\sigma$	$\mu$	$\sigma$	$\xi$
1	0,607	0,084	51,95	0,012	20,55	13,28	0,648	0,077	0,581	0,085	-0,354
2	0,621	0,077	60,98	0,010	24,84	15,21	0,657	0,064	0,601	0,083	-0,490
3	0,612	0,076	59,37	0,010	25,70	16,31	0,647	0,058	0,598	0,082	-0,599
4	0,611	0,070	73,19	0,008	30,72	19,55	0,643	0,054	0,596	0,075	-0,532
5	0,616	0,064	90,63	0,007	36,32	22,70	0,646	0,056	0,597	0,067	-0,401
6	0,618	0,067	81,42	0,008	33,21	20,54	0,649	0,054	0,604	0,073	-0,555
7	0,624	0,067	83,74	0,008	34,06	20,55	0,655	0,052	0,612	0,073	-0,621
8	0,630	0,069	78,88	0,008	31,73	18,66	0,662	0,056	0,614	0,074	-0,505
9	0,637	0,063	93,93	0,007	38,50	21,99	0,665	0,048	0,624	0,068	-0,571
10	0,648	0,070	80,08	0,008	31,26	17,05	0,679	0,053	0,634	0,076	-0,601
11	0,651	0,066	98,53	0,007	33,19	17,81	0,684	0,065	0,627	0,063	-0,246
12	0,640	0,072	78,35	0,008	28,42	15,98	0,675	0,066	0,616	0,073	-0,326
13	0,637	0,082	58,25	0,011	22,03	12,56	0,676	0,069	0,619	0,090	-0,556
14	0,639	0,072	78,14	0,008	27,61	15,62	0,675	0,071	0,612	0,070	-0,250
15	0,634	0,076	68,30	0,009	25,74	14,85	0,671	0,064	0,615	0,081	-0,490
16	0,624	0,096	40,40	0,016	15,82	9,53	0,671	0,082	0,597	0,101	-0,420
17	0,604	0,122	19,26	0,031	9,72	6,46	0,656	0,082	0,586	0,127	-0,694
18	0,579	0,109	24,62	0,024	11,79	8,62	0,629	0,085	0,558	0,119	-0,600
19	0,519	0,105	23,13	0,022	11,58	10,76	0,570	0,094	0,488	0,109	-0,406
20	0,457	0,107	19,19	0,024	9,80	11,61	0,513	0,112	0,405	0,083	0,031
21	0,421	0,099	19,20	0,022	10,58	14,52	0,473	0,102	0,370	0,072	0,119
22	0,399	0,094	19,41	0,021	11,17	16,79	0,448	0,095	0,343	0,057	0,368
23	0,366	0,094	17,00	0,022	10,12	17,47	0,416	0,101	0,308	0,048	0,557
24	0,340	0,072	25,13	0,014	15,55	30,14	0,379	0,083	0,302	0,045	0,241
25	0,306	0,058	31,83	0,010	20,87	47,17	0,338	0,069	0,276	0,034	0,284
26	0,295	0,041	54,89	0,005	37,79	90,14	0,317	0,046	0,277	0,033	-0,030
27	0,284	0,034	76,25	0,004	53,48	134,88	0,302	0,038	0,269	0,027	-0,012
28	0,277	0,029	90,71	0,003	65,62	171,26	0,292	0,027	0,267	0,029	-0,327
29	0,274	0,022	160,73	0,002	115,61	307,07	0,285	0,023	0,263	0,017	0,039
30	0,270	0,021	165,29	0,002	119,47	323,42	0,281	0,023	0,260	0,018	-0,041
31	0,271	0,025	121,84	0,002	88,16	237,60	0,283	0,027	0,260	0,021	-0,079
32	0,280	0,033	80,13	0,004	55,58	142,75	0,298	0,044	0,264	0,022	0,123
33	0,279	0,040	52,16	0,005	36,25	93,75	0,301	0,053	0,262	0,034	-0,077
34	0,303	0,064	26,11	0,012	16,86	38,77	0,338	0,079	0,268	0,027	0,488
35	0,323	0,086	17,09	0,019	10,27	21,41	0,371	0,114	0,277	0,043	0,394
36	0,399	0,146	7,77	0,051	4,45	6,66	0,475	0,152	0,324	0,108	0,112
37	0,493	0,172	7,35	0,067	3,98	4,10	0,577	0,150	0,449	0,183	-0,478
38	0,532	0,141	13,18	0,040	6,50	5,74	0,600	0,123	0,495	0,149	-0,465
39	0,583	0,115	24,27	0,024	10,60	7,60	0,638	0,098	0,559	0,127	-0,574
40	0,615	0,087	46,50	0,013	19,39	12,18	0,655	0,070	0,600	0,097	-0,641
41	0,592	0,110	26,71	0,022	12,08	8,36	0,643	0,084	0,582	0,125	-0,803
42	0,614	0,096	38,56	0,016	16,09	10,16	0,658	0,078	0,595	0,105	-0,589
43	0,594	0,103	31,19	0,019	13,66	9,35	0,643	0,082	0,577	0,114	-0,658
44	0,611	0,088	47,00	0,013	19,13	12,19	0,653	0,074	0,590	0,094	-0,500
45	0,605	0,088	45,75	0,013	18,59	12,17	0,647	0,077	0,582	0,094	-0,470
46	0,620	0,101	35,01	0,018	14,86	9,13	0,666	0,074	0,622	0,124	-1,097

Tabla 4. Parámetros obtenidos por MV de las 5 f.d.p. ´s candidatas en cada uno de los intervalos en los que se divide un año.

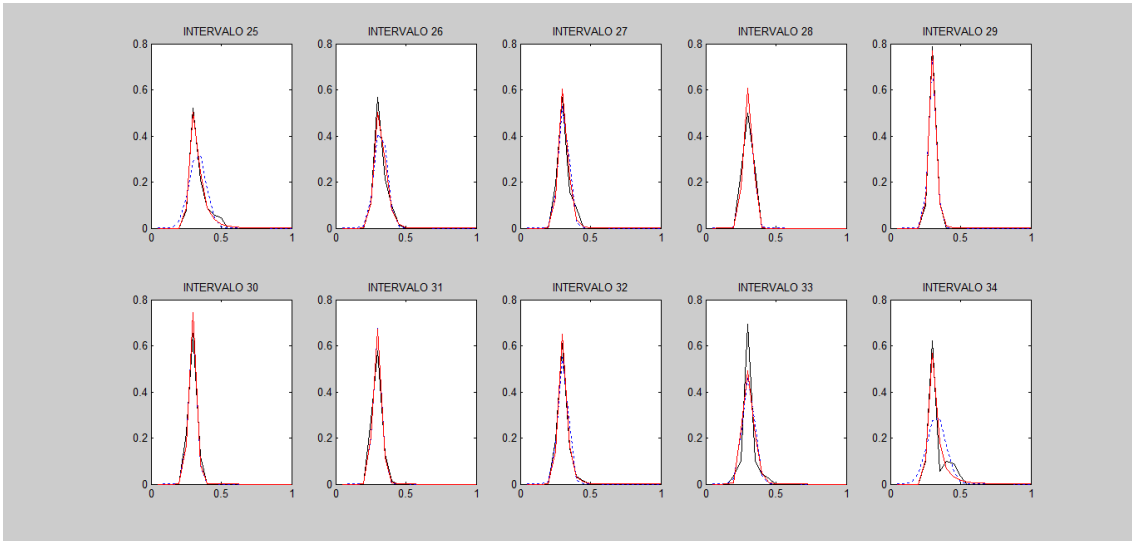




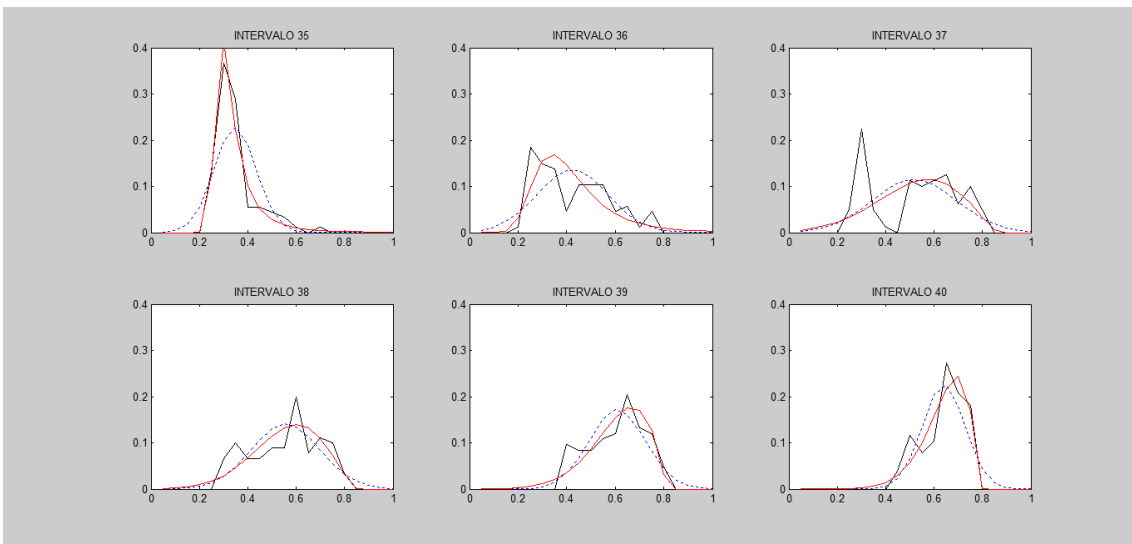
**Figura 33.** F.d.p. observada, Gumbel y normal para todos los intervalo del Tramo A.



**Figura 34.** F.d.p. observada, GEV y normal para todos los intervalo del Tramo B.



**Figura 35.** F.d.p. observada, GEV y normal para todos los intervalos del Tramo C.



**Figura 36.** F.d.p. observada, GEV y normal para todos los intervalos del Tramo D.