



Artículo Original

# Análisis del riesgo de incendios forestales: Un enfoque basado en procesos puntuales

Forest wildfire hazard analyses: A point processes approach

Rafael González de Gouveia<sup>a</sup>, Abelardo Monsalve Cobis<sup>a</sup>

<sup>a</sup>Universidad Centroccidental Lisandro Alvarado. Barquisimeto, Venezuela.

Recibido: 16-03-2017

Aceptado: 15-10-2017

## Resumen

Los procesos estocásticos puntuales representan una herramienta de gran utilidad para el análisis de los factores de riesgo en los incendios forestales. En este artículo se estudia la ocurrencia de los incendios forestales a partir de un proceso de Poisson espacio temporal, en el que se considera la función de intensidad del mismo como una caracterización del riesgo de incendio a partir de técnicas paramétricas y no paramétricas. Finalmente, se considera un conjunto de datos reales, suministrados por el Ministerio del Poder Popular para el Ambiente a través del Instituto Nacional de Meteorología e Hidrología (INAMEH), relativos a los incendios forestales producidos en un día en particular. Se estiman las funciones de riesgo basadas en el modelo propuesto y se generan mapas de riesgo de incendios lo cuales se ajustan a las características geográficas y climáticas de Venezuela.

**Palabras clave:** Procesos puntuales, métodos no paramétricos, procesos de Poisson no homogéneos  
*Código UNESCO: 1209.11*

## Abstract

Point stochastic processes represent a very useful tool for the analysis of hazard factors in wildfire. In this article, the occurrence of wildfire is studied from a spatial-temporal poisson process, in which the intensity function thereof is considered as a wildfire hazard characterization based on parametric and non-parametric techniques. Finally, it is considered a set of real data, provided by the Ministerio del Poder Popular para el Ambiente from Instituto Nacional de Meteorología e Hidrología (INAMEH), relating to wildfire produced on a particular day. The hazard functions are estimated based on the proposed model and wildfire hazard maps are generated, which are adjusted to the geographical and climatic characteristics of Venezuela.

**Key words:** Point process, non-parametrics methods, non-homogeneous Poisson process.  
*UNESCO Code: 1209.11*

## 1. Introducción

En ocasiones es indispensable saber el lugar donde hay mayor riesgo de que se sucedan eventos de interés, para que de forma previsible puedan tomarse las medidas necesarias para evitarlos o al menos minimizar el impacto de sus consecuencias. Es el caso de los incendios forestales, que representan un problema que afecta directa e indirectamente a muchos sectores de la sociedad. Desde el punto de vista ecológico, causan destrucción de la fauna y flora silvestre, además de la contaminación del medio ambiente; y desde el punto de vista económico causan pérdidas cuantiosas en la actividad maderera y sus derivados forestales.

Saber la posible ubicación de los incendios forestales nos permite obtener información acerca de la frecuencia con la que algún incendio forestal ocurre, y también la forma en que ellos producen para obtener patrones y tendencias de su comportamiento. Si agregamos marcas a estas ubicaciones, por ejemplo el área quemada, o la magnitud del incendio forestal, podremos hacer un análisis más riguroso.

Por otro lado, el riesgo de que se produzca un incendio puede ser modelizado como un proceso puntual espacio-temporal  $N$ , en el que se especifica una forma funcional para la intensidad  $\lambda(t, x, y)$ , la cual representa la tasa esperada de eventos (en nuestro caso incendios forestales) por unidad de tiempo y unidades de área, a partir de las observaciones hasta el tiempo  $t$ . Peng et al (2005) [1] proponen estimar la función de riesgo de incendios mediante un proceso de Poisson espacio temporal basados en técnicas paramétricas y no paramétricas. Los autores hacen uso del índice de combustión y evalúan su efectividad en la predicción de las ocurrencias de incendios. Xu y Shoenberg (2011) [2], por su parte, analizan el poder predictivo de un conjunto de modelos de procesos puntuales más complejos que usan información de tendencias de incendios forestales, variables meteorológicas y patrones espaciales históricos de las áreas afectadas por los incendios, como covariables, y que interpolan los registros de diferentes condiciones climáticas. La función de intensidad o riesgo de incendio representa el elemento principal en la caracterización del proceso de Poisson subyacente aportando así información acerca de la probabilidad de ocurrencia de los incendios en una determinada ubicación. En este trabajo se pretende estimar la cantidad de incendios forestales que ocurren por unidad de área y por unidad de tiempo y en consecuencia generar un mapa de riesgo de ocurrencia de estos basados principalmente en los trabajos propuestos por Xu y Shoenberg (2011)[2] y Peng et al (2005) [1].

El trabajo se presenta de la siguiente manera: La sección 2, describe las características principales de los modelos puntuales espacio-temporales y sus premisas. En la sección 3, se presentan los modelos considerados en este trabajo para abordar la caracterización de la ocurrencia de incendios. En la sección 4, se discuten los resultados de la aplicación a datos reales de incendios en Venezuela de los modelos propuestos en la sección 3. Finalmente se presentan unas breves conclusiones.

## 2. Procesos puntuales espacio-temporales

El proceso que abordaremos en este trabajo es el Proceso de Poisson espacio temporal ya que este modelo nos permite caracterizar la ocurrencia de los incendios forestales. Por tanto, a continuación se presentarán algunas definiciones básicas necesarias para el desarrollo de este trabajo.

Un proceso estocástico  $\{N(t), t \geq 0\}$  es llamado un **proceso de conteo** si  $N(t)$  representa el número total de eventos que ocurren hasta el tiempo  $t$ . En este caso con evento nos referimos a un suceso o a la ocurrencia del fenómeno de interés. Por ejemplo, si  $N(t)$  representa el número de personas que entran en un comercio determinado hasta un tiempo  $t$ , entonces  $\{N(t), t \geq 0\}$  es un proceso de conteo en el cual un evento corresponde a una persona entrando en dicha tienda. Un proceso de conteo  $\{N(t), t \geq 0\}$  debe además satisfacer:

- $N(t) \geq 0$ .
- $N(t)$  solo toma valores enteros, es decir,  $N(t)$  es una variable aleatoria discreta.
- Si  $s < t$ , entonces  $N(s) < N(t)$ .
- Para  $s < t$ ,  $N(t) - N(s)$  es igual al número de eventos en el intervalo  $(s, t]$ .

Un proceso de conteo  $\{N(t), t \geq 0\}$  se dice que es un **proceso de Poisson** con tasa constante  $\lambda > 0$  o que es homogéneo cuando:

1.  $N(0) = 0$ .
2. El proceso tiene incrementos independientes.
3. El número de eventos en cualquier intervalo de la forma  $(s, s + t)$  tiene una función de distribución de Poisson con tasa  $\lambda t$ . Esto es, para todo  $s, t \geq 0$

$$P\{N(t + s) - N(s) = n\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad \text{para } n = 0, 1, \dots$$

Una generalización del proceso de Poisson, es el proceso de Poisson no homogéneo, que se caracteriza porque  $\lambda$  es una función que depende del índice  $t \in T$ . El proceso de conteo  $\{N(t), t \geq 0\}$  es un **proceso de Poisson no homogéneo** con intensidad  $\lambda(t), t \geq 0$ , si

1.  $N(0) = 0$ .
2.  $\{N(t), t \geq 0\}$  tiene incrementos independientes.
3.  $P\{N(t+h) - N(t) = 1\} = \lambda(t)h + o(h)$ .
4.  $P\{N(t+h) - N(t) \geq 2\} = o(h)$ .

Este proceso tiene incrementos independientes, mas no incrementos estacionarios, es decir, si tomamos incrementos de la misma longitud en distintos puntos de la recta real no tienen porque seguir una misma distribución, un ejemplo práctico es el estudio de los tiempos de llegada de clientes en un restaurante, donde la intensidad de clientes en la mañana no es la misma que al medio día o en la noche. Matemáticamente esto se expresa a través de la función de intensidad  $\lambda(t)$  que varía en el tiempo.

Teniendo claro las características principales de un proceso de Poisson, estamos listos para dar la definición del proceso que será la base de estudio de este artículo. Así pues, un **proceso puntual espacio temporal** es una colección aleatoria de puntos, donde cada punto representa el tiempo y la ubicación de un evento, ver [1]. Un ejemplo de estos eventos incluyen incidencia de enfermedades, observación o nacimiento de especies animales y vegetales, ocurrencia de incendios, terremotos, rayos, tsunamis, o erupciones volcánicas. Usualmente las ubicaciones espaciales son registradas en tres coordenadas en nuestro caso se trabaja solo con dos coordenadas espaciales, la longitud y la latitud. En este orden de ideas, un **proceso puntual espacio-temporal**  $N$  se define entonces como una medida aleatoria en una región  $S \subset \mathbb{R} \times \mathbb{R}^2$  de espacio-tiempo, tomando valores en los enteros no negativos [1]. En este contexto la medida  $N(A)$  representa el número de puntos en un subconjunto  $A$  de  $S$ . Entonces, el estudio se restringe casi siempre a un conjunto de puntos en un intervalo de tiempo  $[T_0, T_1]$ , y a procesos con solo una cantidad finita de puntos en algún subconjunto compacto de  $S$ . De la misma manera que ocurre con los procesos puntuales descritos inicialmente, cualquier proceso puntual espacio temporal analítico está caracterizado únicamente por su **tasa condicional** asociada  $\lambda$ , ver [3], [2]. En el caso del proceso espacio-temporal, esta tasa asociada  $\lambda(t, x, y)$ , puede ser vista como la frecuencia con la que se espera que los eventos ocurran alrededor de un ubicación específica  $(t, x, y)$  en el espacio-tiempo, condicionada con el historial a priori,  $H_t$ , del proceso puntual hasta el tiempo  $t$ .

Formalmente, la tasa condicional  $\lambda(t, x, y)$  asociada con un proceso puntual espacio temporal  $N$  puede ser definida como una esperanza condicional límite. Fijado cualquier punto  $p = (t, x, y)$  en el espacio-tiempo. Sea  $B_\Delta$  el conjunto  $[t, t + \Delta t] \times [x, x + \Delta x] \times [y, y + \Delta y]$ , y sea  $\Delta$  el vector  $(\Delta t, \Delta x, \Delta y)$ . Entonces

$$\lambda(p) = \lim_{\delta \rightarrow 0} \frac{E[N(B_\Delta) | H_t]}{|\Delta|},$$

siempre que este limite exista.

### 2.1. Proceso de Poisson para el riesgo de incendios

La función de riesgo de incendio es la frecuencia de incendios ocurridos por unidad de área y por unidad de tiempo. Las ocurrencias de incendios pueden verse como observaciones de un proceso espacio temporal de Poisson no homogéneo, con intensidad  $\lambda(t, x, y)$ .

Dada una forma paramétrica para  $\lambda(t, x, y; \theta)$ , los estimados del parámetro  $\theta$  pueden ser obtenidos maximizando la función de log-verosimilitud [2]. La cuál para un modelo espacio temporal tiene la forma

$$L(\theta) = \sum_{i=1}^n \ln \lambda(t_i, x_i, y_i; \theta) - \int_{T_0}^{T_1} \int_x \int_y \lambda(t, x, y; \theta) dy dx dt.$$

Para más detalle sobre esta forma de la verosimilitud ver [?] . En el caso de un proceso de Poisson, la intuición detrás de esta fórmula es que  $\sum_{i=1}^n \ln \lambda(t_i, x_i, y_i; \theta)$  refleja la verosimilitud asociada a los eventos, y  $\int_{T_0}^{T_1} \int_x \int_y \lambda(t, x, y; \theta) dy dx dt$  representa la probabilidad de que no ocurran eventos en ninguna otra porción de la región espacio temporal.

### 3. Los Modelos

Se pretende caracterizar la ocurrencia de los incendios a partir de un proceso Poisson como el descrito en la sección anterior. Así pues, el problema fundamental se centra en estimar la cantidad de incendios forestales por unidad de área y por unidad de tiempo. A dicha cantidad se le denomina riesgo de incendio, y se puede modelizar mediante un proceso de Poisson homogéneo o bien no homogéneo. A continuación se describen los modelos considerados en este trabajo:

#### 3.1. Modelo homogéneo

Se considera un primer modelo de referencia, el proceso puntual espacio-temporal de Poisson homogéneo, cuya función de riesgo o intensidad viene dada por la expresión:

$$\lambda(t, y, x) = \mu,$$

donde  $\mu$  es un valor constante [1]. Este tipo de proceso puntual homogéneo suele llamarse modelo de **aleatoriedad espacial completa** o **CSR** (por sus siglas en inglés, *complete spatial randomness*). Es importante recalcar que es solo un modelo para comparar con otros, ya que este modelo sugiere que los datos están distribuidos de manera uniforme sobre la superficie y en muchas situaciones, incluida la relativa a incendios, esto no ocurre.

#### 3.2. Modelo no homogéneo

En un segundo modelo, la función de riesgo de incendio a considerar se caracteriza por ser más realista. En este caso hemos considerado un proceso de Poisson no homogéneo, donde la intensidad condicional en el tiempo  $t$  y en la ubicación  $(x, y)$  dependa solamente de la temporada asociada con  $t$ , así como la tasa del espectro espacial  $m(x, y)$  de incendios ocurridos en esta zona [1]. Esto es, se considerará como el modelo:

$$\lambda_1(t, x, y) = \gamma m(x, y) + \alpha S(t),$$

donde  $m(x, y)$  es una función para los datos espaciales, llamada espectro o patrón espacial, y  $S(t)$  es una función asociada a los datos temporales, llamada patrón temporal. Aquí  $\gamma$  y  $\alpha$  son parámetros reales a ser estimados en el ajuste del modelo.

Para describir el espectro espacial  $m(x, y)$ , debemos tomar en cuenta la inhomogeneidad espacial de la ocurrencia de incendios (ver [1]). A esta característica se incorpora el conocimiento previo que tenemos de los incendios. Se propone el uso de una estimación no paramétrica, esto es,

$$m(x, y) = \frac{1}{n_0} \sum_{j=1}^{n_0} \frac{1}{\phi_x \phi_y} K\left(\frac{x - x_j}{\phi_x}\right) K\left(\frac{y - y_j}{\phi_y}\right),$$

donde  $K$  es una función kernel o de núcleo bidimensional,  $n_0$  es el número de datos,  $x_j$  es la longitud del  $i$ -ésimo dato,  $y_j$  es la latitud del  $i$ -ésimo dato,  $\phi_x$  es el ancho de banda correspondiente a la longitud geográfica y  $\phi_y$  es el ancho de banda correspondiente a la latitud geográfica. Es bien conocido que los métodos no-paramétricos son sensibles a la selección del ancho de banda más allá de la propia selección de la función kernel [4]. Para obtener una selección adecuada de dicho parámetro, vital en la estimación no-paramétrica se hará uso del método de validación cruzada

Respecto de la componente temporal  $S(t)$ , la cual es utilizada para describir el cambio a lo largo del tiempo de la ocurrencia de incendios forestales (ver [1]). Se procederá de forma similar que en el caso anterior. Así pues, la expresión, usando técnicas no-paramétricas para obtener dicha función es: que

$$S(t) = \frac{1}{n_0} \sum_{j=1}^{n_0} \frac{1}{\sigma_t} K\left(\frac{t^* - t_j^*}{\sigma_t}\right),$$

en donde  $K$  es una función kernel o de núcleo,  $n_0$  es el número de datos,  $t^*$  representa el tiempo desde el inicio del día y  $t_j^*$  es el tiempo desde el inicio del día para el  $j$ -ésimo incendio y  $\sigma_t$  es el ancho de banda a ser estimado.

Similar al caso del espectro, la selección del ancho de banda, se hará usando el método de validación cruzada. A continuación se describen los datos que serán analizados siguiendo los modelos antes descritos.

#### 4. Aplicaciones y resultados

Para esta investigación, se utilizaron 220 datos de incendios forestales, suministrados por el Ministerio del Poder Popular para el Ambiente a través del Instituto Nacional de Meteorología e Hidrología (INAMEH), recolectados por esta institución el 31 de marzo de 2014 [5]. Estos datos constan de la latitud y longitud de los focos de incendios forestales detectados en todo el territorio nacional a través del satélite de órbita polar AQUA, específicamente usando el sensor MODIS. Este satélite produce un conjunto de datos por día que indican la presencia de fuego con aproximadamente 30 metros de extensión por 1 metro de ancho.

Los datos están distribuidos sobre todo el territorio nacional por lo que estamos hablando de alrededor de 12 grados de latitud, que para nuestra posición equivalen a 1300 km aproximadamente, y 14 grados de longitud, que en nuestra región equivale a casi 1500 km. Dada la magnitud de estas medidas si pretendiésemos hacer un análisis estricto del riesgo de incendio forestal en esta área, tendríamos que tomar en cuenta la curvatura de la Tierra (Figura 1, izquierda), sin embargo, los modelos están contemplados para regiones planas que en la práctica son pequeñas en comparación al tamaño de nuestro país, como bosques naturales o regiones con menos de un grado de latitud o de longitud, en donde podemos suponer sin perder mucho detalle una geometría totalmente plana.

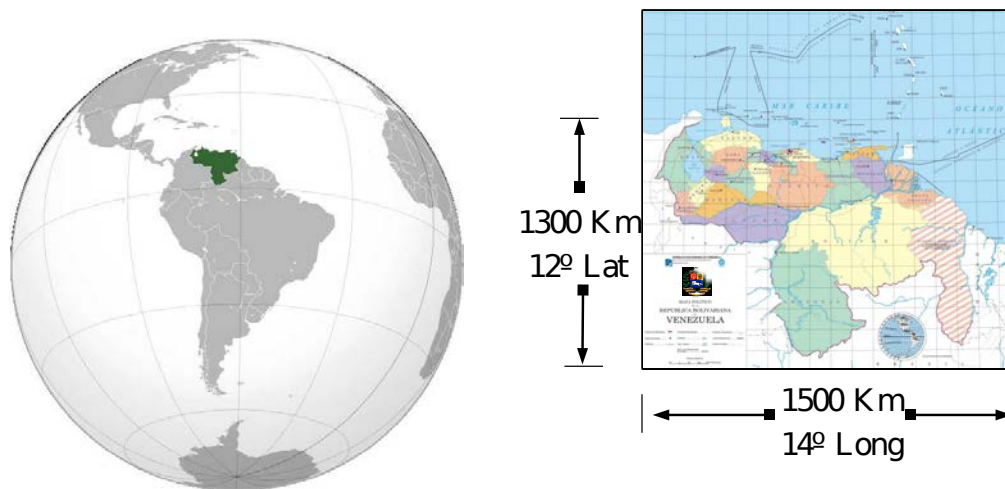


Figura 1. Mapa de Venezuela.

Supondremos que sobre el área completa del país, tanto la latitud como la longitud son coordenadas planas. Dado que Venezuela está cercana al ecuador, los grados de latitud y de longitud son similares. Asignaremos el valor de 108 km a un 1 grado de latitud o de longitud, y llamaremos a esta cantidad **unidad espacial**. Este valor parte del hecho de que en 12 grados de latitud hay 1300 km, y en 14 grados de longitud hay 1500 km, la proporción de los kilómetros entre los grados se aproxima en ambos casos a 108 km por grado (Figura 1, derecha). Este valor será de utilidad más adelante para entender los resultados obtenidos.

Se simularon de manera aleatoria los tiempos de ocurrencia de incendios forestales en un día, siguiendo una distribución normal con media 12 (representativa al medio día) y desviación estándar de 1 hora. En la Figura 2 se muestra un histograma de los tiempos utilizados.

En la Figura 3 se muestra una gráfica de los datos espaciales y en el cuadro 1 se extraen parte de los datos utilizados y se muestra el tiempo simulado (en horas) como una columna adicional agregada a los incendios ubicados en las coordenadas de longitud  $x$  y latitud  $y$ .

Para el patrón temporal, el ancho de banda se encuentra minimizado la función de pérdida:

$$\hat{J}(h) = \int (\hat{f}_n(x))^2 dx - \frac{2}{n} \sum_{i=1}^n \hat{f}_{(-i)}(x_i),$$

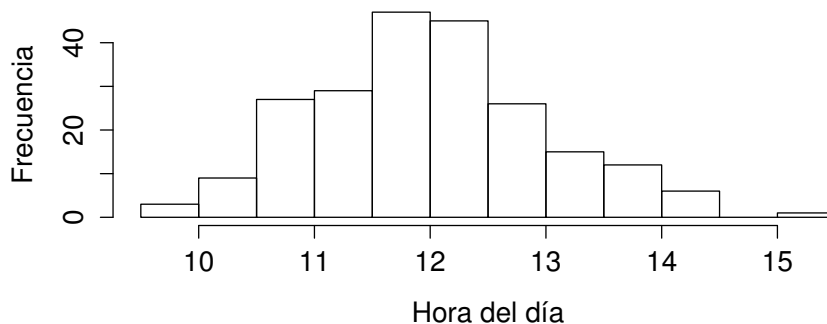


Figura 2. Histograma de ocurrencia de incendios.

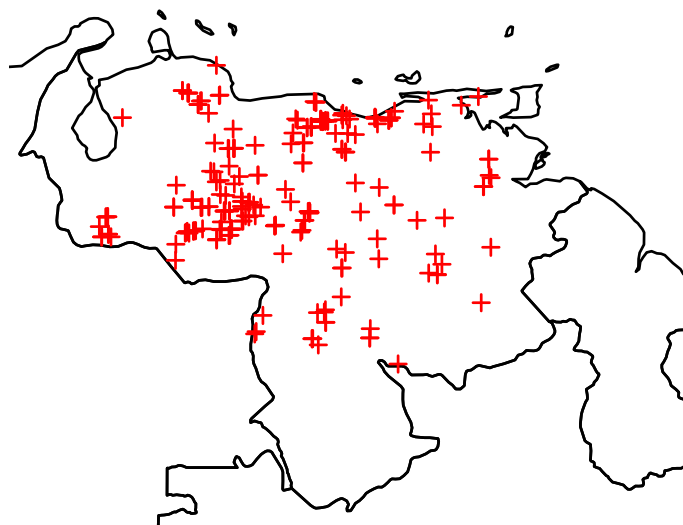


Figura 3. Datos suministrados por el INAMEH para el 31 de marzo de 2014[5].

El ancho de banda resultante es  $\sigma_t = 0,3292192$  horas. Este valor está medido en horas y representa el parámetro de suavidad con que se grafica la densidad.

Haciendo uso del paquete estadístico R, y su librería ks, la cual proporciona algoritmos para encontrar los anchos de banda óptimos mediante la validación cruzada sesgada o BCV, se obtuvieron los siguientes resultados:

$$\phi_x = 0,9, \quad \phi_y = 0,45.$$

Estos valores están medidos en unidades espaciales, donde una unidad espacial equivale a 108 km, por lo que  $\phi_x = 48,6$  km y  $\phi_y = 97,2$  km. La gráfica del espectro espacial, junto con los datos de incendios forestales se muestran en la Figura 4.

El modelo homogéneo estimado da como resultado: de  $\hat{\mu} = 0,05456349 \frac{\text{incendios forestales}}{(\text{unidad espacial})^2 \text{ hora}}$ , esta es la tasa constante de incendios por unidad de área por unidad de tiempo.

En el caso del modelo no-homogéneo, los estimados por máxima verosimilitud para este modelo fueron  $\hat{\gamma} = 1,930265$ ,  $\hat{\alpha} = 1,033772$ , Estas cantidades indican que la intensidad de incendios forestal se ve mayormente afectada por el espectro espacial, es decir, la ubicación.

Por otro lado, los resultados obtenidos para el espectro espacial son muy interesantes en los estados Apure y Miranda, y en menor grado el estado Falcón la intensidad de incendios es alta. Hernández (2014) [5] muestra que estos estados

Tiempo	Lon O	Lat N
11.05	69.36	10.65
12.84	69.33	10.64
12.32	69.33	10.65
10.53	68.60	10.59
13.18	68.52	10.60

Cuadro 1. Parte de los datos utilizados en esta investigación [5].

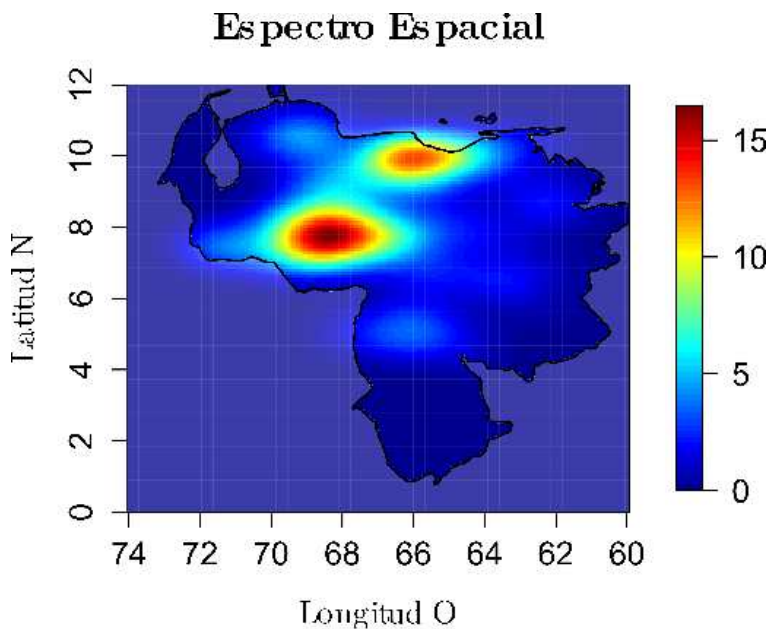


Figura 4. Espectro Espacial para datos de incendios forestales en Venezuela.

presentan la mayor sequía del país para en el periodo en que fueron obtenidos los datos (marzo 2014).

La Figura 5 tomada de la página oficial del INAMEH, a través del trabajo de Hernández (2014) [5] muestra la caracterización de la sequía para el periodo semestral de Febrero a Julio de 2014. La sequía ocurre cuando las precipitaciones son inferiores al promedio histórico registrado en un lugar o una región.

Es importante notar que nuestro modelo solo toma como información la ubicación espacial de los incendios forestales, y la fecha en que estos ocurrieron, no toma en cuenta factores climáticos como la humedad, la temperatura o la precipitación. Pese a esto, las zonas donde nuestro modelo predice mayor intensidad de incendios forestales coinciden con las zonas de sequía vista en la Figura 5.

## 5. Conclusiones

Los resultados obtenidos para el patrón temporal coinciden con las simulaciones realizadas, por lo que el modelo con componentes espacial y temporal se ajusta satisfactoriamente a los datos.

Los procesos puntuales de Poisson espacio temporal, representan una alternativa interesante y de gran potencial en la caracterización de la ocurrencia de incendios en Venezuela dado que corrobora las conclusiones obtenidas por los estudios de sequía y el impacto de estas.

Es importante destacar que se pueden considerar modelos que incorporen información relativa a variables climáticas, tal y como lo propone Xu y Schoenberg (2011) [2].

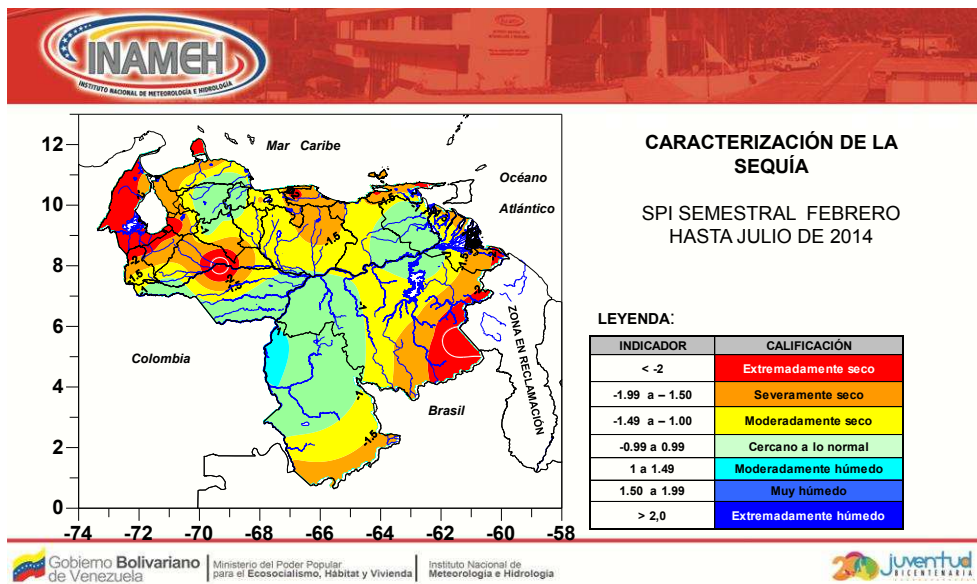


Figura 5. Caracterización de la sequía [5].

## Referencias

- [1] R.D. Peng; F.P. Schoenberg; J.A. Woods. A space-time conditional intensity model for evaluating a wildfire hazard index. *Journal of the American Statistical Association*, 100(469):26–35, 2005.
- [2] H. Xu; F.P. Schoenberg. Point process modeling of wildfire hazard in los angeles county, california. *The Annals of Applied Statistics*, 5(2a):684–704, 2011.
- [3] F.P. Schoenberg; D.R. Brillinger; P.M. Guttorp. Point processes, spatial-temporal. *Encyclopedia Environmetrics*, Abdel El-Shaarawi and Walter Piegorsch, editors. Wiley, 3:1573–1577, 2002.
- [4] B.W. Silverman. Density estimation for statistics and data analysis. *Monographs on Statistics and Applied Probability*, London: Chapman and Hall, pages 1–22, 1986.
- [5] R. Hernández. Caracterización espacial de la sequía meteorológica (spi) a nivel semestral febrero-julio, marzo-agosto y abril-septiembre 2014. *INAMEH - Gerencia de Meteorología Coordinación de Meteorología Aplicada*, 1:2–3, 2014.