



INFERENCIA ESTADÍSTICA EN PROCESOS PUNTUALES ESPACIALES. APLICACIÓN AL ANÁLISIS DE INCENDIOS FORESTALES EN GALICIA

PROYECTO FIN DE MÁSTER

AUTORA:

Isabel Fuentes Santos

DIRECTORES:

Wenceslao González Manteiga

Manuel F. Marey Pérez

Curso 2008-2009

ÍNDICE

INTRODUCCIÓN.....	5
INCENDIOS FORESTALES.....	7
1. ESTUDIO DE LOS INCENDIOS FORESTALES.....	7
1.1. Comportamiento del ciclo de los incendios.....	7
2. DESCRIPCIÓN DE LA BASE DE DATOS.....	9
2.1. Los incendios forestales en el distrito Fonsagrada- Ancares.....	11
2.2. Análisis espacial de los incendios a nivel de parroquia.....	15
2.3. Localización espacial de los fuegos.....	16
PROCESOS PUNTUALES ESPACIALES.....	19
1. ALEATORIEDAD ESPACIAL COMPLETA (CSR).....	19
2. PROPIEDADES DE LOS PROCESOS PUNTUALES.....	20
3. CARACTERÍSTICAS DE PRIMER Y SEGUNDO ORDEN.....	21
3.1. Procesos univariantes.....	21
3.2. Procesos multivariantes.....	25
4. DISTRIBUCIONES DE ESPACIO VACIO Y DEL VECINO MÁS PRÓXIMO.....	26
5. INDEPENDENCIA Y ETIQUETADO ALEATORIO.....	27
TESTS DE CSR.....	29
1. CONTRATOS DE CSR.....	29
1.1. Mosaico de Dirichlet.....	30
1.2. Tests de Monte Carlo.....	30
2. TESTS DE DISTANCIAS.....	32
2.1. Distancia entre eventos.....	32
2.2. Distribución de la distancia al vecino más próximo.....	34
2.3. Distribución de la función de espacio vacío.....	35

3.	TESTS DE CUADRANTES.....	36
3.1.	Cuadrat couts	36
3.2.	Escalas en el diseño.....	37
3.3.	Escalas de dependencia	38
4.	CONTRASTE BASADO EN LA K-FUNCIÓN:.....	38
5.	CONTRASTES DE CSR PARA LOS FUEGOS REGISTRADOS.	39
	MODELOS	45
1.	CSR: LOS PROCESOS HOMOGÉNEOS DE POISSON	45
2.	PROCESOS DE POISSON CON CLÚSTERS	47
3.	PROCESOS INHOMOGÉNEOS DE POISSON.....	49
4.	PROCESOS DE COX	50
5.	PROCESOS DE INHIBICIÓN SIMPLE.....	53
6.	PROCESOS PUNTUALES DE MARKOV	54
6.1.	Intensidad condicional de Papangelou y condiciones de estabilidad.....	58
6.2.	Procesos con interacciones pares	59
7.	OTROS MODELOS	60
8.	MODELOS MULTIVARIANTES	62
	ESTIMACIÓN DE LAS FUNCIONES NOTABLES.....	67
1.	ESTIMACIÓN DE LAS PROPIEDADES DE SEGUNDO ORDEN.....	67
1.1.	Estimadores empíricos para procesos homogéneos.....	67
1.2.	Estimadores empíricos para procesos inhomogéneos.....	71
1.3.	Estimadores empíricos para procesos multivariantes.	73
1.4.	Estimación de la intensidad de segundo orden con integrales ponderadas.	73
2.	ESTIMACIÓN NO PARAMÉTRICA DE LA INTENSIDAD VARIABLE.	74
2.1.	Estimador núcleo de la intensidad.	74
2.2.	Estimación bootstrap de la intensidad de un proceso puntual inhomogéneo en \mathbb{R}	77
2.3.	ESTIMACIÓN DE LA INTENISDAD PARA LOS DATOS DE FUEGOS.	79
	PROCESOS INHOMOGENEOS	87
1.	K- FUNCIÓN PARA PROCESOS INHOMOGÉNEOS.	87
2.	APLICACIÓN A FUEGOS.	89

3.	PROCESOS PUNTUALES MULTITIPO	95
3.1.	Función K-cross	95
3.2.	Análisis de interacción de los datos.	96
AJUSTE DE MODELOS		103
1.	AJUSTE DE MODELOS POR MÁXIMA VEROSIMILITUD	103
1.1.	Verosimilitud.	103
1.2.	Pseudoverosimilitud.	104
2.	ANÁLISIS DE BONDAD DE AJUSTE	106
2.1.	Intensidad condicional de Papangelou.	107
2.2.	Residuos para procesos puntuales espaciales.	108
2.3.	Gráficos de diagnóstico.	111
3.	TEST DE BONDAD DE AJUSTE PARA PROCESOS INHOMOGÉNEOS DE POISSON.....	112
4.	AJUSTE PARAMÉTRICO PARA LOS FUEGOS.	113
CONCLUSIÓN		123
1.	CONCLUSIONES DEL ANÁLISIS DE LOS INCENDIOS FORESTALES.	123
2.	FUTURAS LINEAS DE INVESTIGACIÓN.	124
BIBLIOGRAFÍA.....		125
1.	PROCESOS PUNTUALES ESPACIALES	125
2.	INCENDIOS FORESTALES.....	128

INTRODUCCIÓN

Las formaciones de conjuntos de datos distribuidos de modo aleatorio en una región plana aparecen en distintos contextos; encontramos ejemplos en campos tan importantes como la ecología: localización de árboles en una región, nidos de una colonia de pájaros; o la epidemiología: distribución de enfermos en una región. Llamamos *diseño puntual espacial* a cualquiera de estos conjuntos de datos y *eventos* a las localizaciones, para distinguirlos de puntos arbitrarios de la región.

Los *procesos puntuales espaciales* son modelos estocásticos complejos que describen la localización de eventos de interés y, en ocasiones alguna información relativa a estos eventos. Los modelos más habituales son aquellos donde las localizaciones se recogen en dos dimensiones, (x, y) . Pero podemos usar técnicas similares en tres dimensiones (por ejemplo, localización de galaxias en el espacio), o para datos unidimensionales (animales a lo largo de la costa o en la ribera de un río). Los procesos puntuales univariantes incluyen solo la localización de los eventos, los procesos puntuales con marcas (o marcados) incluyen información adicional sobre cada evento; por ejemplo nos puede interesar la especie de cada árbol, o el diámetro de su tronco.

Estos conjuntos de datos se pueden usar para responder a una gran variedad de preguntas. El contexto científico de estas preguntas depende del área de aplicación, pero podemos clasificar las cuestiones en tres grandes grupos. En primer lugar, nos interesa el patrón espacial de los datos observados: ¿están agrupados?, ¿están distribuidos regularmente?, o ¿es su localización aleatoria?. Un segundo grupo de preguntas atiende a la relación entre distintos tipos de eventos en un proceso con marcas: ¿tienden dos especies de árboles a aparecer juntas?. El tercero y último grupo de preguntas se centra en la densidad (número de eventos por unidad de área).

Los procesos puntuales espaciales se aplican en ciencias como agricultura, astronomía, biología, climatología, ecología, epidemiología, geografía, geología, etc. La estadística proporciona una serie de herramientas para analizar un conjunto de puntos determinado. Y los procesos puntuales espaciales proporcionan una gran variedad de diseños para modelar comportamientos particulares. Además de la descripción de datos en el espacio, la estadística ofrece modelos o procesos que pueden generar datos puntuales según leyes estocásticas. Estos modelos son, generalmente, el resultado de considerar situaciones reales; por ejemplo, en el caso biológico, la localización de animales o plantas depende de la existencia de depredadores o comida. Una vez que tenemos una familia de modelos es importante determinar cuando un conjunto de datos dado se ajusta a uno de esos modelos, para esto es necesaria la inferencia, tanto paramétrica como no paramétrica.

Los incendios forestales suponen uno de los principales problemas medioambientales a los que se enfrenta la sociedad actual y en particular Galicia. El objetivo de este trabajo es utilizar la metodología estadística desarrollada en procesos puntuales espaciales para analizar y modelar el comportamiento de los incendios forestales en Galicia.

Nótese que si a cada incendio forestal se le asocian sus coordenadas espaciales, longitud y latitud del centroide de la zona quemada o del lugar donde se detecta el fuego, junto con otras variables como magnitud del incendio, tipo de terreno quemado o causa del incendio y además se controla el momento temporal en el que se ha producido, podemos identificar el conjunto de incendios forestales con un proceso estocástico espacio-temporal. Este tipo de procesos suelen presentar dependencia entre las posiciones espaciales y los instantes temporales, así como interdependencia entre ambas. En este trabajo nos centraremos en el aspecto espacial, entendiendo el conjunto de datos como un proceso puntual espacial con marcas e incluyendo el aspecto temporal (año del incendio) como una marca.

En este trabajo se aplicarán distintas técnicas de análisis e inferencia en procesos puntuales espaciales al conjunto de datos correspondiente a los incendios forestales registrados en el distrito Fonsagrada- Ancares en el periodo 1991-2008. Para esto se va combinando el desarrollo metodológico de las distintas técnicas con su aplicación a los datos. Así, en el primer capítulo se introduce la problemática de los incendios forestales, comentando brevemente los distintos estudios realizados al respecto, y se describe la situación de Galicia y en concreto del área de estudio, presentando el conjunto de datos que vamos a analizar. En el capítulo 2 se introducen los procesos puntuales espaciales. En los capítulos siguientes nos centramos en cada uno de los pasos del análisis de un diseño puntual espacial: tests de CSR, modelos, estimación no paramétrica de las funciones notables, análisis de procesos inhomogéneos y ajuste paramétrico de modelos.

INCENDIOS FORESTALES

1. ESTUDIO DE LOS INCENDIOS FORESTALES.

Entender el comportamiento del fuego es esencial para mejorar la gestión forestal. En concreto, conocer con precisión la distribución espacial de los posibles focos para después modelar y predecir el comportamiento de los incendios. En esta sección se muestra, en primer lugar, una revisión de los principales aspectos asociados a los incendios forestales, se presentan los principales trabajos desarrollados en relación al comportamiento completo de los incendios forestales y sus repercusiones, partiendo desde las causas hasta la recuperación de las zonas quemadas. Nos vamos a centrar en aquellas técnicas y modelos que han sido utilizadas para predecir la probabilidad de ocurrencia de un fuego. A continuación se caracteriza la Comunidad Autónoma de Galicia en relación con los incendios forestales y la zona objeto de estudio, distrito Fonsagrada-Ancares.

1.1. Comportamiento del ciclo de los incendios.

El fuego es una componente importante en muchos ecosistemas forestales (Moritz, 2000) con una gran influencia en las consecuencias ecológicas (Minnich & Bahre, 1995), estos fenómenos, al igual que otros (avalanchas, terremotos, tormentas de arena, etc.), tienen la propiedad de que cuando se producen por encima de un determinado umbral provocan una cascada de actividad ambiental, social y económica (Malamud, et al. 1998) y a su vez experimentan comportamientos muy distintos según el ámbito donde se localizan (Reed & McKelvey, 2002), variando en función de variables naturales y socioeconómicas ligadas a los paisajes forestales y presentando comportamientos difícilmente parametrizables.

A pesar de ello, muchos han sido los autores que total o parcialmente han intentado mostrar cuáles son los modelos matemáticos que rigen el comportamiento de lo que definiríamos como “ciclo del fuego” y como éste ha experimentado cambios en su comportamiento a lo largo del tiempo y el espacio (Minnich & Bahre, 1995; Pereira et al., 1998; Hessburg, et al., 2000; Vázquez & Moreno, 2001; Moreira, et al. 2001; Lloret, et al., 2002; Schoenberg, et al. 2003; Floyd, et al. 2004; Mermoz, et al., 2005; Viedma, et al., 2006; González & Pukkala, 2007).

Centrando el ámbito del trabajo en relación con la Comunidad Autónoma de Galicia, dentro de lo que a escala mundial se define como los países Mediterráneos, el fuego se ha convertido en un problema muy grave durante las últimas tres décadas y es actualmente la principal causa de la destrucción de los bosques, con una media de área quemada de 500.000 hectáreas por año (Vélez, 2002). En Galicia han ardido 312.000 ha entre los años 1996 y 2005 sobre un total de 2.060.500 ha de superficie forestal. El máximo anual se alcanzó en el año

2006 en el que se quemaron un total de 95.945 Ha (DGCN, 2006). Por ello los incendios forestales son percibidos como el principal problema medioambiental de la región (Alonso-Betanzos et al., 2003; Chas, 2007).

Ejemplos de variables utilizadas en los modelos de incendios con parámetros conocidos y controlables por los sistemas de planificación forestal son las relacionadas con la densidad del arbolado, composición de especies y la estructura vertical del dosel. Diferentes investigaciones han demostrado su correlación con la probabilidad de daños por incendios (Van Wagner, 1977; Finney, 1994; Moreira, et al. 2001; Pollet, 2002; González & Pukkala, 2007). La variación en la estructura forestal a lo largo del tiempo como consecuencia de actuaciones silvícolas o del fuego impone una variabilidad espacial en la distribución de la -- dimensión del área ardida (Cumming, 2001), suponiendo el fuego también cambios importantes en el paisaje. En algunos casos incrementando la heterogeneidad (Lloret, et al. 2002), en otros casos no afectando a las especies y ecosistemas existentes (Minnich, 1983). El cambio en el clima también parece estar influyendo directamente en el comportamiento del fuego (Vázquez & Moreno, 2001). También es conocida la estacionaridad de los incendios (Vázquez & Moreno, 2001 y Moritz, 2003), bien sea por causas naturales (Teensma, 1987) o por prácticas socioeconómicas (Lloret, et al. 2002), dándose la circunstancia de que muchas áreas son castigadas por el fuego por debajo del período de rotación de la zona, concentrándose los mayores fuegos en zonas específicas (Vázquez & Moreno, 2001). Un aspecto muy importante para la gestión forestal está en conocer cuáles son las especies con mayor nivel de supervivencia después de un incendio y cuáles son los factores que inciden en ello (Kushla & Ripple, 1997), junto con las causas que provocan un incremento en el tamaño de los mismos (Díaz-Delgado, et al., 2004). Así pues, la probabilidad de incendio depende de dos factores complejos como son:

1. Condiciones climáticas ([Andrews, 1986; Van Wagner, 1987, Dickson, et al. (2006)
2. Tipo y estructura de la vegetación (Mermoz, et al. 2005) que en algunos casos favorece y en otros casos parece frenar su avance (Nunes, et al. 2005) y que en las zonas de referencia están totalmente relacionadas con las condiciones socioeconómicas (Chas, 2007; Marey-Pérez, et al. 2006; Marey-Pérez & Rodríguez-Vicente, 2008).

Como es lógico las condiciones meteorológicas de un bosque en el futuro no se conocen, lo que significa que los modelos basados en factores meteorológicos no se puede utilizar en planificación forestal, cuyo objetivo es predecir las consecuencias a largo plazo de las diferentes alternativas de gestión, pero sí pueden ser utilizados para la simulación de escenarios previsibles. Por otra parte, gran cantidad de datos de valores medios de las diferentes variables climáticas son conocidos, lo que significa que pueden utilizarse en los modelos predictivos. La existencia de correlaciones entre variables climáticas y parámetros de posición tales como longitud, latitud, altitud, pendiente y orientación implica que estos parámetros pueden sustituir a las variables climáticas en la elaboración de modelos.

El tipo, cantidad y distribución de los combustibles forestales afectan a la ocurrencia de incendios y daños (Velez, 1990; Loehle, 2004). La cantidad y el estado de los combustibles

apenas se conoce exactamente en la planificación forestal ya que la cantidad de madera muerta y su estado de descomposición no suelen ser medidos o previstos. Sin embargo, cabe suponer que la cantidad total y las características de los combustibles se correlaciona con el crecimiento de población y características de cada sitio, lo que significa que los modelos basados en las características de posición pueden, en cierta medida, también representar la variación en la probabilidad de incendios causados por la cantidad y las propiedades de los combustibles. Las masas forestales pueden ser de la misma o diferente edad y con mezcla de especies. Un modelo útil de riesgo para la planificación forestal debe tener en cuenta la estructura y composición de las especies (Díaz Delgado et al., 2004) que son obtenidas en los inventarios forestales (González, et al., 2007). Los modelos, basados en datos empíricos, deben ser sensibles a los efectos de tratamientos silvícolas y tipos de gestión para establecer el riesgo de incendio en una determinada situación. Los modelos existentes no parecen cumplir todos estos requisitos.

En algunos modelos el riesgo de incendio depende de la edad (Reed, 1994; Moritz, 2003), y otros, como la clasificación de riesgo de incendio siguiendo el modelo utilizado por Thompson et al. (2000), se basan en evaluaciones cualitativas de expertos (González, et al. 2006) o en encuestas a ciudadanos (Tábara et al., 2003; Riera & Mogas, 2004) También resulta interesante conocer qué vegetación resulta de una serie continuada de fuegos y si ésta puede estar formando parte del ciclo natural del mismo (Viedma, et al. 2006).

2. DESCRIPCIÓN DE LA BASE DE DATOS.

La base de datos analizada en este trabajo recoge los fuegos registrados en el distrito Fonsagrada-Ancares en el periodo 1991-2008. Junto con la parroquia donde se ha iniciado cada uno de los fuegos, se presentan las siguientes variables:

- Fechas de inicio, control y fin del incendio.
- Detección: medio por el que se ha detectado el incendio.
- Causa: intencionado, natural, desconocida, basurero, reproducción, negligencia, quema y otros.
- Tipo: clasificación en conato, quema o incendio en función de la superficie quemada.
- Inicio: arbolado o raso, clasificación del terreno en función de su uso.
- Superficie quemada: total y según tipo de uso.
- Intervención: número de técnicos, agentes y autoridades que han intervenido en la extinción del fuego.

El área de estudio la compone el distrito forestal Fonsagrada-Ancares, formada por un total de 9 municipios y situado como muestra la ilustración 1 en la zona oriental de la provincia

de Lugo, Sus coordenadas geográficas son Norte 43º 18', 7º 0', Sur 42º 36', 7º 3', Oeste 42º54' N, 7º 19' y Este 42º 51', 6º 44'.



Ilustración 1: Localización del distrito Fonsagrada- Ancares.

Concello	Sup. Arbolada (ha)	% Arb s/ Total	Sup. Matorral (ha)	% Matorral s/ Total	Sup. Total Forestal (ha)	% Forestal s/ Total	% Arb s/ Forestal
Baleira	6.631,49	39,29	6.134,26	36,34	12.765,75	75,64	51,95
Baralla	5.416,31	38,60	3.355,44	23,55	8.771,75	62,16	61,75
Becerreá	7.386,41	42,94	5.219,87	30,34	12.606,28	73,29	58,59
Cervantes	13.657,09	49,19	10.434,58	37,58	24.091,67	86,79	56,69
A Fonsagrada	27.225,91	62,12	7.967,86	18,07	35.193,77	80,28	77,36
Navia de Suarna	11.392,92	46,97	9.275,40	38,24	20.668,32	85,22	55,12
Negueira de Muñiz	3.798,54	52,57	2.638,76	36,52	6.437,30	89,09	59,01
As Nogais	6.553,81	59,40	2.715,16	24,62	9.268,97	84,02	70,71
Pedrañita	5.765,05	54,98	2.500,84	23,84	8.265,89	78,81	69,75
TOTAL DISTRITO	87.827,53	50,83	50.242,17	29,08	138.069,70	79,91	63,61
TOTAL LUGO	463.818,00	47,08	192.429,00	19,53	656.248,00	66,61	70,68
TOTAL GALICIA	1.405.452,00	47,54	634.123,00	21,45	2.039.575,00	68,99	68,91

Tabla 1: Superficie forestal desglosada por municipios y comparada con los totales de Lugo y Galicia.

La superficie forestal representa un total de 138.069 ha lo que supone el 79,9% de la superficie total del distrito, de esta un total de 87.828 ha están arboladas, lo que supone el 63,6% de la superficie forestal. La distribución por municipios se presenta en la tabla 1.

2.1. Los incendios forestales en el distrito Fonsagrada- Ancares.

Los incendios forestales pueden considerarse uno de los grandes problemas sociales, ambientales y económicos a los que se enfrenta cualquier sociedad (CES, 2005). En esta sección se presenta un breve análisis descriptivo de la evolución de los incendios forestales (número de fuegos y hectáreas quemadas) y sus causas en el es distrito Fonsagrada- Ancares.

La prevención y defensa contra los incendios forestales en Galicia está regulada en la *Ley 3/2007, del 9 de abril de prevención y defensa contra los incendios forestales de Galicia* (DOG, nº74, del 17 de abril de 2007), junto con su corrección de errores (DOG, nº 96, del 21 de mayo de 2007).

En el artículo 2 de la citada Ley, se definen los términos de incendio forestal y quema; siendo:

- **Incendio forestal:** el fuego que se extiende sin control sobre combustibles forestales situados en el monte, incluyéndose los enclaves forestales localizados en terrenos agrícolas cualquiera que fuera su extensión.
- **Quema:** fuego que se extiende sin control sobre el monte raso.

En el PLADIGA 2008 (Plan de Defensa contra Incendios de Galicia) se definen los conceptos de conato, incendio y quema de acuerdo al esquema que se muestra en la ilustración 2, así tenemos:

- **Quema:** cuando la superficie quemada rasa sea inferior a 1 ha y no se quemase superficie arbolada.
- **Conato:** cuando la superficie quemada, arbolada más rasa, sea igual o menor a 1 ha, siempre que el arbolado quemado sea igual o menor a 0,5 ha, y cuando afectando solamente a la superficie rasa esta no supere 1 ha.
- **Incendio:** cuando la superficie quemada arbolada sea superior a 0,5 ha o la suma de superficie arbolada y rasa quemada supere la hectárea.

A partir de ahora adoptaremos la nomenclatura definida por el PLADIGA 2008, de modo que nos referiremos a los incendios forestales como fuegos, dejando el término incendio para los fuegos que alcanzan la superficie que se indica en la definición.

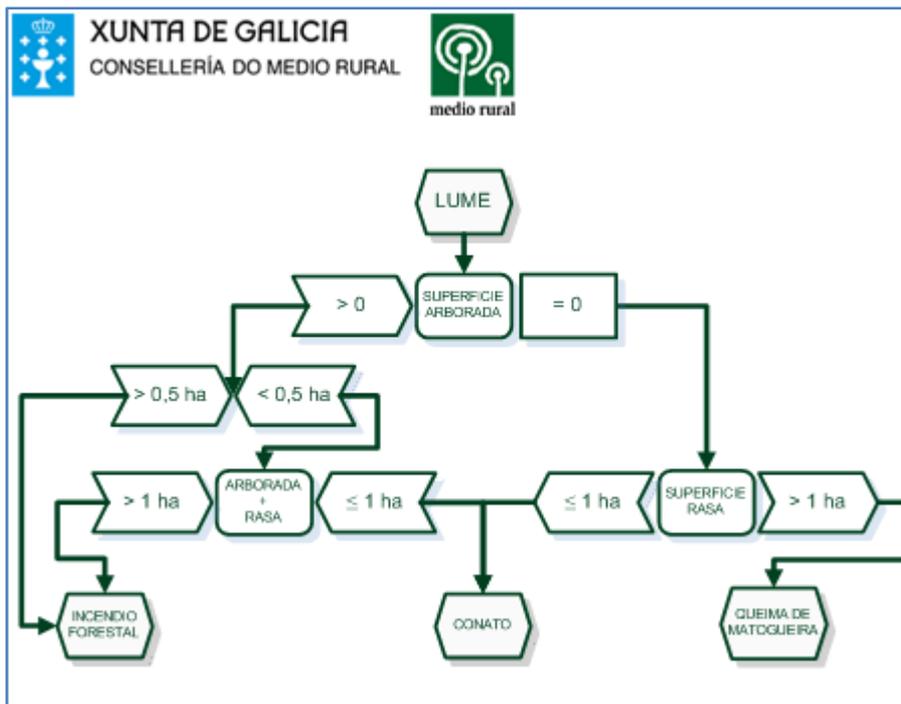


Ilustración 2: Clasificación de fuegos según superficie quemada.

Teniendo en cuenta los términos anteriores, el tipo de fuego predominante en el distrito Fonsagrada-Ancares entre 1991 y 2008 son los conatos que constituyen el 57% del los fuegos registrados, le siguen las quemas con un 23% y por último los incendios con el 19%, estos datos se pueden observar en el gráfico 1.

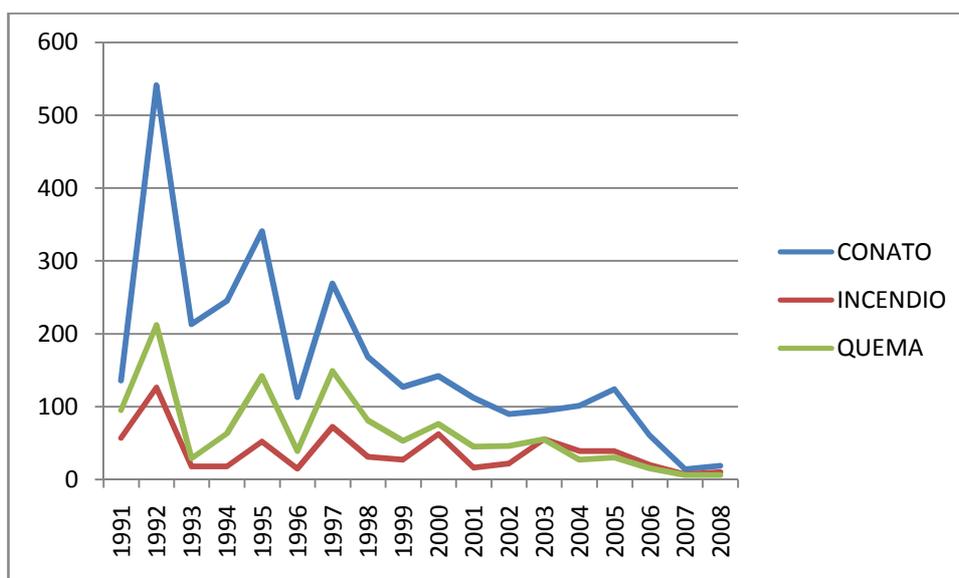


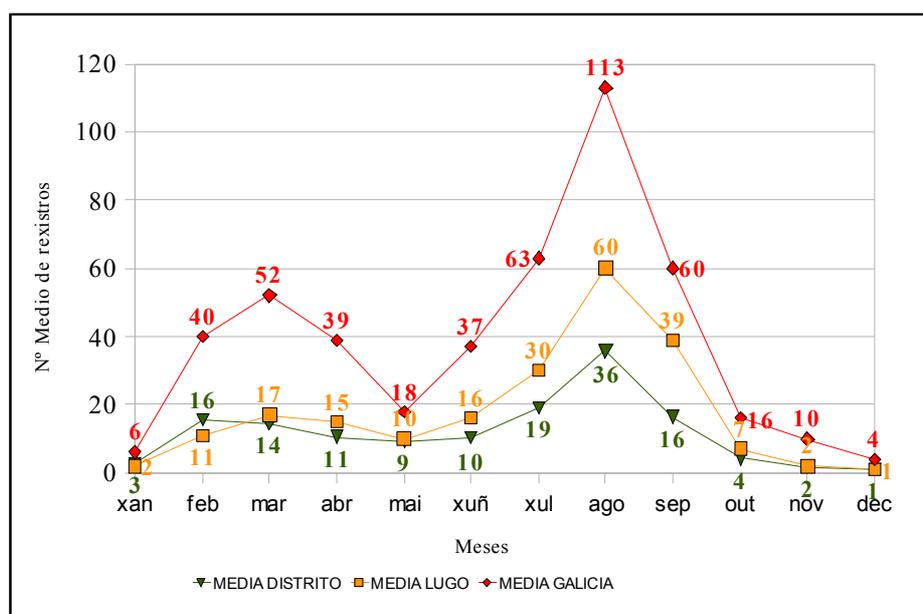
Gráfico 1: Evolución de los distintos tipos de incendios forestales en el periodo 1991-2008.

Fuente – Elaboración propia a partir de datos de la Consellería de Medio Rural

A lo largo de los últimos años, Galicia se ha caracterizado por la existencia de dos períodos con alto número de incendios: finales de invierno-primavera temprana y el período estival. Como refleja el gráfico 2, es necesario destacar que la evolución del número medio de registros para el período de referencia 2002 – 2007 en la época estival sigue la misma tendencia en Galicia, Lugo y el Distrito, produciéndose el mayor número de incendios en el mes de agosto. Sin embargo, es la época de finales de invierno-primavera temprana y en el mes de febrero cuando se producen mayor número de incendios en el Distrito en contra de lo que ocurre en Lugo y Galicia que es en el mes de marzo.

Respecto a la superficie quemada y número de fuegos por municipio en el Distrito, tomando como referencia el período 1991-2008 son Cervantes y Navia de Suarna los municipios que presentan mayor superficie quemada, mientras que es en Cervantes y A Fonsagrada donde se registran el mayor número de incendios forestales (43%). Hay que destacar que Negueira es el municipio con menor número de registros (2%) y con menor superficie quemada (3%).

Gráfico 2: Evolución del número medio de registros para diferentes zonas entre 2002 a 2007
Fuente – Elaboración propia a partir de datos de la Consellería de Medio Rural



A continuación se analizan las causas de los incendios forestales en el Distrito y la superficie afectada por los mismos, tomando como período de referencia 1991-2008. Como refleja la tabla 2, de los 4.764 fuegos que se registraron en el Distrito, el 87% son intencionados. La superficie quemada por estos incendios provocados es de 15.398,56 ha, lo que significa que algo más de la tercera parte de la superficie quemada en el Distrito se debe a fuegos intencionados.

Le siguen en importancia los fuegos debidos a causas naturales, que representan el 5,86% del total y el 7,86% de la superficie quemada, y los que tienen causa desconocida, que suponen el 4,14% del total y el 8,42% de la superficie.

Hay que destacar que los fuegos producidos por basureros suponen la causa menos importante en todo el Distrito. En el año 2005 se llevó a cabo la clausura y el sellado del vertedero de Becerreá, pero siguen funcionando los basureros de Baleira, Baralla, Fonsagrada (A), Navia de Suarna, Negueira de Muñiz, Nogais (As) y Pedrafita do Cebreiro.

Gráfico 3: Superficie quemada y número de registros por municipio en el periodo 1991-2008.

Fuente – Elaboración propia a partir de datos de la Consellería de Medio Rural.

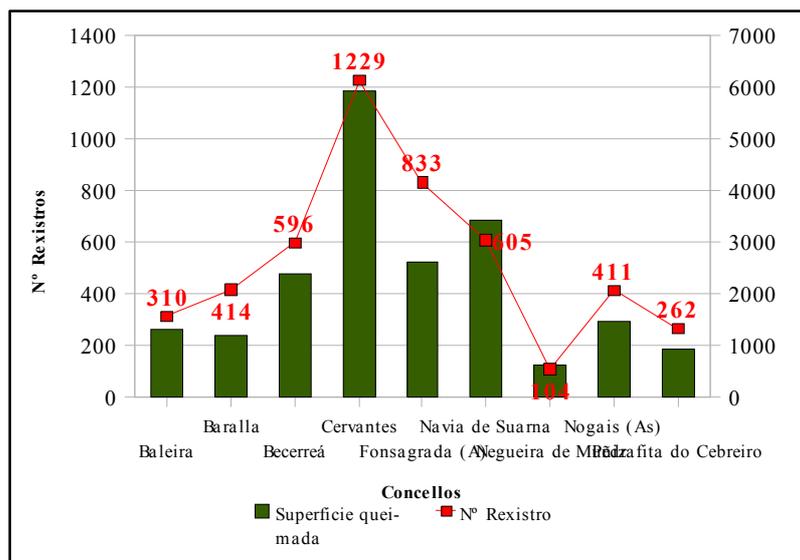


Tabla 4: Causas de incendios forestales para el distrito Fonsagrada- Ancares en el periodo 1991-2008.

Fuente – Elaboración propia a partir de datos de la Consellería de Medio Rural.

Causa	Nº Registros	% Registros	Superficie	% Superficie
Basurero	11	0,23%	1,42	0,01%
Desconocido	197	4,14%	1.664,41	8,42%
Reproducción	24	0,50%	243,30	1,23%
Intencionado	4.153	87,17%	15.398,56	77,86%
Negligencia	42	0,88%	265,32	1,34%
Otros	41	0,86%	225,48	1,14%
Quema	17	0,36%	423,12	2,14%
Causa natural	279	5,86%	1.555,33	7,86%
TOTAL	4.764	100%	19.776,94	100%

2.2. Análisis espacial de los incendios a nivel de parroquia.

Se ha fijado una zonificación del Distrito por parroquias de acuerdo a su conflictividad en lo referente a incendios forestales en función de los datos del número de incendios y la superficie total quemada por parroquia en los últimos 18 años.

Para hacer este análisis el primer paso ha sido dividir el período de estudio (18 años) en tres etapas definidas por la evolución temporal del número de incendios a nivel Distrito:

1º período: 1991-1996

2º período: 1997-2002

3º período: 2003-2008

Para cada una de las etapas se trabajó con los datos de superficie total quemada (ha/km²) y el número de incendios clasificando los datos en cuatro grupos delimitados por la mediana y el percentil 25 y 75. A cada rango de datos se le asignó un número en el intervalo 1-4 en orden creciente según su nivel de conflictividad.

Si los datos están dentro del rango marcado por el percentil 25 se les asigna el valor 1, si los datos están en el rango descrito por el percentil 25 y la mediana les corresponde el valor 2, en caso de estar entre la mediana y el percentil 75 se les asigna el valor 3 y al resto les corresponde el valor 4.

Finalmente, se calculó un valor para el total del período de estudio mediante un sistema de ponderación que da mayor representatividad a los datos más recientes de acuerdo con la siguiente fórmula:

$$\text{Grado de conflictividad} = \frac{\text{Valor}1^{\circ}\text{período} + 1,5 \cdot \text{Valor}2^{\circ}\text{período} + 2 \cdot \text{Valor}3^{\circ}\text{período}}{4,5}$$

Como resultado, cada parroquia tiene asignados 2 valores que indican su conflictividad en función del número de fuegos y la superficie total quemada (ha/km²) de acuerdo estos dos valores las parroquias se clasificarán como vemos en la ilustración 3:

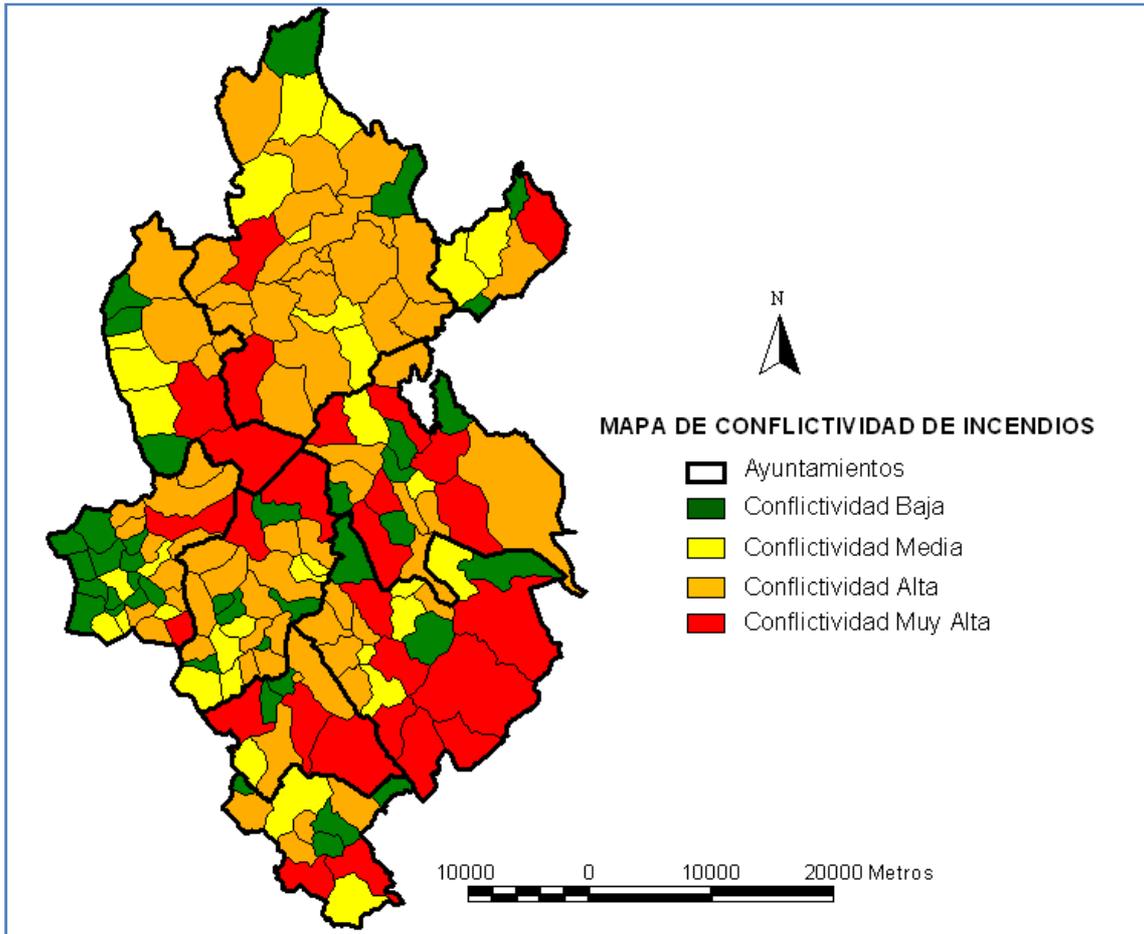


Ilustración 3: Clasificación de las parroquias según conflictividad.

Fuente – Elaboración propia a partir de datos de la Consellería de Medio Rural.

2.3. Localización espacial de los fuegos.

En la base de datos original la localización de los fuegos viene dada por la parroquia en la que se ha registrado, pero para poder interpretar estos datos como un proceso puntual espacial necesitamos asociar coordenadas espaciales a los eventos. A continuación se describe el método utilizado para asociar coordenadas a cada fuego.

Actualmente los Sistemas de Información Geográfica (SIG) son una herramienta imprescindible para el manejo de los datos asociados al terreno, ya que la gran cantidad de variables que interactúan en el medio natural hace necesario el empleo de programas informáticos en los que se recoja simultáneamente información gráfica y bases de datos asociadas a ésta (información no gráfica o alfanumérica). Así, permiten trabajar con gran cantidad de información y realizar complejas operaciones sobre un determinado escenario.

Partiendo de la potencialidad de los SIG, se pretende localizar geográficamente los 4764 incendios forestales ocurridos en la zona de estudio, para poder aplicar análisis espacial. De esta forma también podremos asociar a estos fuegos información relativa a otras variables

del terreno, tales como orientación, pendiente o distancias a carreteras o a núcleos de población, para ver cuál es su influencia en el comportamiento del fuego (Yang et al, 2007)

En un SIG se tienen que relacionar los atributos gráficos y no gráficos (alfanuméricos), esto se consigue mediante atributos de unión. Dado que se dispone de toda la información gráfica, el nexa de unión para localizar geográficamente los fuegos son los datos disponibles de parroquia y uso del suelo (forestal arbolado y desarbolado o raso) donde se iniciaron. De este modo, atribuimos al centro de cada tesela (definida por la parroquia y el uso) los incendios que le correspondan, según el procedimiento expuesto en la ilustración 3:

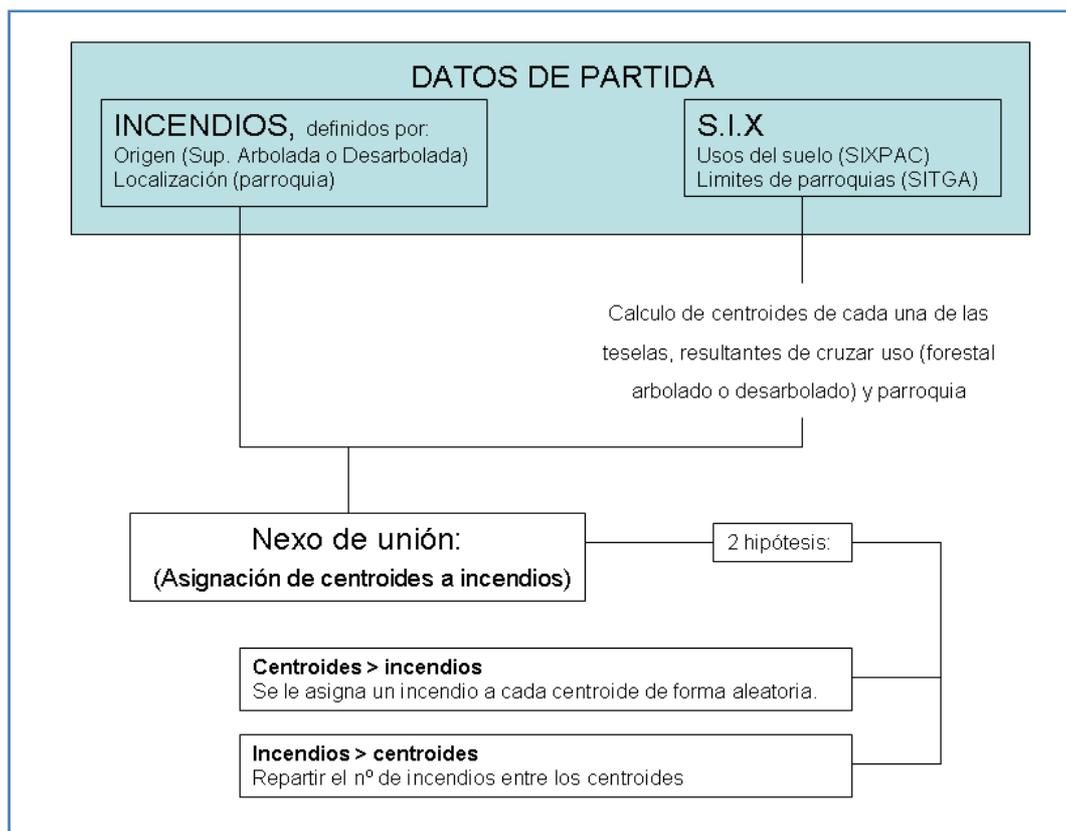


Ilustración 4: Esquema de localización geográfica de los incendios forestales

Por tanto, conocemos para cada incendio forestal sus coordenadas UTM de origen, cumpliendo así con la misión de localizarlos en el terreno. Con esta información estamos en condiciones de asociar a cada evento variables espaciales que influyen tanto en el comportamiento del fuego, como en el tiempo de actuación para su extinción.

El resultado obtenido tras realizar dicho análisis geográfico es una base de datos con todos los incendios codificados y con toda la información de los mismos (municipio, parroquia, fecha, tipo, causa, etc.), a la cual vinculamos variables físicas tales como la pendiente, la orientación, tipo de propiedad, distancia a núcleos urbanos y carreteras, y modelos de combustible existentes en el punto de inicio de cada incendio.

PROCESOS PUNTUALES ESPACIALES

Un **proceso puntual** es un modelo estocástico que genera un número finito de eventos $\{x_i, i=1\dots n\}$ en un conjunto X . Si los eventos tienen asociadas medidas o marcas tenemos un **proceso puntual con marcas**.

Como estamos interesados en procesos puntuales espaciales, X será una región acotada de \mathbb{R}^d o un toro. De modo más general X puede ser un espacio de Hausdorff localmente compacto con una topología segundo numerable.

Si un evento situado en $s \in \mathbb{R}^d$ está asociado a una marca $z \in \mathcal{F}$, entonces (s, z) es un punto en $\mathbb{R}^d \times \mathcal{F}$. Por tanto, un proceso puntual con marcas se puede representar como un proceso puntual en $\mathbb{R}^d \times \mathcal{F}$. Del mismo modo, si observamos un proceso puntual con marcas en los instantes $(t_1, t_2, \dots, t_n) \in T$, se puede tomar como un proceso puntual en $\mathbb{R}^d \times \mathcal{F} \times T$, dando lugar a los procesos puntuales espacio-temporales

Sea χ la σ -álgebra de Borel en $X \subset \mathbb{R}^d$ y ν la medida de Lebesgue en X ($\nu(B) = |B|$, volumen de B). La forma más natural de definir un diseño puntual (una realización de un proceso puntual espacial) es considerando la localización espacial de los eventos x_1, x_2, \dots en X . Matemáticamente es más adecuado definir un diseño puntual mediante una medida ϕ en X . Para cada conjunto de Borel B , $\phi(B)$ es el número de eventos en B , así $\phi(B) \in \{1, 2, \dots\}$ para todo $B \in \chi$. Esta medida se supone localmente finita, es decir $\phi(B) < \infty$ para todo $B \in \chi$ acotado. Conocer $\phi(B)$ para todo $B \in \chi$ equivale a conocer la localización espacial de todos los eventos x_1, x_2, \dots en X . Tenemos que

$$\phi(B) = \sum_{i=1}^{\infty} I(s_i \in B), \quad \forall B \in \chi$$

Recíprocamente, x es un evento del diseño puntual si $\phi(\{x\}) > 0$. Estas dos definiciones de los diseños puntuales espaciales nos llevan a dos caracterizaciones equivalentes de los procesos puntuales.

1. ALEATORIEDAD ESPACIAL COMPLETA (CSR)

La aleatoriedad espacial completa (CSR, Complete Spatial Radomness), es el “ruido blanco” de un proceso puntual espacial, pues caracteriza la ausencia de estructura en los datos. Por tanto, habitualmente es la hipótesis nula de un test estadístico para determinar

cuando hay estructura espacial en un diseño puntual dado. Un proceso completamente aleatorio se identifica con un proceso homogéneo de Poisson y está definido por los siguientes postulados:

1. El número de eventos en una región plana A , con área $|A|$ sigue una distribución de Poisson con media $\lambda|A|$, donde λ es la intensidad del proceso, es decir, el número esperado de eventos por unidad de área.
2. Dados n eventos $\{x_i\}_{i=1}^n$ en la región A , los x_i son una muestra aleatoria de la distribución uniforme en A . Por tanto, no hay interacción entre eventos.

Por 2, la densidad condicional de una n -upla ordenada (x_1, x_2, \dots, x_n) dado $N(A) = n$, es

$$f(x_1, x_2, \dots, x_n) = 1/|A|^n$$

Así, teniendo en cuenta ambos postulados, la densidad conjunta de n y (x_1, x_2, \dots, x_n) , viene dada por

$$f((s_1, s_2, \dots, s_n), n) = \frac{\lambda^n e^{-\lambda|A|}}{n!}$$

El principal interés de la CSR es que representa un modelo ideal que, a pesar de no alcanzarse en la práctica, resulta muy útil como primera aproximación de algunos patrones observados. En general, el análisis de un diseño puntual observado se inicia con un test de CSR, por dos razones fundamentales. En primer lugar, un diseño que verifica CSR, apenas necesita más análisis estadístico formal. En segundo lugar, además del interés intrínseco que tiene rechazar CSR, los tests se utilizan como método de exploración de los datos al proporcionar información del sentido en que nos estamos desviando de la hipótesis nula.

2. PROPIEDADES DE LOS PROCESOS PUNTUALES

Sea $N(A)$ la variable aleatoria que representa el número de eventos en una región plana A .

$$N(A) = \#(x_i \in A)$$

Veamos algunas propiedades de interés que pueden verificar los procesos puntuales espaciales:

- a) El proceso es *estacionario* si, para cualquier entero k y regiones $A_i, i=1, \dots, k$, la distribución conjunta de $N(A_1), N(A_2), \dots, N(A_k)$ es invariante por translaciones de los A_i por cualquier conjunto x .

b) El proceso es *isotrópico* si, para cualquier entero k y regiones $A_i, i=1, \dots, k$, la distribución conjunta de $N(A_1), N(A_2), \dots, N(A_k)$ es invariante por rotación de la unión de los A_i .

c) El proceso es *ordenable* cuando no hay eventos coincidentes, es decir:

$$\lim_{|dx| \rightarrow 0} \frac{P(N(dx) > 1)}{|dx|} = 0$$

d) El proceso es *ordenable de segunda orden* cuando para cualquier par de eventos x e y se verifica

$$\lim_{\substack{|dx| \rightarrow 0 \\ |dy| \rightarrow 0}} \frac{P(N(dx) > 1)P(N(dy) > 1)}{|dx||dy|} = 0$$

3. CARACTERÍSTICAS DE PRIMER Y SEGUNDO ORDEN

3.1. Procesos univariantes

A continuación vamos a ver cómo definir en procesos puntuales estructuras análogas a la media y la covarianza para procesos con valores reales.

Las características de primer y segundo orden de la variable aleatoria $N(B)$, para $B \subseteq S$, están descritas por la medida de intensidad y el momento factorial de orden dos, definidos como sigue:

Definición:

La *medida de intensidad* μ en \mathbb{R}^d viene dada por

$$\mu(B) = E[N(B)], \quad B \subseteq \mathbb{R}^d$$

El *momento factorial de orden 2*, $\alpha^{(2)}$ en $\mathbb{R}^d \times \mathbb{R}^d$ viene dado por

$$\alpha^{(2)}(C) = E \left[\sum_{\xi \neq \eta \in X} I((\xi, \eta) \in C) \right], \quad C \subseteq \mathbb{R}^d \times \mathbb{R}^d$$

Si la medida de intensidad μ se puede escribir como:

$$\mu(B) = \int_B \lambda(\xi) d\xi$$

donde λ es una función no negativa, entonces λ será la *función de intensidad*. Sea dx un entorno infinitesimal del punto x :

Definición:

La *función de intensidad de primer orden* de un proceso puntual espacial es

$$\lambda(x) = \lim_{|dx| \rightarrow 0} \left\{ \frac{E[N(dx)]}{|dx|} \right\}$$

Definición:

La *función de intensidad de segundo orden* de un proceso puntual espacial es

$$\lambda_2(x, y) = \lim_{\substack{|dx| \rightarrow 0 \\ |dy| \rightarrow 0}} \left\{ \frac{E[N(dx)N(dy)]}{|dx||dy|} \right\}$$

Definición:

La *densidad de covarianza* de un proceso puntual espacial es

$$\gamma(x, y) = \lambda_2(x, y) - \lambda(x)\lambda(y)$$

Para un proceso puntual estacionario e isotrópico, tenemos:

1. $\lambda(x) = \lambda = E[N(A)]/|A|$ (constante para todo A)
2. $\lambda_2(x, y) = \lambda_2(\|x - y\|)$
3. $\gamma(u) = \lambda_2(u) - \lambda^2$

Definición:

La *intensidad condicional* de un proceso puntual espacial es

$$\lambda_c(x/y) = \lim_{|dx| \rightarrow 0} \left\{ \frac{E[N(dx)/\text{evento en } y]}{|dx|} \right\}$$

Lema:

Para un proceso estacionario isotrópico y ordenable, se verifica:

$$\lambda_c(x/y) = \lambda_2(\|x - y\|)/\lambda$$

Baddeley *et al.* (2000), estudiaron los procesos estacionarios reponderados de orden 2, que verifican $\lambda_2(x, y)/\lambda(x)\lambda(y) = \rho(t)$, depende solo de $t = \|x - y\|$. Nótese que $\lambda(\cdot)$ no puede tomar valores muy próximos a 0. La estacionariedad reponderada en procesos puntuales es análoga a suponer en un proceso espacial con valores reales que la media varía en el espacio mientras que las variaciones en torno a la media local son estacionarias.

A partir de ahora, salvo indicación específica se asume estacionariedad, isotropía y ordenabilidad. Bajo estas condiciones λ tiene fácil interpretación como el número esperado de eventos por unidad de área. La intensidad de segundo orden es más difícil de interpretar. Consideramos una nueva definición:

Definición:

La *función de segundo momento reducida* de un proceso puntual estacionario e isotrópico viene dada por:

$$K(t) = \lambda^{-1} E[N_0(t)] \quad (2.1)$$

donde $N_0(t)$ denota el número de eventos situados a una distancia menor que t de un evento arbitrario. Se entiende por arbitrario un evento elegido al azar.

La importancia práctica de esta función está en que se puede expresar como la media de una cantidad observable, esto sugiere un método para estimarla a partir dos datos.

Para establecer una relación entre $K(t)$ y $\lambda_2(t)$ se asume que el proceso es ordenable, esto significa que $E[N(dx)] \sim P[N(dx)=1]$ pues su cociente tiende a 1 cuando $|dx| \rightarrow 0$. Suponemos también que $E[N(dx)N(dy)] \sim P[N(dx)=N(dy)=1]$ en el mismo sentido. Bajo estas condiciones, el número esperado de eventos a una distancia menor que t de un evento arbitrario se puede calcular integrando la intensidad condicional en el disco de centro el origen y radio t . Así

$$\lambda K(t) = \int_0^{2\pi} \int_0^t \lambda_c(x/0) x dx d\theta$$

teniendo en cuenta que $\lambda_c(x/0) = \lambda_2(x)/\lambda$

$$K(t) = 2\pi\lambda^{-2} \int_0^t \lambda_2(x) x dx \quad (2.2)$$

o, recíprocamente

$$\lambda_2(t) = \lambda^2 (2\pi t)^{-1} K'(t) \quad (2.3)$$

A pesar de que $K(t)$ tiene una interpretación más intuitiva que $\lambda_2(t)$, desde el punto de vista teórico es recomendable trabajar con $\lambda_2(t)$.

Si podemos expresar el momento factorial de orden 2 como

$$\alpha^{(2)}(B_1 \times B_2) = \int_{B_1} \int_{B_2} l^{(2)}(s, t) ds dt$$

donde $l^{(2)}$ es una función no negativa, entonces $l^{(2)}$ se denomina *densidad producto de segundo orden*.

Para ds y dt infinitesimales, $l^{(2)}(s,t)dsdt$ se puede interpretar como la probabilidad de tener un punto del proceso en cada uno de los conjuntos con áreas ds y dt . Si el proceso es estacionario $l^{(2)}$ depende solo de $h=t-s$. En el caso isotrópico $l^{(2)}$ depende solo de la distancia $r = \|t-s\|$. Podemos normalizar $l^{(2)}$ dividiendo por el cuadrado de la intensidad de primer orden. Definiendo así, bajo isotropía, la *función de correlación par*, g

$$g(r) = \frac{l^{(2)}(r)}{\lambda(x)^2}$$

Para un proceso completamente aleatorio $g(r)=1$. Si $g(r)>1$, tenemos distancias próximas a r con más frecuencia que en un proceso completamente aleatorio, esto es típico de procesos con clústers. Cuando $g(r)<1$, la presencia de esta distancia no es habitual, característica propia de los procesos de inhibición.

Bajo estacionariedad e isotropía tenemos

$$\lambda(x)^2 dK(r) = 2\pi r l^{(2)}(r) dr$$

En consecuencia

$$g(r) = \frac{1}{2\pi r} \frac{d}{dr} K(r)$$

o, recíprocamente

$$K(r) = \int_0^r g(u) 2\pi u du, \quad r > 0$$

Un *refinamiento aleatorio* de un proceso puntual P , es un proceso puntual que tiene por eventos un subconjunto de eventos de P generado manteniendo o eliminando elementos de P en función de tiradas independientes de un proceso de Bernouilli. La K -función es invariante por refinamientos aleatorios, es decir la K -función de cualquier refinamiento de un proceso puntual coincide con la del proceso original.

En lugar de observar la localización exacta de los eventos en una región plana, en ocasiones resulta más fácil en la práctica observar solo contadores $N(A)$ en subregiones adecuadas. La distribución del contador resultante

$$p_n(A) = P(N(A) = n); \quad n = 0, 1, \dots$$

proporciona un posible estadístico descriptivo del proceso. Debido a la naturaleza aleatoria de A puede resultar problemático estudiar esta distribución. Una posible solución es estudiar la distribución del contador a través de sus momentos, calculados en función de la región de observación. En particular

$$E[N(A)] = \int_A \lambda(x) dx$$

que se reduce a $\lambda|A|$ para procesos estacionarios. Además,

$$E[N(A)^2] = E\left[\left\{\int_A N(dx)\right\}^2\right] = E\left[\int_A N(dx) + \int_A \int_A N(dx)N(dy)\right].$$

Intercambiando la esperanza con las integrales tenemos, para un proceso estacionario,

$$E[N(A)^2] = \lambda|A| + \int_A \int_A \lambda_2(x-y) dx dy$$

de donde

$$Var[N(A)] = \int_A \int_A \lambda_2(x-y) dx dy + \lambda|A|(1 - \lambda|A|) \quad (2.4)$$

y, directamente

$$Cov(N(A), N(B)) = \int_A \int_B \lambda_2(x-y) dx dy + \lambda|A \cap B| - \lambda^2|A||B|$$

3.2. Procesos multivariantes

Un proceso multivariante es un proceso puntual en el que hay distintos tipos de eventos. La definición de las características de primer y segundo orden para este tipo de procesos es una generalización natural del caso univariante. Se mantiene la suposición de que el proceso es estacionario, isotrópico y ordenable, y denotamos por $N_j(A)$ al número de eventos de tipo j en la región A .

Definición:

Las *funciones de intensidad de primer orden* para cada tipo de evento son

$$\lambda_j = \frac{E[N_j(A)]}{|A|}$$

Definición:

Las *funciones de intensidad de segunda orden* vienen dadas por

$$\lambda_{ij}(r) = \lim_{\substack{|dx| \rightarrow 0 \\ |dy| \rightarrow 0}} \left\{ \frac{E[N_i(dx)N_j(dy)]}{|dx||dy|} \right\}, \quad (\|x-y\|=r)$$

Definición:

Las *densidades de covarianza* de un proceso puntual multivariante son

$$\gamma_{ij}(r) = \lambda_{ij}(r) - \lambda_i \lambda_j$$

Definición:

La *función de segundo momento reducida* de un proceso puntual multivariante viene dada por

$$K_{ij}(t) = \lambda^{-1} E[N_{0ij}(t)]$$

donde $N_{0ij}(t)$ denota el número de eventos de tipo j situados a una distancia menor que t de un evento de tipo i arbitrario.

Nótese que $\lambda_{ij}(r) = \lambda_{ji}(r)$, por tanto $\gamma_{ij}(r) = \gamma_{ji}(r)$. Siguiendo el mismo argumento que en el caso unidimensional vemos que:

$$K_{ij}(t) = 2\pi(\lambda_i \lambda_j)^{-1} \int_0^t \lambda_{ij}(x) x dx \quad (2.5)$$

de donde se sigue que $K_{ij}(r) = K_{ji}(r)$.

4. DISTRIBUCIONES DE ESPACIO VACIO Y DEL VECINO MÁS PRÓXIMO

A pesar de que la K-función es un recurso muy útil para analizar un proceso puntual espacial, es importante recordar que, en la mayoría de los casos, estos procesos no son Gaussianos, por lo que una descripción de segundo orden resultará incompleta.

Para contrastar una hipótesis básica o para estudiar la bondad de ajuste de un determinado modelo paramétrico, van a ser de gran utilidad las funciones de distribución de las mínimas distancias evento-evento y evento-punto, $G(s)$ y $F(s)$ respectivamente.

Definición:

La *función de distribución del vecino más próximo* de un proceso puntual espacial es

$$G(s) = P(\text{distancia de un evento arbitrario al evento más próximo} \leq s)$$

Definición:

La *función de distribución del espacio vacío* de un proceso puntual espacial es

$$F(s) = P(\text{distancia de un punto arbitrario al evento más próximo} \leq s)$$

La forma analítica de estas funciones es bastante intratable salvo para el caso de un proceso homogéneo de Poisson plano. Esta intratabilidad de $G(s)$ y $F(s)$ limita su valor como herramientas para la construcción de modelos. Sin embargo, se pueden usar para construir tests de bondad de ajuste muy efectivos para modelos particulares y discriminan modelos que serían indistinguibles comparando las propiedades de segundo orden.

5. INDEPENDENCIA Y ETIQUETADO ALEATORIO.

Para determinar la relación entre dos tipos de eventos en un proceso puntual bivalente, podemos considerar dos hipótesis de partida.

- a) **Independencia:** los dos tipos de eventos han sido generados por dos procesos univariantes independientes.
- b) **Etiquetado aleatorio:** los dos tipos de eventos se han generado etiquetando los eventos de un proceso puntual univariante según tiradas independientes de un proceso de Bernouilli.

Cada una de estas hipótesis da lugar a distintas funciones K_{12} . En primer lugar, para dos procesos independientes con eventos tipo 1 y tipo 2, se tiene

$$K_{12}(t) = \pi t^2$$

Esta expresión se debe a que si las dos componentes del proceso son independientes, entonces los eventos tipo 1 y tipo 2 están en la misma situación respecto a un evento tipo 1, por tanto el número de eventos de tipo 2 en el disco de radio t centrado en un evento tipo 1 arbitrario, es $\lambda_2 \pi t^2$, el número esperado de eventos de tipo 2 por unidad de área multiplicado por el área del disco. Así, $K_{12}(t) = \pi t^2$.

En segundo lugar, para cualquier proceso etiquetado aleatoriamente con eventos tipo 1 y tipo 2

$$K_{11}(t) = K_{22}(t) = K_{12}(t) = K(t)$$

donde $K(t)$ es la K-función del proceso univariante sin etiquetar. Para ver esto, sea $K(t)$ la K-función del proceso unitario sin etiquetar formado por todos los eventos independientemente del tipo. Entonces, bajo etiquetado aleatorio, los procesos unitarios tipo 1 y tipo 2 son refinamientos aleatorios del proceso de partida. Así, dado que la K-función es

invariante por refinamientos aleatorios, tenemos que $K_{11}(t) = K_{22}(t) = K(t)$. Por el mismo motivo $K_{12}(t) = K(t)$.

La independencia y el etiquetado aleatorio son equivalentes cuando ambas componentes son procesos de Poisson.

TESTS DE CSR

1. CONTRATOS DE CSR

A pesar de que el interés científico de la aleatoriedad espacial completa es limitado, hay varias razones para empezar el análisis de un diseño espacial con un test de CSR: en primer lugar, rechazar esta hipótesis es un requisito mínimo previo al inicio de la construcción de un modelo para un patrón observado, en segundo lugar, estos tests se usan para explorar los datos y, por último, sirven de referencia para formular posibles alternativas a la CSR. La aleatoriedad espacial completa actúa como una hipótesis intermedia entre diseños regulares y agregados.

Teniendo en cuenta lo anterior, debemos destacar dos aspectos: el valor de los métodos gráficos, que casi siempre aportan información y que en ocasiones hacen que no sea necesario aplicar un test formal; y la importancia de combinar varios tests que indiquen la naturaleza de la desviación de CSR. Con respecto a la última observación, si queremos un único nivel de significación debemos tener en cuenta el siguiente resultado: Supongamos que los niveles de significación de k tests de CSR, no necesariamente independientes, son $p_j; j = 1, \dots, k$ y sea p_{\min} el menor de los p_j , que corresponde a la desviación más importante de la aleatoriedad espacial completa. Entonces, bajo CSR:

$$p \leq P(p_{\min} \leq p) \leq kp \quad (3.1)$$

Si los k tests son independientes, el resultado es:

$$P(p_{\min} \leq p) \leq 1 - (1 - p)^k$$

Cox (1977) apuntó que el uso de varios tests como parte de un diagnóstico solo tiene sentido práctico si cada uno examina un aspecto distinto del diseño, así la obtención de un resultado significativo para uno de los contrastes no evita el análisis de los demás.

Contrastar la aleatoriedad espacial completa es un objetivo poco ambicioso en sí mismo y debe verse como un punto de partida natural en el análisis de todo diseño espacial observado. Desde el punto de vista pedagógico, aporta una perspectiva histórica del desarrollo previo de esta materia y, a su vez, permite ilustrar aspectos generales en el conjunto más sencillo posible. Entre estos aspectos destacan el papel de los métodos de Monte Carlo, la necesidad de valorar la importancia de los métodos intuitivos y, posiblemente lo más importante, la necesidad de tener en cuenta la dependencia entre múltiples medidas del mismo diseño puntual. Antes de empezar a describir distintos tests y comentar sus ventajas e inconvenientes, se introducen los *Mosaicos de Dirichlet* y los *Tests de Monte Carlo*, que son herramientas clave en muchos contrastes.

1.1. Mosaico de Dirichlet

Dados n eventos distintos, $x_i, i = 1, \dots, n$, en una región plana A , podemos asignar a cada x_i un territorio que consiste en la parte de A que está más próxima a x_i que a cualquier otro x_j . Esta construcción, denominada tanto **Mosaico de Dirichlet** como **Mosaico de Voronoi** de los eventos en A , se incorpora a los modelos estocásticos de fenómenos naturales, como competición entre plantas, donde se asume que las plantas situadas en territorios vecinos están en competencia directa por el alimento. Para n grande, el mosaico también es la base de una solución eficiente a los problemas del cálculo de distancias entre eventos.

Cada territorio o celda es una región poligonal, excepto posiblemente en la frontera de A . Los eventos, x_i y x_j , que comparten un segmento frontera se denominan contiguos. Normalmente, cada vértice de las celdas es común a tres territorios y las líneas que comparten los pares de eventos vecinos definen una triangulación de los x_i , denominada **Triangulación de Delaunay**. Así, las fronteras de cada celda se pueden obtener como bisecciones perpendiculares de las fronteras de las triangulaciones y los vértices como los correspondientes circuncentros.

La construcción del *Mosaico de Dirichlet* y los *Triángulos de Delaunay* asociados deja de ser un ejercicio trivial a medida que n aumenta. Green y Sibson (1978) propusieron un algoritmo de eficiencia notable cuyo coste computacional aumenta a razón de $n^{1/5}$.

1.2. Tests de Monte Carlo

Incluso los modelos estocásticos más simples para diseños puntuales espaciales tienen distribuciones teóricas intratables. Por tanto, para contrastar el modelo respecto a los datos, se hará un uso extenso de los tests de Monte Carlo.

Los **tests de Monte Carlo** suponen un método extremadamente simple y útil con aplicaciones en muchas áreas de la estadística. La idea básica es la siguiente:

Sea u_1 el valor observado de un estadístico U y sean $u_i, i = 2, \dots, s$ los valores de la distribución de U para muestras aleatorias independientes generadas bajo la hipótesis nula H_0 . Sea $u_{(j)}$ el j -ésimo mayor de los $u_i, i = 1, \dots, s$. Entonces, bajo \mathcal{H}_0 :

$$P(u_1 = u_{(j)}) = s^{-1}; j = 1, \dots, s$$

Rechazar H_0 cuando u_1 ocupa la k -ésima posición o una mayor proporciona un test unilateral exacto de nivel de significación k/s . Estamos asumiendo que los u_i son todos distintos, si no tendríamos una clasificación ambigua. La extensión de este test al caso bilateral es trivial.

Hope (1968) mostró, a través de ejemplos, que la pérdida de eficiencia debida a la implementación de Monte Carlo es leve, por lo que no necesitamos que s sea muy grande. Para un test unilateral y un nivel del 5% es suficiente tomar $s = 100$.

La pérdida de eficiencia de este método está relacionada con la investigación de las “regiones críticas borrosas” de Marriott, que surge al haber valores de u_1 que son significativos para un test clásico pero no lo son para el de Monte Carlo, y viceversa. Sea $F(u)$ la distribución (desconocida) de U bajo H_0 . Para un nivel de significación del 5% ($s = 20k$), la probabilidad de rechazar H_0 , dado $U = u_1$, es:

$$p(u_1) = \sum_{r=0}^{k-1} \binom{s-1}{r} (1-F(u_1))^r (F(u_1))^{s-1-r} \quad (3.2)$$

Para un test clásico, representado aquí por el límite cuando $s \rightarrow \infty$, $p(u_1)$ es 1 ó 0 según $F(u_1)$ sea mayor o menor que 0,95.

En la aplicación práctica de los métodos de Monte Carlo, las muestras aleatorias se sustituyen por muestras pseudo-aleatorias.

Otro punto crítico de los tests de Monte Carlo es que fomentan el “data-dredging”, dado que el usuario puede elegir un estadístico U que se centre en una característica aparentemente notable de los datos que en realidad no lo es. Es obvio que los resultados significativos obtenidos a partir de estadísticos patológicos no son válidos.

El uso de los tests de Monte Carlo está restringido al contraste de una hipótesis nula simple. Para contrastar una hipótesis nula compuesta necesitamos construir muestras pseudo-aleatorias condicionadas a los valores de estadísticos suficientes para los parámetros desconocidos, pero este método no es aplicable en la práctica.

Cuando se puede aplicar la teoría asintótica, los tests de Monte Carlo proporcionan una alternativa exacta para muestras pequeñas y también se utilizan como control de la aplicabilidad de la teoría asintótica. Si los resultados de los tests clásicos y de Monte Carlo coinciden, no hay pérdida de información. La situación contraria se debe a que los tests clásicos parten de alguna hipótesis incorrecta sobre la distribución.

2. TESTS DE DISTANCIAS

2.1. Distancia entre eventos

Un posible método de descripción del diseño de n eventos en una región A es la distribución empírica de las $\frac{1}{2}n(n-1)$ distancias entre eventos, t_{ij} . La distribución teórica correspondiente a la distancia T entre dos eventos distribuidos uniforme e independientemente en A depende del tamaño y la forma de la región, pero tiene una expresión cerrada en casos comunes, como son las regiones circulares o cuadradas.

Si A es el cuadrado unitario, la distribución de T es:

$$H(t) = \begin{cases} \pi t^2 - 8t^3/3 - t^4/2 & ; 0 \leq t \leq 1 \\ 1/3 - 2t^2 - t^4/2 + 4(t^2 - 1)^{1/2} (2t^2 + 1)/3 \\ \quad + 2t^2 \sin^{-1}(2t^2 - 1) & ; 1 < t \leq \sqrt{2} \end{cases} \quad (3.3)$$

mientras que para el círculo de radio 1 la expresión será:

$$H(t) = 1 + \pi^{-1} \left\{ 2(t^2 - 1) \cos^{-1}(t/2) - r(1 + t^2/2) \sqrt{1 - t^2/4} \right\} \quad (3.4)$$

para todo $0 \leq t \leq 2$.

A continuación se desarrolla un test de CSR basado en las distancias entre eventos, el método descrito se puede aplicar a otros estadísticos.

Supongamos $H(t)$ conocida para una región dada, A . Se calcula la función de distribución empírica de las distancias entre eventos. Esta función, $\hat{H}_1(t)$, representa la proporción de distancias entre eventos, t_{ij} menores o iguales que t , así

$$\hat{H}_1(t) = \left\{ \frac{1}{2}n(n-1) \right\}^{-1} \#(t_{ij} \leq t)$$

Tras obtener $\hat{H}_1(t)$, se representa frente a $H(t)$. Si los datos son compatibles con CSR, el gráfico será prácticamente lineal. Para calcular el nivel de significación o la distancia a la linealidad el procedimiento habitual es calcular la distribución empírica de $\hat{H}_1(t)$ bajo CSR, pero esto se complica debido a la dependencia de las distancias entre eventos con un punto final común. Por tanto, se procede del modo siguiente: calculamos funciones de distribución empíricas $\hat{H}_i(t)$, $i = 2, \dots, s$ para s simulaciones independientes de n eventos distribuidos uniforme e independientemente en A , y se definen las envolturas superior e inferior:

$$U(t) = \max \{ \hat{H}_i(t) \} \quad (3.5)$$

$$L(t) = \min \{ \hat{H}_i(t) \} \quad (3.6)$$

con $2 \leq i \leq s$. Estas envolturas se pueden representar frente a $H(t)$ y, bajo CSR, verifican que, para cada t

$$P(\hat{H}_i(t) > U(t)) = P(\hat{H}_i(t) < L(t)) = s^{-1} \quad (3.7)$$

Las envolturas simuladas facilitan la interpretación del gráfico de $\hat{H}_1(t)$ frente a $H(t)$. También es posible aplicar múltiples tests de Monte Carlo, veamos dos ejemplos

- (i) Se elige un valor t_0 y se define $u_i = \hat{H}_i(t_0)$, la posición de u_i entre los u_i es la base de un test, porque bajo CSR todas las posiciones de u_i en el ranking son equiprobables.
- (ii) Sea u_i una medida de discrepancia entre $\hat{H}_1(t)$ e $H(t)$ en todo el rango de t , por ejemplo:

$$u_i = \int (\hat{H}_1(t) - H(t))^2 dt \quad (3.8)$$

de nuevo, se aplicaría un test basado en la posición de u_i .

El primer método solo tiene sentido cuando se puede elegir un t_0 adecuado para cada problema concreto. El segundo tiene la ventaja de la objetividad, pero proporciona un test débil para los contrastes basados en distancias entre eventos.

Cuando no se conoce la distribución teórica de $H(t)$ en la región A se puede aplicar el test sustituyendo en (3.8) $H(t)$ por

$$\tilde{H}_i(t) = (s-1)^{-1} \sum_{j \neq i} \hat{H}_j(t)$$

ahora los u_i ya no son independientes bajo CSR, pero se pueden intercambiar y se mantiene la propiedad de que todas las posición de u_i en la clasificación son equiprobables. El procedimiento gráfico consiste, de nuevo, en representar $\hat{H}_1(t)$, $U(t)$ y $L(t)$ frente a $\tilde{H}_1(t)$. Nótese que, como $\tilde{H}_1(t)$ incluye solo las simulaciones bajo CSR y no los datos observados, es un estimador insesgado de $H(t)$ bajo la hipótesis nula.

2.2. Distribución de la distancia al vecino más próximo

Para n eventos en la región A , sea y_i la distancia del i -ésimo evento al evento más próximo en A . Las y_i se denominan distancias al vecino más próximo e incluyen medidas duplicadas entre pares de vecinos recíprocos. Podemos calcular la función de distribución empírica, $\hat{G}_1(t)$, de las distancias al vecino más próximo de modo análogo al visto para obtener $\hat{H}_1(t)$. Así:

$$\hat{G}_1(t) = n^{-1} \#(y_i \leq t)$$

En muchas aplicaciones prácticas hay interacción entre eventos aunque solo sea a pequeña escala. Por ejemplo, los árboles pueden competir por la luz solar o por alimento en una región delimitada por su copa o por su raíz. En estos casos, las distancias al vecino más próximo proporcionan una medida objetiva de la concentración de pequeñas distancias entre eventos cuando no se puede especificar la distancia mínima.

Bajo CSR, la distribución teórica de la distancia al vecino más próximo, Y , depende de n y de A y no tiene expresión explícita debido a las complicaciones derivadas del efecto frontera. Ignorando este efecto tenemos una aproximación, si $|A|$ denota la superficie de A , entonces $\pi y^2 |A|^{-1}$ es la probabilidad, bajo CSR, de que la distancia de un evento a otro arbitrario sea inferior a y . Si la localización de los eventos es independiente, la distribución aproximada de Y es:

$$G(y) = 1 - \left(1 - \pi y^2 |A|^{-1}\right)^{n-1}$$

Para n grande, tomando $\lambda = n |A|^{-1}$ se tiene la siguiente aproximación:

$$G(y) = 1 - \exp(-\lambda \pi y^2); \quad y \geq 0 \quad (3.9)$$

La función de distribución empírica se puede comparar con las envolturas superior e inferior de las distribuciones empíricas de simulaciones bajo CSR $\hat{G}_i(y)$; $i = 2, \dots, s$. En un diseño agregada, $G(y)$ estará por encima de la envoltura superior, mientras que si $G(y)$ está por debajo de $L(y)$ estamos ante una distribución regular de los datos. Una posible base para un test de Monte Carlo incluye la media muestral de las distribuciones de distancia al vecino más próximo, o

$$u_i = \int \left(\hat{G}_i(t) - \tilde{G}_i(t) \right)^2 dt \quad (3.10)$$

donde

$$\tilde{G}_i(t) = (s-1)^{-1} \sum_{j \neq i} \hat{G}_j(t)$$

definida de modo análogo a $\tilde{H}_1(t)$ en la sección anterior. Clark y Evans (1954) propusieron un test basado en la media muestral, \bar{y} , de las distancias al vecino más próximo, que no permitía dependencia entre ellas. Una posible ventaja del test basado en \bar{y} es que, al igual que en los tests de distancia mínima entre eventos no es necesaria la simulación. Donnelly (1978) mostró que, para una buena aproximación, la distribución de \bar{y} bajo CSR es normal con media y varianza:

$$E[\bar{y}] = 0.5(n^{-1}|A|)^{1/2} + (0.051 + 0.042n^{-1/2})n^{-1}P \quad (3.11)$$

$$Var(\bar{y}) = 0.070n^{-2}|A| + 0.037(n^{-5}|A|)^{1/2}P \quad (3.12)$$

donde P es el perímetro de A . Valores de \bar{y} significativamente grandes o pequeños indican diseño agregado o regularidad, respectivamente. Para regiones con formas poco regulares, la aproximación se deteriora, causando una pérdida de eficacia del test, por lo que es necesario implementar un método de Monte Carlo. Seguimos manteniendo que la inspección del gráfico de las funciones de distribución empíricas es, al menos, tan importante como un contraste formal.

Un método obvio para calcular la distancia al vecino más próximo es buscar entre todas las distancias entre eventos. Para n suficientemente grande, será más eficiente construir el Mosaico de Dirichlet de los n eventos y buscar las distancias al vecino más próximo dentro del mosaico. Así, usamos que para cualquier n suficientemente grande, cada evento es vecino de otros seis eventos y uno de ellos será el vecino más próximo. Por tanto, solo hay que calcular una pequeña fracción de las distancias entre eventos. La búsqueda directa es más eficiente para $n < 500$, pero según aumenta n pierde eficiencia.

2.3. Distribución de la función de espacio vacío.

Un análisis similar al anterior usa las distancias, x_i , de m puntos aleatorios en A al evento más próximo. La función de distribución empírica $\hat{F}_1(x) = m^{-1} \#(x_i \leq x)$ mide los espacios vacíos en A , pues $1 - \hat{F}_1(x)$ es un estimador del área, $|B(x)|$, de la región $B(x)$ formada por todos los puntos de A a una distancia mayor que x de cualquiera de los n eventos de A . Repitiendo el argumento que nos llevaba a (3.9) se tiene que, bajo CSR:

$$F(x) = 1 - \exp(-\lambda \pi x^2), \quad x \geq 0$$

aproximadamente, donde $\lambda = n|A|^{-1}$.

Lotwick (1981) describe un algoritmo, basado en el Mosaico de Dirichlet de Green-Sibson, para calcular $|B(x)|$ cuando A es un rectángulo. En la práctica tomando m puntos en una red regular $k \times k$ se obtiene una aproximación adecuada si k es suficientemente grande. La elección de un k adecuado depende de la localización de los n eventos en A y de para que se use el estimador. Diggle y Matérn (1981) recomiendan tomar $k \approx \sqrt{n}$ para estimar $F(x)$ a partir de realizaciones simuladas de un proceso puntual, en este contexto se puede elegir tanto el número de puntos por realización como el número de simulaciones.

En la actualidad no es necesario estudiar el esfuerzo computacional del cálculo de $\hat{F}_1(x)$ y tampoco hay razones estadísticas para limitar la elección de k . Aun así, debemos recordar que cuanto mayor sea k más suave será la curva $\hat{F}_1(x)$, esta precisión estadística está limitada por el número de eventos.

De modo análogo al caso de distancia al vecino más próximo se puede representar $\hat{F}_1(x)$ junto con las envolturas superior e inferior para diseños simulados bajo CSR. En este caso $\hat{F}_1(x) > U(x)$ indica regularidad y $\hat{F}_1(x) < L(x)$ indica una distribución con datos agregados. También podemos construir un test de Monte Carlo basado en el estadístico

$$u_i = \int (\hat{F}_i(t) - \tilde{F}_i(t))^2 dt \quad (3.13)$$

3. TESTS DE CUADRANTES

3.1. Cuadrat couts

Una alternativa a los contrastes basadas en distancias consiste en dividir A en m subregiones o cuadrados de igual área y utilizar el contador de número de eventos en cada cuadrado para el test de CSR. Las subregiones se eligen de forma arbitraria. Supongamos que A es el cuadrado unidad y está dividido en una red $k \times k$ de subregiones cuadradas, por tanto $m = k^2$. Sean $n_i, i = 1, \dots, m$ los contadores de eventos en los cuadrados resultantes de esta partición de A y denotamos $\bar{n} = n/m$ a la media muestral de los n_i . Un estadístico obvio para contrastar CSR en A es el criterio χ^2 de Pearson

$$X^2 = \sum_{i=1}^m (n_i - \bar{n})^2 / \bar{n} \quad (3.14)$$

Donde m es el número de regiones y $\bar{n} = \sum_{i=1}^m n_i / m$, es el número esperado de eventos por región si la intensidad es homogénea. Bajo la hipótesis nula, este estadístico tiene distribución χ_{m-1}^2 , asumiendo que \bar{n} no es muy pequeño, en la práctica se pide $\bar{n} > 5$

En este contexto, la hipótesis de CSR puede fallar tanto por una distribución no uniforme de eventos en A , intensidad no constante, como por la existencia de relaciones de dependencia entre eventos. En particular, valores de X^2 suficientemente grandes o pequeños indican, respectivamente, distribuciones agregadas o regulares.

3.2. Escalas en el diseño

Greig-Smith (1952) propuso este método para analizar datos a partir de sus contadores en una red de cuadrados contiguos. Se calcula el índice de dispersión, cociente entre varianza y media, de los contadores para la red de partida y para redes sucesivas obtenidas por combinación de cuadrados en bloques 2×2 , 4×4 , etc. Se representa el índice de dispersión frente al tamaño de los bloques y los picos y depresiones se interpretan como una evidencia de *escalas en el diseño* (datos agregados o regularidad, respectivamente).

Este sistema de análisis tiene su origen en la ecología vegetal, donde se hizo muy popular. Un posible problema de este método es que los test formales para determinar la significación de picos y depresiones no se pueden evaluar para una secuencia de índices a diferentes escalas. Según el desarrollo hecho en la sección anterior, cada índice es proporcional a un estadístico chi-cuadrado para contrastar que los eventos son muestras independientes de la distribución uniforme sobre la región de estudio A .

Mead (1974) soluciona este defecto formal del proceso de Greig-Smith estableciendo una serie de tests independientes para el modelo a diferentes escalas. Este proceso requiere hacer sucesivas particiones de los datos en 1, 4, 16,... bloques, formado cada uno por 16 contadores en una red 4×4 . En cada etapa, la hipótesis nula es que, en cada bloque, el conjunto de contadores en los cuatro subbloques 2×2 asociados es una selección aleatoria entre las $16!/(4!)^5 = 2.627.625$ posibilidades equiprobables, como implica la aleatoriedad espacial completa. El estadístico sugerido por Mead es la suma absoluta de las seis diferencias entre los cuatro contadores que hay en el bloque, sumados en todos los bloques. Un valor significativamente grande de este estadístico implica que, dentro de los bloques, los contadores en cuadros vecinos son relativamente distintos, esto se puede interpretar como clústers a una escala adecuada. Un valor significativamente pequeño, indica contadores relativamente similares en cuadrados vecinos, esto es más difícil de interpretar. Una situación extrema puede ser un diseño tipo tablero de ajedrez que alterna valores altos y bajos, pero esta configuración no es realista. Una vez elegido el estadístico U , el test se puede implementar mediante muestreo de Monte Carlo de la distribución de U bajo la hipótesis nula. Se tiene independencia de los tests a diferentes escalas porque los bloques de contadores aleatorios 2×2 a una escala pasan a ser los contadores fijos 4×4 en el siguiente paso.

3.3. Escalas de dependencia

Besag (1978) describe un test bivariante análogo al de Mead para investigar lo que denominaremos *escalas de dependencia* entre dos diseños. Ahora la unidad básica es un bloque 2×2 de cuadrados, cada cuadrado proporciona un par de contadores, uno para cada especie. Besag sugiere contrastar la hipótesis de que los dos conjuntos de cuatro contadores en cada bloque son independientes usando como estadístico el rango de coeficientes de correlación de Spearman para los dos conjuntos de cuatro contadores y sumar en todos los bloques. El test se implementa por Monte Carlo, generando en cada bloque contadores aleatorios. Cada bloque 2×2 se integra en un único cuadrado y así sucesivamente, para construir una secuencia de tests independientes para contrastar la hipótesis de independencia entre los dos conjuntos de contadores en cada bloque.

4. CONTRASTE BASADO EN LA K-FUNCIÓN:

Las características de segundo orden definen la dependencia entre pares de puntos del proceso. La función de segundo momento reducida, o K de Ripley, de un proceso puntual estacionario e isotrópico viene dada por:

$$K(t) = \lambda^{-1} E[N_0(t)]$$

Donde $N_0(t)$ denota el número de eventos situados a una distancia menor que t de un evento arbitrario. Así, $\lambda K(t)$ se puede interpretar como el número esperado de eventos a una distancia menor que t de un evento dado.

Para un proceso homogéneo de Poisson en \mathbb{R}^2 , $K(t) = \pi t^2$. Así, al comparar el valor de $K(t)$ para un proceso puntual observado con el valor que tiene para el proceso homogéneo de Poisson tenemos que:

$$K(t) > \pi t^2 \Rightarrow \text{proceso agregado}$$

$$K(t) < \pi t^2 \Rightarrow \text{proceso regular}$$

Para un diseño espacial observado en una región A, se puede estimar su K- función mediante:

$$\hat{K}_1(t) = \frac{|A|}{n(n-1)} \sum_{x \neq y} I(|y-x| \leq t) w_A(x, y)$$

Donde $w_A(x, y)$ es un corrector del efecto frontera.

Una vez estimada la K-función para el proceso observado, se aplica un contraste de CSR similar a los propuestos para F y G. Es decir, se calcula la K-función empírica

$\hat{K}_i(t)$, $i = 2, \dots, s$, para $s-1$ simulaciones independientes de un proceso homogéneo de Poisson de tamaño n en A , y se definen las envolturas superior e inferior:

$$U(t) = \max_{i=2, \dots, s} \{ \hat{K}_i(t) \}$$

$$L(t) = \min_{i=2, \dots, s} \{ \hat{K}_i(t) \}$$

Representar gráficamente $\hat{K}_1(t)$ junto con las envolturas superior e inferior, proporciona un test gráfico de CSR. Así, valores de la K-función empírica mayores que la envoltura superior o menores que la inferior indican, respectivamente, diseño agregado o regular.

Por comodidad a la hora de interpretar los datos, se puede utilizar la función:

$$L(t) = \sqrt{K(t)/\pi}$$

Pues, para un proceso homogéneo de Poisson se tiene que $L(t) = t$.

5. CONTRASTES DE CSR PARA LOS FUEGOS REGISTRADOS.

La base de datos consiste en 4764 fuegos forestales registrados entre 1991 y 2008. Estos datos se pueden clasificar según su causa o tipo. Según la causa, tenemos 4153 fuegos intencionados, 279 provocados por causas naturales, 197 de origen desconocido y otras categorías que no vamos a tener en cuenta debido a las pocas observaciones registradas en cada una de ellas. Respecto al tipo, se han observado 2909 conatos, 1169 quemas y 686 incendios. En esta sección se han aplicado distintos contrastes de CSR, tanto al total de los datos como a cada uno de las categorías indicadas.

El primer objetivo del análisis es contrastar la homogeneidad de los diseños. Con este fin, se ha aplicado un test chi-cuadrado de independencia para determinar si la intensidad de los eventos en una región es significativamente distinta de la observada en otras regiones. Se ha dividido el área de interés en subregiones y se han contado los eventos observados en cada una. Dado que la región de estudio tiene forma irregular, para tratar de obtener subregiones de tamaños similares, se han realizado solo dos divisiones verticales. Como se ha visto en la sección 3, el estadístico de contraste es:

$$X^2 = \sum_{i=1}^m (n_i - \bar{n}_i)^2 / \bar{n}_i$$

Donde $\bar{n}_i = n \times w_i$, siendo w_i la proporción del i -ésimo cuadrante rectangular que está dentro de la región de observación, así se corrige el efecto frontera producido al no tener una región de estudio cuadrada.

Tabla 1: Test chi-cuadrado.

		Número	4x2	8x2
Total		4764	821.2179	1060.203
Causa	Intencionado	4153	940.7118	1174.127
	Natural	279	38.8212	48.4056*
	Desconocida	197	7.4746	24.5604*
Tipo	Incendio	686	61.3163,	144.3209
	Conato	2909	555.2473	660.7257
	Quema	1169	308.5264	390.0507

(*) Valores esperados inferiores a 5, test no fiable.

En la tabla 1 se muestran los valores del estadístico de contraste obtenidos para cada uno de los procesos observados. Se han hecho 4 y 8 divisiones horizontales para evaluar, respectivamente, la homogeneidad global y la homogeneidad local. Si comparamos los resultados obtenidos para cada uno de los patrones observados con los valores bajo la hipótesis nula $\chi^2_{7,0.01} = 18.475$ y $\chi^2_{15,0.01} = 30,578$, vemos que los primeros, en la mayoría de los casos, son notablemente superiores a los valores críticos. Por tanto, podemos concluir que, salvo en el caso de fuegos con causa desconocida, los diseños observados no tienen distribución homogénea.

Para completar este análisis se han aplicado también F-test, G-test y K-test tanto para todos los fuegos como para las distintas categorías según tipo y causa. Los gráficos correspondientes a todos los fuegos registrados, gráfico 1, confirman la ausencia de CSR indicada por el test chi-cuadrado y muestran evidencias de diseño agregado en el proceso.

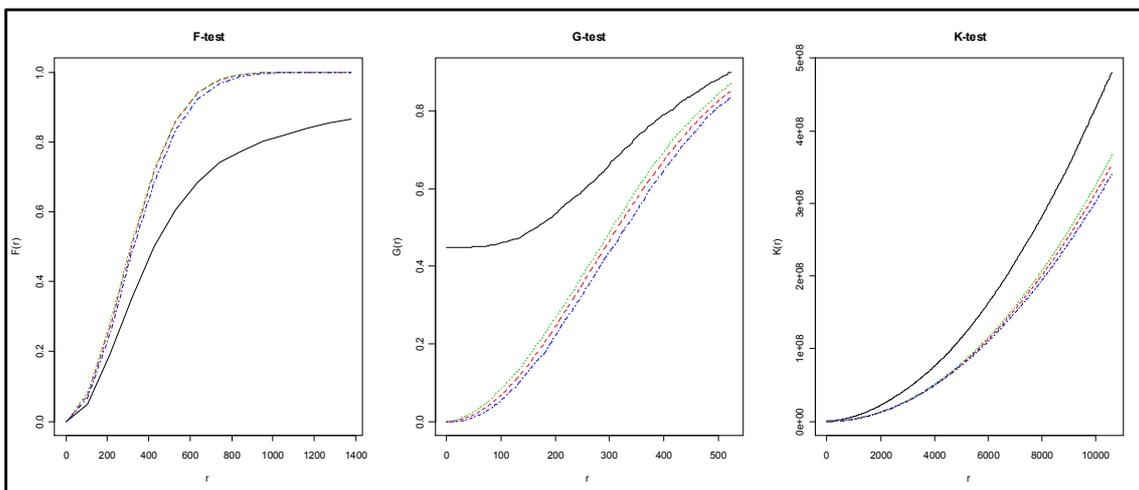


Gráfico 1: F-test, G-test y K-test para todos los fuegos. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

Si nos fijamos en los gráficos obtenidos para los fuegos registrados entre 1991 y 2008 en función de la causa que los provocó, gráficos 2-4, tanto para fuegos intencionados como para aquellos debidos a causas naturales o cuya causa se desconoce, los tres tests nos llevan a suponer un diseño agregado de los datos.

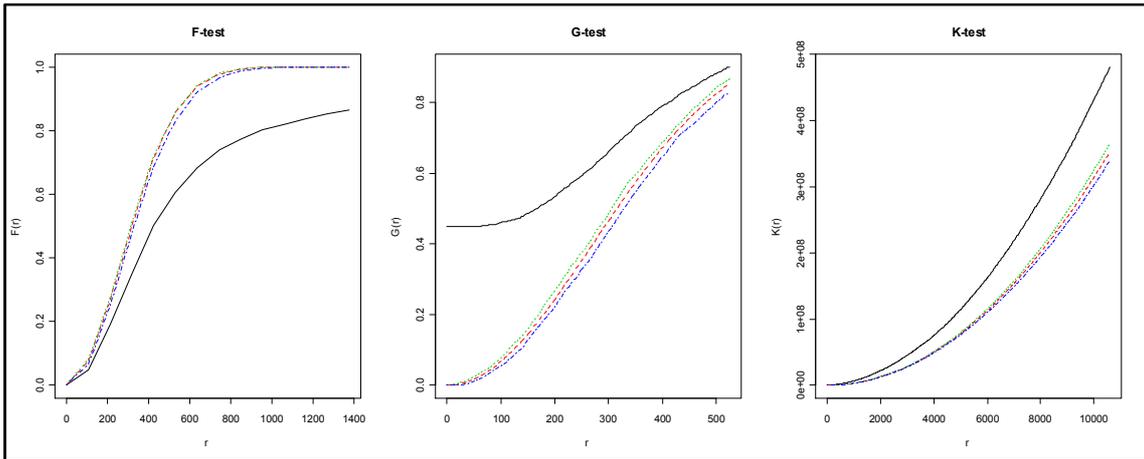


Gráfico 2: F-test, G-test y K-test para los fuegos intencionados. Distribución empírica (negro), $L(t)$ (azul), $U(t)$ (verde) distribución media de las simulaciones (rojo)

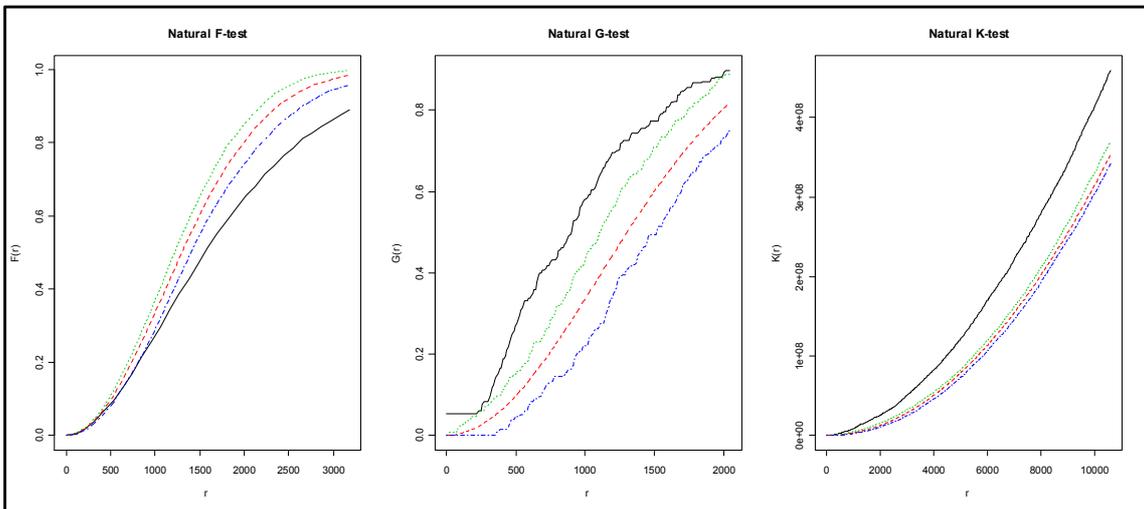


Gráfico 3: F-test, G-test y K-test para los fuegos provocados por causas naturales. Distribución empírica (negro), $L(t)$ (azul), $U(t)$ (verde) distribución media de las simulaciones (rojo)

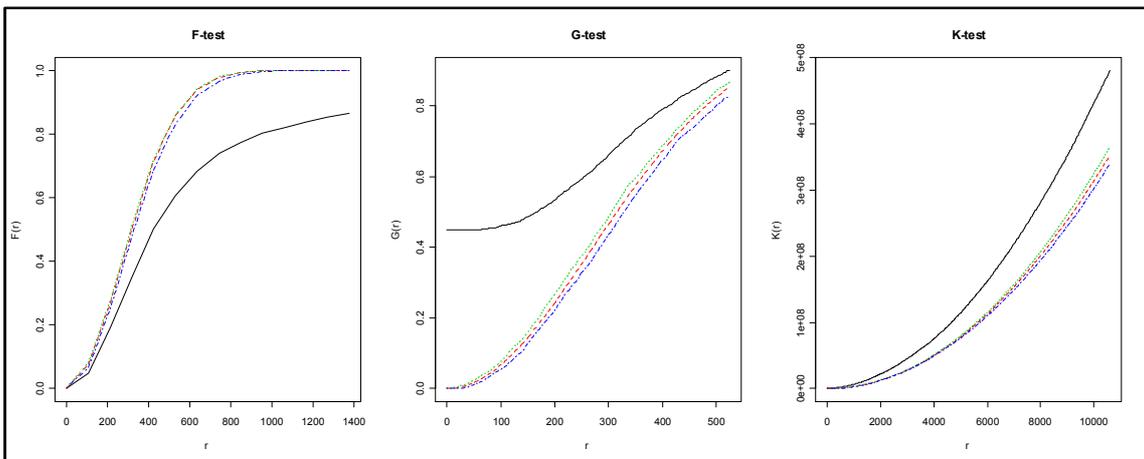


Gráfico 4 F-test, G-test y K-test para los fuegos con causa desconocida. Distribución empírica (negro), $L(t)$ (azul), $U(t)$ (verde) distribución media de las simulaciones (rojo)

Por último si analizamos los fuegos en función de su tipo, gráficos 5-7, los tres test de rechazan la CSR y vuelven a mostrar evidencias de diseño agregado para conatos, incendios y quemas.

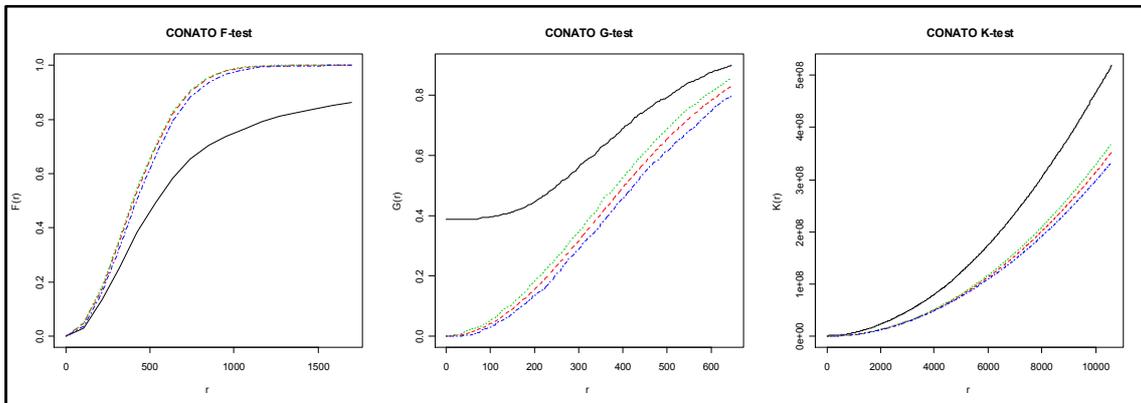


Gráfico 5: F-test, G-test y K-test para conatos. Distribución empírica (negro), $L(t)$ (azul), $U(t)$ (verde) distribución media de las simulaciones (rojo)

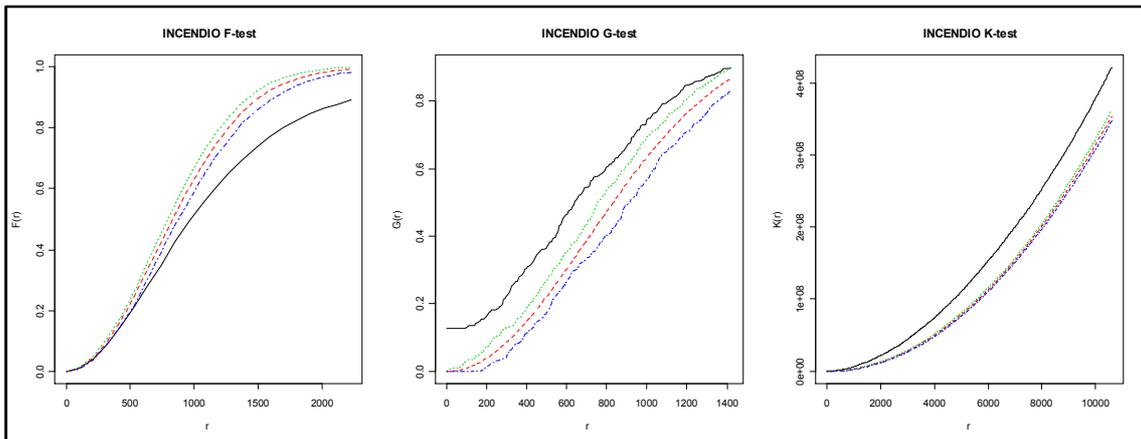


Gráfico 6: F-test, G-test y K-test para incendios. Distribución empírica (negro), $L(t)$ (azul), $U(t)$ (verde) distribución media de las simulaciones (rojo)

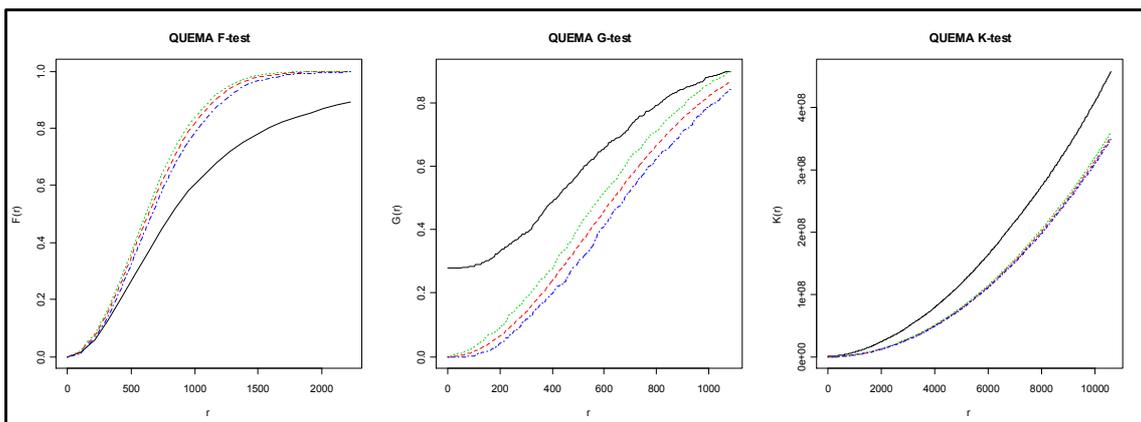


Gráfico 7: F-test, G-test y K-test para quemas. Distribución empírica (negro), $L(t)$ (azul), $U(t)$ (verde) distribución media de las simulaciones (rojo)

Teniendo en cuenta tanto los resultados formales obtenidos mediante el test chi-cuadrado, como los contrastes basados en la distribución de espacio vacío, F-test, distribución del vecino más próximo, G-test, y función de segundo momento reducida, K-test, podemos concluir que ninguno de los procesos que estamos estudiando verifican la condición de aleatoriedad espacial completa. Además los test gráficos sugieren que la discrepancia respecto a la aleatoriedad espacial completa se debe a la existencia de un diseño agregado, es decir formación de clústers.

MODELOS

1. CSR: LOS PROCESOS HOMOGÉNEOS DE POISSON

Los procesos homogéneos de Poisson planos, que habitualmente se denominan simplemente procesos de Poisson, son la base a partir de la que se construye la teoría de los procesos puntuales espaciales. Representan el mecanismo estocástico más sencillo que puede generar un diseño puntual espacial y se aplican como modelo ideal de CSR que, a pesar de no alcanzarse en la práctica, resulta muy útil como descripción aproximada de algunos patrones observados.

Los procesos de Poisson están definidos por las siguientes propiedades:

Definición:

Un proceso puntual es un proceso homogéneo de Poisson plano de intensidad λ si:

PP1) El número de eventos en una región A , $N(A)$ sigue una distribución de Poisson con media $\lambda|A|$ donde $|A|$ denota el área de A .

PP2) Dado $N(A) = n$ los n eventos forman una muestra aleatoria de la distribución uniforme en A .

PP3) Para dos regiones disjuntas A y B , las variables aleatorias $N(A)$ y $N(B)$ son independientes.

Sea $C = A \cup B$, siendo A y B disjuntas. Sean $p = |A|/|C|$ y $q = 1 - p = |B|/|C|$. Entonces PP2 aplicada a C implica:

$$P[N(A) = x, N(B) = y / N(C) = n] = \binom{x+y}{x} p^x q^y$$

para enteros $0 \leq x \leq n$ e $y = n - x$. PP1 nos da la distribución incondicional de $N(A)$ y $N(B)$

$$\begin{aligned} P[N(A) = x, N(B) = y] &= \binom{x+y}{x} p^x q^y \left(e^{-\lambda|C|} (\lambda|C|)^n / n! \right) = \\ &= \left(e^{-\lambda|A|} (\lambda|A|)^x / x! \right) \left(e^{-\lambda|B|} (\lambda|B|)^y / y! \right) \end{aligned} \quad (4.1)$$

para todo $x, y \geq 0$. La expresión anterior establece PP3 y muestra también que $N(A)$ y $N(B)$ tienen la distribución indicada por PP1.

Si PP2 es cierta para una región C , es inmediato que se verifica para toda subregión. Recíprocamente, la propiedad aditiva de las variables aleatorias independientes de Poisson, X e Y , y la distribución condicional binomial de X , dado $X+Y$, nos dan PP1 y PP2, respectivamente, para cualquier región formada por la unión de dos disjuntas que verifican ambas propiedades.

El parámetro λ del proceso de Poisson es su intensidad. Por PP3 tenemos:

$$\lambda_2(t) = \lambda^2, \quad t > 0 \quad (4.2)$$

de donde

$$K(t) = 2\pi\lambda^{-2} \int_0^t \lambda_2(x) x dx = \pi t^2, \quad t > 0 \quad (4.3)$$

Directamente de PP1, se obtiene la varianza

$$Var[N(A)] = \lambda|A| \quad (4.4)$$

Las funciones de espacio vacío y vecino más próximo, $F(s)$ y $G(s)$, son idénticas, puesto que la existencia de un evento en un punto determinado, x_0 , no afecta a la distribución del resto de eventos en un disco centrado en x_0 . Por PP1 se deduce

$$F(t) = G(t) = P[N(\pi t^2) > 0] = 1 - \exp(-\pi\lambda t^2), \quad t > 0 \quad (4.5)$$

Para simular una realización parcial de un proceso de Poisson en A condicionando a un número fijo de eventos $N(A)$, se generan eventos uniforme e independientemente en A . Si la forma de la región A es compleja, se realiza el proceso de simulación en una región más grande con forma más simple, como un rectángulo o un disco, y se consideran solo los eventos que caen en A . Hsuan (1979) proporcionó un algoritmo para simular directamente eventos uniformemente distribuidos en un polígono arbitrario.

Si queremos que $N(A)$ varíe aleatoriamente, se utiliza el proceso anterior precedido de la simulación de $N(A)$ según la distribución de Poisson correspondiente. En algunas implementaciones, la simulación directa de $N(A)$ tiene un coste computacional elevado: Lewis & Shedler (1979) propusieron un método alternativo que se puede utilizar cuando A es un rectángulo, por ejemplo $(0, a) \times (0, b)$. Este método está basado en el hecho de que la localización de la coordenada X de cada evento en la banda $0 \leq y \leq b$ forma un proceso de Poisson con intensidad λb , por tanto se tiene que la diferencia entre coordenadas X sucesivas

son realizaciones independientes de una variable aleatoria exponencial con función de distribución

$$F(v) = 1 - \exp(-\lambda bv); \quad v \geq 0$$

Las coordenadas Y correspondientes están independientemente distribuidas según la distribución uniforme en $(0, b)$. Este método genera directamente (x_i, y_i) en orden creciente de las coordenadas X y termina cuando la última coordenadas X supera a .

2. PROCESOS DE POISSON CON CLÚSTERS

Los procesos de Poisson con clústers son modelos que incorporan explícitamente la formación de grupos y, por tanto, proporcionan la base adecuada para modelar patrones puntuales con eventos agrupados. Su definición parte de tres postulados:

Definición:

Un proceso puntual es un proceso de Poisson con clústers si:

PCP1) Los eventos principales forman un proceso de Poisson homogéneo con intensidad ρ .

PCP2) Cada evento principal produce un número aleatorio, S , de descendientes, que se generan de forma independiente e idénticamente distribuidos, para cada evento principal, según las distribuciones de probabilidad $p_s, s = 1, 2, \dots$

PCP3) Las posiciones de cada descendiente con respecto a su antecesor están independiente e idénticamente distribuidas según una función de distribución de probabilidad bivalente $h(\cdot)$.

Por convenio, el diseño final está formado solo por la superposición de descendientes de todos los eventos principales.

Los procesos de Poisson con clústers así definidos, son estacionarios con intensidad $\lambda = \rho\mu$, donde $\mu = E[S]$, son isotrópicos cuando se toma en PCP3 una función de distribución de probabilidad radialmente simétrica.

Para expresar las características de segundo orden, se define

$$h_2(x) = \int h(x)h(x-y) dx$$

como la función de distribución de probabilidad de la distancia entre la posición de dos descendientes del mismo evento principal y $H_2(\cdot)$ la distribución acumulada correspondiente.

Si se considera un evento arbitrario en un grupo de tamaño S , el número esperado de otros eventos del mismo grupo situados a una distancia inferior a t es $(S-1)H_2(t)$. La distribución

de probabilidad del tamaño del clúster al que pertenece un evento dado se obtiene a partir de la distribución de los tamaños de los grupos, $p(s)$, y viene dada por $p^*(s) = sp(s)/\mu$, $s = 1, 2, \dots$. La media de esta distribución nos da el número esperado de eventos relacionados con nuestro evento arbitrario situados a una distancia menor que t de él, esto es $E[S(S-1)]H_2(t)/\mu$.

Consideremos ahora el número esperado de eventos no relacionados, es decir los eventos que pertenecen a otros grupos, situados a una distancia menor que t de nuestro evento arbitrario. Por PCP1, la localización de estos eventos no depende del evento de partida, entonces el número esperado será simplemente $\lambda\pi t^2$.

Teniendo en cuenta la contribución de los dos tipos de eventos, se obtiene

$$\lambda K(t) = \lambda\pi t^2 + E[S(S-1)]H_2(t)/\mu$$

Y dividiendo por $\lambda = \rho\mu$

$$K(t) = \pi t^2 + E[S(S-1)]H_2(t)/(\rho\mu^2) \quad (4.6)$$

Diferenciando la expresión anterior, y teniendo en cuenta la relación entre $\lambda_2(t)$ y $K(t)$, se llega a

$$\lambda_2(t) = \lambda^2 + \rho E[S(S-1)]h_2(t) \quad (4.7)$$

El segundo termino de la expresión de $K(t)$ es monótono decreciente y no negativo, y $K(t) - \pi t^2$ aproxima la constante $c = E[S(S-1)]H_2(t)/(\rho\mu^2)$ cuando $t \rightarrow \infty$. Si S tiene una distribución de Poisson, $c = \rho^{-1}$. Este resultado sugiere un método muy útil para detectar cuando un PCP es un modelo razonable para un patrón observado o si es adecuado para una primera estimación de parámetros.

La varianza de la distribución de un proceso de Poisson con clústers será

$$Var[N(A)] = \rho\mu|A| + \rho E[S(S-1)] \int_A \int_A h_2(x-y) dx dy$$

Bajo isotropía, podemos calcular la expresión general de la función de distribución del vecino más próximo. Sea $q(x, y)$ la probabilidad de no tener ningún descendiente a una distancia menor que x del origen, para un evento principal a una distancia y del origen. Entonces la función de distribución de la distancia de un punto al evento más próximo será

$$F(x) = 1 - \exp\left(-2\pi\rho \int_0^\infty (1 - q(x, y)) y dy\right)$$

Como la localización de los eventos principales es independiente, la distribución de la distancia al vecino más próximo es

$$G(y) = 1 - (1 - F(y))q^*(y)$$

donde $q^*(.)$ denota la probabilidad de no tener otro descendiente del mismo evento principal a una distancia menor que y de un evento arbitrario.

Para simular un proceso de Poisson con clústers en una región rectangular, por ejemplo $A = (0, a) \times (0, b)$, los eventos principales en A se generan como una realización parcial de un proceso de Poisson adecuado. Los descendientes están relacionados con su padre según PCP2 y su localización se define según PCP3, con las siguientes excepciones

1. Cualquier coordenada X de la forma $ka + x$, para un entero k no nulo y $0 < x < a$, se transforma en x .
2. Cualquier coordenada Y de la forma $kb + y$, para un entero k no nulo y $0 < y < b$, se transforma en y .

Así transformamos el rectángulo en un toro, identificando las fronteras opuestas, para evitar el efecto frontera.

Cuando PCP2 indica que el número de descendientes de cada evento principal sigue una distribución de Poisson, se puede simular el proceso condicionando al número total de eventos en A situando aleatoriamente los descendientes entre los eventos principales.

Los procesos de Poisson con clústers se pueden extender a procesos multigeneracionales, donde los descendientes de una etapa pasan a ser los eventos principales de la siguiente generación. Este tipo de construcción tiende a resultar matemáticamente intratable.

3. PROCESOS INHOMOGÉNEOS DE POISSON

Cuando se aplican procesos puntuales en campos como la epidemiología no es realista suponer estacionariedad. El modelo más simple cuando el proceso no es estacionario es el proceso inhomogéneo de Poisson, que se obtiene sustituyendo la intensidad constante, λ , de un proceso de Poisson por una función de intensidad variable $\lambda(x)$. Así un proceso inhomogéneo de Poisson está definido por las siguientes propiedades

Definición:

Un proceso puntual es un proceso inhomogéneo de Poisson si:

IPP1) El número de eventos en una región A , $N(A)$ sigue una distribución de Poisson con media $\int_A \lambda(x) dx$, para alguna función no negativa $\lambda(x)$.

IPP2) Dado $N(A) = n$ los n eventos forman una muestra aleatoria de la distribución en A con función de distribución de probabilidades proporcional a $\lambda(x)$.

IPP3) Para dos regiones disjuntas A y B , las variables aleatorias $N(A)$ y $N(B)$ son independientes.

Los procesos inhomogéneos de Poisson proporcionan un posible marco para introducir covariables en el análisis de un diseño puntual espacial a través de la función de intensidad $\lambda(x) = \lambda(z_1(x), z_2(x), \dots, z_p(x))$. Por ejemplo, supongamos que la localización de los árboles de una determinada especie se ajusta a un proceso de Poisson con intensidad determinada por la altura sobre el nivel del mar, entonces un posible modelo para este proceso será $\lambda(x) = \exp\{\alpha + \beta z(x)\}$, donde $z(x)$ denota la altura de x sobre el nivel del mar. Cox (1972) llama a estos modelos procesos de Poisson modulados.

Los procesos inhomogéneos de Poisson se pueden simular a partir de IPP2, tanto para $N(A)$ fijo como generado aleatoriamente. Lewis & Shedler (1979) propusieron un algoritmo basado en aceptación – rechazo, que consiste en simular un proceso de Poisson con intensidad λ_0 , igual al máximo de $\lambda(x)$ en A , y mantener cada evento en X con probabilidad $\lambda(x)/\lambda_0$.

4. PROCESOS DE COX

Si se quieren modelar fenómenos donde la distribución espacial de los eventos puede ser el resultado de variaciones estocásticas en los factores ambientales, es razonable pensar en la función de intensidad de un proceso de Poisson, $\lambda(x)$, como una realización de un proceso estocástico, $\{\Lambda(x); x \in \mathbb{R}^2\}$

Definición:

Un proceso puntual es un proceso de Cox cuando verifica:

CP1) $\{\Lambda(x); x \in \mathbb{R}^2\}$ es un proceso estocástico con valores no negativos.

CP2) Condicionado a una realización de $\{\Lambda(x)\}$, el proceso puntual es un proceso de Poisson inhomogéneo con función de intensidad $\lambda(x)$.

El proceso resultante hereda de forma natural las propiedades de $\Lambda(x)$, es decir, el proceso puntual es estacionario o isotrópico en función de que $\Lambda(x)$ lo sea.

Las propiedades de primer y segundo orden se obtienen como las de un proceso inhomogéneo de Poisson tomando medias respecto a $\{\Lambda(x)\}$. Así, en el caso estacionario, la intensidad es

$$\lambda = E[\Lambda(x)]$$

La intensidad condicionada de los eventos x e y , dados $\{\Lambda(x)\}$ y $\Lambda(x)\Lambda(y)$, será

$$\lambda_2(x, y) = E[\Lambda(x)\Lambda(y)]$$

Bajo estacionariedad e isotropía

$$\lambda_2(t) = \lambda^2 + \gamma(t) \tag{4.8}$$

donde $\gamma(t) = Cov(\Lambda(x), \Lambda(y))$ y $t = \|x - y\|$.

En las mismas condiciones, la función de segundo momento reducida viene dada por:

$$K(t) = \pi t^2 + 2\pi\lambda^{-2} \int_0^t \gamma(r) r dr$$

Cuando $\gamma(t)$ solo toma valores no negativos, (4.8) es similar a la expresión correspondiente (4.7) para procesos de Poisson con clústers, incluso ambos procesos pueden parecer equivalentes. Para ver esto, sea $h(\cdot)$ una función de distribución de probabilidad bivalente y se genera un proceso de intensidad $\{\Lambda(x)\}$ definiendo

$$\Lambda(x) = \mu \sum_{i=1}^{\infty} h(x - X_i) \tag{4.9}$$

para algún $\mu > 0$, donde los X_i son puntos de un proceso de Poisson. El proceso de Cox definido por (4.9) también es un proceso de Poisson con clústers donde PCP2 especifica una distribución de Poisson con media μ y PCP3 especifica la función de distribución de probabilidad $h(\cdot)$. Intuitivamente, esto se debe a que una distribución de Poisson para el número de descendientes por evento principal se corresponde con la localización aleatoria de descendientes entre los principales. La equivalencia entre ambos procesos se demuestra formalmente en Bartlett (1964).

Matérn (1971) pone de manifiesto que obtener una expresión explícita para la distribución del vecino más próximo de un proceso de Cox general tiene mucha dificultad. Condicionando a la realización $\lambda(x)$ del proceso de intensidad, la probabilidad de no tener eventos a una distancia t del origen es

$$\exp\left(-\int \lambda(x) dx\right) \quad (4.10)$$

siendo la región de integración el disco de centro el origen y radio t . En principio, la distribución de la distancia de un punto al evento más próximo se obtiene calculando la media de (4.10) respecto a la distribución (infinito-dimensional) de la superficie $\Lambda(x)$ en el disco de radio t .

Una construcción relativamente flexible y tratable para procesos de Cox, que no está basada en su relación con los procesos de Poisson con clústers, corresponde a la clase de procesos log-Gaussianos, estudiados por Moller et al. (1998). Estos procesos están definidos tomando $\Lambda(x) = \log Z(x)$, donde $Z(x)$ es un proceso Gaussiano. Las propiedades de segunda orden de estos procesos se deducen de las propiedades de las distribuciones log-Gaussianas. En particular, si $Z(x)$ es estacionario con media μ , varianza σ^2 y función de correlación $\rho(x)$; entonces $\lambda(t) = \exp(\mu + 0.5\sigma^2)$ y $\gamma(t) = \exp(\sigma^2 \rho(t))$.

Kingman (1977) argumentó que los procesos de Cox proporcionan un marco natural para modelar diseños puntuales de poblaciones con reproducción de individuos. Sea G_n el proceso puntual que determina la localización, de los individuos de la n -ésima generación y suponemos las siguientes normas de reproducción

1. El número de descendientes de un individuo x_i es una variable aleatoria de Poisson con media $\mu_i = \mu_i(G_n)$.
2. Las posiciones de los descendientes respecto a sus antecesores están idénticamente distribuidas según una distribución bivalente $h(\cdot)$.

La condición 2 es igual al postulado PCP3 de un proceso de Poisson con clústers, mientras que la primera es similar a PCP2, pero permite que μ_i dependa de la configuración de los eventos principales; por ejemplo, μ_i puede ser función del número de eventos principales en un entorno determinado de x_i . La localización de los descendientes define el proceso G_{n+1} , y así sucesivamente. G_{n+1} es un proceso de Cox con $\{\Lambda(x)\}$ definido por

$$\Lambda(x) = \sum_{i=1}^{\infty} \mu_i(G_n) h(x - x_i)$$

donde los x_i son los eventos.

Todo proceso de Cox se puede simular generando previamente $\{\Lambda(x)\}$ en la región adecuada y utilizando a continuación un algoritmo de aceptación-rechazo para un proceso inhomogéneo de Poisson como vimos en la sección anterior.

5. PROCESOS DE INHIBICIÓN SIMPLE

Los modelos alternativos al de Poisson vistos hasta este momento tienden a producir diseños agrupados. Para describir un patrón regular el camino más natural consiste en imponer una distancia mínima, δ , entre dos eventos. Así, además de mostrar de modo sencillo el comportamiento de las entidades biológicas que se adaptan a estos diseños se pueden poner de manifiesto efectos más sutiles, como son la competición entre plantas o el comportamiento territorial en los animales. Estos modelos se denominan procesos de inhibición simple y hay distintos modos de formularlos. Un concepto importante es la “packing intensity”, τ , definida por

$$\tau = \lambda \pi \delta^2 / 4$$

donde λ es la intensidad. Así, τ es la proporción del plano cubierta por discos de diámetro δ no solapados, o la proporción de cobertura esperada para una región finita A . Nótese que τ alcanza su máximo para discos centrados en los vértices de un triángulo equilátero de lado δ , así $\tau_{\max} = (\pi\sqrt{3})/6 \approx 0.907$.

Matern (1960), fue el primero en describir estos modelos presentando dos tipos de procesos de inhibición simple:

- **Modelo 1:** Se parte de un proceso homogéneo de Poisson con intensidad ρ , que se refina eliminando todos los pares de elementos que distan entre sí menos que δ . Entonces la probabilidad de que un evento arbitrario sobreviva es $\exp(-\pi\rho\delta^2)$ y la intensidad del proceso de inhibición simple es

$$\lambda = \rho \exp(-\pi\rho\delta^2) \quad (4.11)$$

La “packing intensity” correspondiente es menor o igual que $(4e)^{-1} \approx 0.092$, esto representa aproximadamente el 10% de τ_{\max} . Las propiedades de segundo orden se expresan mediante

$$\lambda_2(t) = \begin{cases} 0, & t < \delta \\ \rho^2 \exp(-\rho U_\delta(t)), & t \geq \delta \end{cases} \quad (4.12)$$

donde $U_\delta(t)$ denota el área de la unión de los discos de radio δ con centros situados a una distancia menor que t . Esta expresión se sigue de la probabilidad condicionada de que dos eventos sobrevivan, dado que están situados a una distancia $t \leq \delta$.

- **Modelo 2:** En un proceso homogéneo de Poisson se introducen marcas, $Z(s)$, correspondientes al instante de nacimiento de los eventos. Un evento s se elimina si existe otro evento u con $\|s-u\| < \delta$ tal que $Z(u) < Z(s)$, es decir, si hay algún evento anterior situado a una distancia menor que δ de s . Podemos obtener expresiones para λ y $\lambda_2(r)$ análogas a las del modelo anterior, siempre que ignoremos cuando un

evento mayor ha sido eliminado previamente. Si se tiene en cuenta este último aspecto se genera un proceso secuencial de inhibición simple.

- **Proceso secuencial de inhibición simple:** Se considera una secuencia de n eventos X_i en la región finita A . Entonces el modelo será el resultado del siguiente proceso:
 - SSI1) X_1 está uniformemente distribuido en A .
 - SSI2) Dados $\{X_j = x_j, j = 1, \dots, i-1\}$, X_i está uniformemente distribuido en la intersección de A con el conjunto $\{y; \|y - x_j\| \geq \delta, j = 1, \dots, i-1\}$.

El proceso termina cuando se alcanza la cantidad de eventos predeterminada. Los procesos secuenciales de inhibición simple se parametrizan mediante su “packing intensity”. Si se fija a priori un valor de τ muy grande el proceso secuencial puede terminar prematuramente.

6. PROCESOS PUNTUALES DE MARKOV

Muchos procesos regulares requieren una descripción más flexible que la que proporcionan los procesos de inhibición simple. Por ejemplo, la competición entre plantas puede hacer difícil, aunque no imposible, la supervivencia de dos individuos próximos entre sí.

Para los procesos de inhibición, se ve claramente que la intensidad condicionada de un evento en el punto x , dada la realización del proceso en el resto de la región A , depende solo de la existencia o no de otro evento a una distancia menor que δ del punto. En otras palabras, el proceso de inhibición presenta una especie de dependencia markoviana entre eventos. Por las razones expuestas anteriormente, nos interesa mantener esta dependencia local, pero introducir una mayor flexibilidad en el modelo. Esto motiva la aparición de los procesos puntuales de Markov, introducidos por Ripley y Kelly (1977).

Los procesos puntuales de Markov se definen en una región arbitraria, pero finita, $A \subseteq S$, $|A| < \infty$. Cada proceso está caracterizado por su razón de verosimilitud $f(\cdot)$ respecto al proceso homogéneo de Poisson con intensidad unitaria. Así, si $X = \{x_1, \dots, x_n\}$ denota un conjunto finito de puntos en A , $f(X)$ indica, intuitivamente, en qué medida se parece más la configuración de eventos a un proceso particular que al proceso de Poisson de intensidad unitaria. Siempre se puede factorizar $f(\cdot)$ como un producto de la forma

$$f(X) = \alpha \prod_{i=1}^n g_i(x_i) \prod_{j>i} g_{ij}(x_i, x_j) \dots g_{12\dots n}(x_1, \dots, x_n) \quad (4.13)$$

donde α es una constante normalizadora tal que

$$\sum_{n=0}^{\infty} \int_{X^n} f(x_1, \dots, x_n) = 1$$

Para dar una definición formal de un proceso puntual de Markov, es necesario tener claro el significado del término vecino. Se dice que dos puntos, x e y en A , son vecinos si $\|x - y\| < \rho$, donde $\rho > 0$ es un valor prefijado que se denomina rango del proceso. A continuación se define un grupo como un conjunto de vecinos, y la vecindad de x será el conjunto $\{y \in A; 0 \leq \|x - y\| < \rho\}$.

Definición:

Un proceso puntual es un proceso puntual de Markov de rango ρ si la intensidad condicional en el punto x , dada la configuración de eventos en el resto de la región A , depende solo de la configuración en la vecindad de x .

En este contexto, la intensidad condicional se define como una extensión natural de la intensidad condicional de segundo orden definida en el capítulo 2, excepto cuando en lugar de una configuración de eventos en una región determinada, tenemos un único evento en esa región.

Ripley y Kelly (1977) establecieron un resultado fundamental que dice que para que un proceso puntual sea de Markov de rango ρ es necesario que cada g_i de la expresión de $f(\cdot)$ sea idénticamente unitaria, salvo que sus argumentos constituyan un grupo. Posteriormente se establecieron más condiciones, principalmente para asegurar que $f(\cdot)$ sea integrable.

Para un proceso de Poisson de intensidad unitaria, la verosimilitud de n eventos situados en x_1, \dots, x_n es $\exp(-|A|)$, dado que $N(A)$ sigue una distribución de Poisson con media $|A|$, la distribución de eventos condicionada a $N(A)$ es uniforme y hay $n!$ permutaciones de x_1, \dots, x_n equiprobables. Así, la función de verosimilitud de un proceso puntual de Markov se puede escribir como $f(X) \exp(-|A|)$. En la práctica esto tiene poca utilidad porque se desconoce la constante normalizadora α .

A continuación se presentan algunos ejemplos de procesos puntuales de Markov:

- **Proceso de Strauss:** este proceso es un ejemplo simple de procesos de Markov con familia exponencial. Sea r un número real fijo. Se asume que la densidad del proceso depende solo del número de pares de vecinos, definido por

$$S(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j>i} I(\|x_i - x_j\| \leq r)$$

tomando $g_1(x) = \log \beta$,

$$g_{12}(x_i, x_j) = \begin{cases} \log \gamma, & \text{se } \|x_i - x_j\| \leq r \\ 0, & \text{se } \|x_i - x_j\| > r \end{cases}$$

Y $g_{1,2,\dots,k}(x) = 0$ si $k \geq 3$, se obtiene la función de verosimilitud

$$f(x; \beta, \gamma) = \frac{1}{c(\beta, \gamma)} \beta^{n(x)} \gamma^{S(x)} = \alpha \beta^{n(x)} \gamma^{S(x)}$$

donde β , γ son parámetros no negativos y α la constante normalizadora. Para expresarlo en forma exponencial, sea $t(x) = (n(x), S(x))$ e $\theta = (\theta_1, \theta_2)$, entonces

$$f(x; \theta) = \frac{1}{c(\theta)} g(x; \theta)$$

y

$$g(x, \theta) = e^{\langle t(x), \theta \rangle} = \exp(n(x)\theta_1 + S(x)\theta_2)$$

donde $\theta_1 = \log \beta$ e $\theta_2 = \log \gamma$.

$c(\theta)$ es finita si y solo si $\theta_2 \leq 0$, es decir $\gamma \leq 1$. Entonces el espacio canónico de parámetros será $\Theta = \{\theta \in \mathbb{R}^2; \theta_2 \leq 0\}$, el caso $\gamma = 1$ corresponde a los procesos de Poisson con intensidad β ; mientras que $\gamma = 0$ define un proceso de inhibición donde cada realización es una realización de un proceso de Poisson condicionado a que no puede haber pares de eventos vecinos en A . Este último proceso tiene una formulación distinta a la del proceso secuencial de inhibición simple, pero tiene propiedades estadísticas similares. Valores de γ entre 0 e 1 representan una forma de inhibición no estricta. Strauss propuso este modelo con $\gamma > 1$ para diseños con clústers, pero el proceso explota dando lugar a un número infinito de eventos en A .

Condicionando al número de eventos fijando n , se obtiene una distribución de probabilidad en X válida para todo γ . De todos modos, cuando $\gamma > 1$ los diseños resultantes muestran una agregación extrema, de modo que se adaptan a una configuración de eventos formada por un único grupo de vecinos en la región A .

- **Proceso de tripletas:** la idea de los proceso de Markov sugiere añadir las interacciones de orden siguiente para obtener un proceso que permita atracción positiva de pares de puntos. Sea

$$W(x) = \frac{1}{6} \sum_{i=1}^n \sum_{j \neq i} \sum_{k \neq j \neq i} I(\|x_i - x_j\| \leq r) I(\|x_j - x_k\| \leq r) I(\|x_i - x_k\| \leq r)$$

el número de tripletas de puntos mutuamente independientes. Se define $t(x) = (n(x), S(x), W(x))$, donde $S(x)$ es el mismo que en los procesos de Strauss. La familia exponencial que tiene a $t(x)$ como estadístico canónico y densidad no normalizada

$$g(x, \theta) = e^{\langle t(x), \theta \rangle}$$

se denomina proceso tripleta.

$c(\theta)$ es finita si y solo si $\theta_2 \leq 0$ o $\theta_3 \leq 0$, o cuando $\theta_3 < 0$. Entonces, el espacio canónico de parámetros puede ser $\Theta = \{\theta \in \mathbb{R}^3; \theta_3 < 0\}$. Este es un buen modelo para agrupaciones de eventos.

- **Proceso de saturación:** otra forma sencilla de obtener un modelo para datos agrupados es definiendo para cada $s \in x$

$$m_s(x) = \sum_{u \neq s \in x} I(\|s - u\| \leq r)$$

Entonces $\sum_{s \in x} m_s(x) = 2S(x)$. Se fija una cota superior, $d > 0$, en la influencia de cada punto y se define

$$U(x) = \sum_{s \in x} \min\{d, m_s(x)\}$$

Sea ahora $t(x) = (n(x), U(x))$. La familia exponencial que tiene a $t(x)$ como estadístico canónico y densidad no normalizada

$$g(x, \theta) = e^{\langle t(x), \theta \rangle} = \exp(n(x)\theta_1 + U(x)\theta_2)$$

es un proceso de saturación. En este caso $c(\theta)$ es finita para cualquier valor de θ .

- **Proceso escalonado:** cuando el radio de interacción fijo del proceso de Strauss es sustituido por una función escalonada con interacción constante en cada intervalo, obtenemos una familia exponencial de gran interés. Sean $0 = r_1 < r_2 < \dots < r_p$, $p \geq 1$ números reales fijos. La función análoga a la de Strauss se define como

$$e(t) = \begin{cases} \gamma_i, & \text{se } t = \|x_i - x_j\| \in (r_{i-1}, r_i], i = 1, \dots, p \\ 1, & \text{en otro caso} \end{cases}$$

y proporciona la siguiente función de verosimilitud, para $x = (x_1, \dots, x_n)$

$$f(x; \beta, \gamma_1, \dots, \gamma_p) = \frac{1}{c(\beta, \gamma_1, \dots, \gamma_p)} \beta^{n(x)} \prod_{i=1}^p \gamma_i^{S_i(x)}$$

donde $S_i(x)$ es el número de pares de vecinos con distancia en $(r_{i-1}, r_i]$, $i = 1, \dots, p$. El estadístico canónico es $t(x) = (n(x), S_1(x), \dots, S_p(x))$ y el espacio de parámetros $\theta = (\theta_1 = \log \beta, \theta_2 = \log \gamma_1, \dots, \theta_{p+1} = \log \gamma_p)$.

La constante de normalización es finita para todo $\gamma_i \leq 1$.

6.1. Intensidad condicional de Papangelou y condiciones de estabilidad.

A continuación se definen características fundamentales de los procesos puntuales con densidad que van a jugar un papel muy importante en sus algoritmos de simulación.

Definición:

La intensidad condicional de Papangelou para un proceso puntual X con densidad f está definida por

$$\lambda(\xi; x) = f(x \cup \xi) / f(x), \quad x \in N_f, \quad \xi \in S \setminus x$$

tomando $a/0 = 0$, $a \geq 0$ (Kallenberg 1984), donde $N_f = \{x \subset S; n(x) < \infty\}$

Nótese que $\lambda(\cdot; x)$ no depende de la constante normalizadora de f . Se puede interpretar $\lambda(\xi; x) d\xi$ como la probabilidad condicionada de que X tenga un punto en un entorno infinitesimal de ξ , $d\xi$, dado que el resto de X es x . Se dice que X (o f) es atractiva cuando

$$\lambda(\xi; x) \leq \lambda(\xi; y) \quad \text{si } x \subset y$$

y repulsiva, se

$$\lambda(\xi; x) \geq \lambda(\xi; y) \quad \text{si } x \subset y$$

Intuitivamente, dado $X \setminus \xi = x$, la atracción implica que la probabilidad de que $\xi \in X$ es una función creciente, mientras que el carácter repulsivo significa lo contrario.

Habitualmente se consideran funciones $h: N_f \rightarrow [0, \infty)$ hereditarias, esto es

$$h(x) > 0 \Rightarrow h(y) > 0 \quad y \subset x$$

Si f es hereditaria, hay una correspondencia uno a uno entre ella y $\lambda(\cdot; x)$.

A continuación se introduce una condición de estabilidad que será útil en la simulación de procesos puntuales con densidad.

Definición:

Sea $\phi^* : S \rightarrow [0, \infty)$ una función tal que $c^* = \int_S \phi^*(\xi) d\xi$ es finita, Una función $h : N_f \rightarrow [0, \infty)$ es localmente estable, o ϕ^* -localmente estable, si

$$h(x \cup \xi) \leq \phi^*(\xi) h(x)$$

para todo $x \in N_f$, $\xi \in S \setminus x$.

La estabilidad local implica que h es hereditaria. Si $f \propto h$ es ϕ^* -localmente estable, entonces $\lambda(\xi; x) \leq \phi^*(\xi)$. La mayoría de los modelos de procesos puntuales presentan estabilidad local.

6.2. Procesos con interacciones pares

La clase de los procesos con interacciones pares está definida por

$$f(X) = \alpha \beta^n \prod_i \prod_{i \neq j} h(\|x_i - x_j\|)$$

donde α es la constante normalizadora y β refleja la intensidad del proceso, $h(u)$ es una función no negativa para todo u y se hace el producto en todos los pares de puntos distintos en X . Se ve claramente que el proceso de Strauss es un caso particular de estos procesos. Para definir un proceso puntual válido necesitamos que $h(\cdot)$ sea acotado y que $h(u) = 0$ para todo u inferior a un $\delta > 0$, con esta última condición, al imponer una distancia mínima entre eventos, se limita el número de eventos en una región finita A . Aún así, los modelos con $h(u) > 1$ tienden a producir diseños poco realistas, salvo cuando es natural considerar un proceso con un número fijo de eventos en una determinada región finita.

6.3. Formas de interacción más generales

Un modo satisfactorio de modelar diseños agregados en el marco de los procesos de Markov son los procesos con área de interacción, propuestos por Baddeley y Van Lieshout (1995). En el caso más sencillo, estos procesos están definidos por

$$f(X) = \alpha \beta^{n(x)} \gamma^{-A(x, \delta)}$$

onde $A(x, \delta)$ es el área de la unión de discos de radio δ centrados en puntos de X .

Cuando $\gamma=1$ tenemos un proceso de Poisson con intensidad β ; si $\gamma < 1$ o $\gamma > 1$ el proceso es agregado o regular, respectivamente. A diferencia de los procesos de Strauss, la potencia del parámetro γ es sublineal en n , número de eventos en X , y esto evita la explosión del proceso cuando $\gamma > 1$, generando así un modelo estable de agrupaciones espaciales.

7. OTROS MODELOS

7.1. Procesos refinados

Muchos procesos biológicos presentan mortalidad, en ocasiones como reacción ante un ambiente desfavorable. Por ejemplo, la probabilidad de supervivencia de un cultivo depende de la cantidad de alimento disponible en su entorno. Los procesos puntuales refinados proporcionan una clase de modelos para diseños fruto de la variación espacial de la mortalidad.

Un proceso puntual refinado se define mediante un proceso puntual inicial $N_0(dx)$ y un "campo de refinamiento" $Z(x)$ que es un proceso estocástico, independiente de $N_0(dx)$, con valores en $[0,1]$ para todo x . Una vez que tenemos $N_0(dx)$ y $Z(x)$, cada evento x_i , de $N_0(dx)$, tiene probabilidad $Z(x)$ de permanecer en el proceso. La realización correspondiente del proceso refinado, $N(dx)$, está formada por los eventos que permanecen.

Las características de segundo orden de $N(dx)$ se obtienen a partir de $N_0(dx)$ y $Z(x)$. En particular, en el caso estacionario e isotrópico, sean μ y $\gamma(t)$ la media y la función de covarianzas de $Z(x)$. Entonces, la función de intensidad de segunda orden de $N(dx)$ es

$$\lambda_2(t) = \lambda_{02}(t)(\gamma(t) + \mu^2)$$

donde $\lambda_{02}(t)$ es la intensidad de segundo orden de $N_0(dx)$. Esta expresión de $\lambda_2(t)$ se obtiene porque dado un par de eventos x e y , en $N_0(dx)$, a distancia t entre sí, la probabilidad de que ambos permanezcan en el sistema es $Z(x)Z(y)$. Utilizando la relación entre $\lambda_{02}(t)$ y $K_0(t)$, junto con la expresión anterior, se llega a la siguiente relación entre las K-funciones de $N(dx)$ y $N_0(dx)$

$$K(t) = K_0(t) + \mu^{-2} \int_0^t \gamma(u) K_0'(u) du$$

Si $N_0(dx)$ es un proceso homogéneo de Poisson, entonces el proceso refinado es un proceso de Cox. Cuando $N_0(dx)$ es un proceso de Markov, el proceso refinado nos permite combinar los efectos de la interacción entre eventos con los de las variaciones ambientales.

7.2. Superposiciones

Otro posible modo de construir un modelo consiste en la superposición de dos o varios procesos.

Asumiendo que los procesos que forman el modelo son independientes, se obtienen fácilmente las características de segundo orden y las funciones de vecino más próximo del modelo a partir de cada componente.

Si consideramos el caso bivariante, es decir un modelo resultante de la superposición de dos modelos independientes, la suma de las intensidades de cada componente nos da la intensidad del modelo superpuesto $\lambda = \lambda_1 + \lambda_2$. Teniendo en cuenta que $\lambda K(t)$ es el número esperado de eventos a una distancia menor que t de un evento arbitrario, junto con la independencia entre componentes, se llega a la siguiente expresión:

$$\lambda K(t) = p[\lambda_1 K_1(t) + \lambda_2 \pi t^2] + (1-p)[\lambda_2 K_2(t) + \lambda_1 \pi t^2]$$

donde $p = \lambda_1/\lambda$ es la probabilidad de que un evento pertenezca a la componente 1. Así

$$K(t) = \lambda^{-2} [\lambda_1^2 K_1(t) + \lambda_2^2 K_2(t) + 2\lambda_1 \lambda_2 \pi t^2] \quad (4.14)$$

Análogamente, denotando las distribuciones de mínima distancia punto-evento y evento-evento por $F(\cdot)$ y $G(\cdot)$ respectivamente, se tiene

$$F(x) = 1 - (1 - F_1(x))(1 - F_2(x)) \quad (4.15)$$

y

$$G(x) = 1 - p(1 - G_1(x))(1 - F_2(x)) + (1-p)(1 - G_2(x))(1 - F_1(x))$$

Estas construcciones ponen de manifiesto como las características de segundo orden no describen por completo un proceso. Si se considera la superposición de un proceso homogéneo de Poisson y un proceso de Poisson con clústers, con K-funciones $K_1(t) = \pi t^2$ y

$$K_2(t) = \pi t^2 + \rho^{-1} H(t)$$

donde $H(\cdot)$ es la función de distribución de la distancia entre dos descendientes del mismo evento principal. Entonces la K-función del proceso superpuesto será

$$K(t) = \pi t^2 + \lambda^{-2} \lambda_2^2 \rho^{-1} H(t)$$

Vemos que $K(t)$ tiene la misma forma que $K_2(t)$, basta tomar $\rho^* = \lambda^2 \rho / \lambda_2^2$. Por tanto, al comparar las propiedades de segundo orden, la superposición de un proceso de Poisson y un proceso de Poisson con clústers es indistinguible de un proceso de Poisson con clústers. Para

diferenciar un proceso del otro tendríamos que comparar, por ejemplo, las distribuciones del vecino más próximo.

7.3. Interacciones en un ambiente inhomogéneo

Los procesos puntuales de Markov se utilizan para modelar interacciones entre eventos, mientras que los procesos de Poisson inhomogéneos modelan heterocedasticidad. Para combinar ambos efectos, basta cambiar la constante β de la densidad de un proceso con interacciones pares, por una función $\beta(x)$ que varía con la posición. Entonces la densidad para una configuración de eventos $X = (x_1, \dots, x_n)$, viene dada por

$$f(X) = \alpha \sum_{i=1}^n \beta(x_i) \sum_{j \neq i} h(\|x_i - x_j\|)$$

8. MODELOS MULTIVARIANTES

8.1. Procesos puntuales con marcas

Una forma de construir un modelo multivariante consiste en etiquetar cada evento de un proceso puntual univariante con una variable de categoría que define los distintos tipos de eventos. La variable categórica se denomina marca y el proceso resultante es un ejemplo de proceso puntual con marcas.

En este contexto, el punto de partida más sencillo para construir el modelo es asumir que las marcas son mutuamente independientes e idénticamente distribuidas. Bajo esta hipótesis para las etiquetas, todas las K-funciones son idénticas. Permitiendo interacciones entre marcas se elaboran modelos más generales. Por ejemplo, para modelos con marcas se puede tomar un campo aleatorio con marcas que define, para cada evento x_i , la distribución condicionada de las marcas en x_i , dadas las marcas en los demás eventos x_j .

8.2. Procesos puntuales multivariantes.

Considerando interdependencia entre los procesos puntuales univariantes, se llega a una construcción distinta para modelos multivariantes. Como se ha visto en el capítulo anterior, si los procesos unitarios son estacionarios e independientes dos a dos, la K-función bivalente correspondiente es πt^2 , independientemente de cuáles sean las propiedades marginales de cada componente.

8.3. Como formular procesos multivariantes

A pesar de que cualquier modelo multivariante se puede formular como un proceso puntual con marcas o un proceso puntual multivariante, en la práctica la elección entre una formulación u otra nos conduce a distintos modelos. En particular, las hipótesis de etiquetado aleatorio e independencia son distintas, salvo para procesos homogéneos de Poisson. Hay un rango tan amplio de posibles modelos para procesos multivariantes que es más adecuado usar la información que nos proporciona el contexto científico de cada aplicación particular para modelar el proceso, que tratar de establecer un catálogo de modelos estándar.

Como ejemplo, podemos contrastar tres situaciones reales que contienen un proceso puntual bivalente o, equivalentemente, un proceso puntual con marcas binarias. En este último contexto, sean P el proceso puntual sin marcas y M el proceso con marcas binarias, la distribución conjunta de P y M se puede factorizar de dos formas equivalentes.

$$[P, M] = [P][M|P] = [M][P|M]$$

donde $[]$ denota "la distribución de " .

En primer lugar, situémonos en el área de epidemiología. En este caso, P indica la localización de los miembros de la población de riesgo de una enfermedad y M indica que individuos contraen la dolencia. En esta situación, el proceso sin marcas, P , se puede construir físicamente, por tanto es natural desarrollar un modelo a partir de la factorización $[P, M] = [P][M|P]$. Además, no tiene interés científico especificar $[P]$, por tanto los epidemiólogos no están interesados en modelos elaborados para esta distribución. Aquí la atención se centra en la distribución condicional $[M|P]$.

En segundo lugar pensemos en la exploración en busca de un determinado mineral. Ahora P indica la localización de una serie de perforaciones exploratorias y M indica donde hay mineral con calidad apta para el comercio. A diferencia del ejemplo anterior, ahora el proceso binario M procede de un campo aleatorio $\{M(x), x \in A\}$ definido en la región de estudio, A . En este caso es natural formular un modelo marginal para M a partir de la factorización $[P, M] = [M][P|M]$. Estos modelos se crean en el campo de la estadística espacial denominado geoestadística, aunque que en este contexto no es habitual considerar la distribución conjunta $[P, M]$ sino que se asume que P y M son independientes y, por tanto, $[P, M] = [P][M]$, aunque esta hipótesis se viola en el momento en que se realizan las perforaciones en zonas con probabilidad alta de contener mineral apto para el comercio.

Por último, se considera la distribución conjunta de las localizaciones de los nidos de dos especies de hormigas. En este caso, ninguna de las posibles factorizaciones parece particularmente útil. En lugar de centrar nuestra atención en modelar un proceso puntual para las hormigas sin tener en cuenta su especie, o en especificar la especie de una hormiga situada en un punto arbitrario, es más natural modelar componente, P_1 y P_2 , por separado junto con las posibles interacciones entre hormigas de la misma o distinta especie.

8.4. Procesos de Cox

Las componentes de un proceso de Cox multivariante son procesos de Poisson mutuamente independientes condicionados a las intensidades correspondientes $\lambda_j, j = 1, \dots, k$, que son realizaciones de un proceso estocástico multivariante con valores no negativos, $\Lambda(x) = \{\Lambda_1(x), \dots, \Lambda_k(x)\}$. vamos a estudiar el caso bivalente, $k = 2$.

En primer lugar, nótese que cualquiera que sea la estructura de dependencia entre las componentes del proceso de Cox, ésto se traduce directamente en dependencia entre los procesos $\{\Lambda_1(x)\}$ y $\{\Lambda_2(x)\}$. En este sentido los procesos de Cox proporcionan un marco natural para modelar la reacción conjunta de dos tipos de eventos en un entorno heterogéneo, pero no incluyen interacción estocástica directa entre eventos.

Igual que en el caso univariante, las distribuciones de vecino más próximo y espacio vacío resultan intratables, pero las propiedades de segundo orden se pueden expresar en términos de las propiedades correspondientes de $\Lambda(x)$. Para $\Lambda(x)$ estacionario

$$\lambda_j = E[N_j(x)] \quad (4.16)$$

Y

$$\gamma_{ij}(u) = Cov(\Lambda_i(x), \Lambda_j(y))$$

donde $u = \|x - y\|$. Entonces, λ_1 y λ_2 son las intensidades marginales del proceso de Cox definido por $\Lambda(x)$. Las funciones de intensidad de segundo orden son

$$\lambda_{ij}(u) = \gamma_{ij}(u) + \lambda_i \lambda_j \quad (4.17)$$

y las K-funciones vienen dadas por

$$K_{ij}(t) = \pi t^2 + 2\pi (\lambda_i \lambda_j)^{-1} \int_0^t \gamma_{ij}(u) u du \quad (4.18)$$

Para dar una noción intuitiva de lo que sucede cuando hay correlación positiva extrema entre las componentes del proceso de Cox, Diggle y Milne (1983b) definen un proceso de Cox enlazado como aquel que verifica

$$\Lambda_1(x) = v \Lambda_2(x)$$

Para $v = \lambda_1 / \lambda_2 > 0$. Combinando (4.16), (4.17) y (4.18) se deduce que

$$\lambda_{11}(u) = v \lambda_{12}(u) = v^2 \lambda_{22}(u)$$

con

$$\lambda_{12}(u) = \nu(\gamma_{22}(u) + \lambda_{22}(u))$$

Esto muestra que la covarianza entre las componentes del proceso es simplemente la covarianza de $\Lambda_2(x)$ reescalada. Sustituyendo en (4.18) tenemos

$$K_{11}(t) = K_{22}(t) = K_{12}(t) = \pi t^2 + 2\pi\lambda_2^{-2} \int_0^t \gamma_{22}(u) u du$$

la coincidencia entre las tres K-funciones para un proceso bivalente de Cox enlazado nos recuerda que estos procesos son ejemplos de mecanismos de etiquetado aleatorio que, en otros contextos, se corresponden con componentes del proceso no asociadas entre sí. Es fácil construir procesos distintos a los de Cox con $K_{12}(t) > K_{jj}(t)$; por ejemplo podemos considerar la siguiente definición de un proceso de Poisson bivalente enlazado (Linked pairs bivariate Poisson process):

- LP1) Los eventos de tipo 1 forman un proceso de Poisson homogéneo con intensidad λ .
- LP2) Cada evento de tipo 1 tiene asociado un evento de tipo 2.
- LP3) Las posiciones de cada evento de tipo 2 respecto al evento asociado de tipo 1 están determinadas por realizaciones mutuamente independientes de una distribución bivalente radialmente simétrica.

Cada componente es un proceso de Poisson, entonces $K_{11}(t) = K_{22}(t) = \pi t^2$. Si $H(t)$ es la función de distribución de la distancia entre dos eventos asociados, es fácil ver que $K_{12}(t) = \pi t^2 + \lambda^{-1}H(t)$.

Para ilustrar la correlación negativa extrema, Diggle y Milne (1983b) definen el proceso bivalente de Cox equilibrado como aquel para el que $\Lambda_1(x) + \Lambda_2(x) = \nu$, una constante positiva. Para cualquier par de procesos la superposición de las componentes es un proceso homogéneo de Poisson.

Los proceso de Cox enlazados y equilibrados son extremos, pues las correlaciones puntuales entre $\Lambda_1(x)$ y $\Lambda_2(x)$ son, respectivamente, 1 y -1. Para generar procesos con niveles de correlación intermedios acudimos al marco log-Gaussiano, donde $\Lambda_j(x) = \exp\{Z_j(x)\}$ e $\{Z_1(x), Z_2(x)\}$ es un proceso Gaussiano bivalente.

8.5. Procesos puntuales de Markov

La definición formal de los procesos puntuales de Markov se extiende directamente al caso multivalente mediante un cambio de notación. Por ejemplo, para un proceso bivalente

de interacción par es necesario especificar las tres funciones de interacción $h_{11}(\cdot)$, $h_{22}(\cdot)$ e $h_{12}(\cdot)$ en lugar de una única $h(\cdot)$.

Sea (X, Y) una configuración bivalente de puntos, donde $X = \{x_1, \dots, x_n\}$ e $Y = \{y_1, \dots, y_n\}$ son las configuraciones de los eventos tipo 1 y tipo 2, respectivamente. Entonces en un proceso bivalente de interacción par, la densidad conjunta de (X, Y) es

$$f(X, Y) = \alpha \beta_1^{n_1} \beta_2^{n_2} \prod_{i \neq h} h_{11}(\|x_i - x_j\|) \prod_{k \neq l} h_{22}(\|y_k - y_l\|) \prod_{p, q} h_{12}(\|x_p - y_q\|)$$

Las componentes son independientes si $h_{12}(u) = 0$, para todo u , y están etiquetadas aleatoriamente si $h_{11}(u) = h_{12}(u) = h_{22}(u)$, para todo u .

Funciones de interacción más simples, con parámetros que pueden tomar distintos valores, dan lugar a un amplio rango de comportamientos bivariantes. Por ejemplo, consideramos inhibición simple dada por

$$h_{ij}(u) = \begin{cases} 0 & u \leq \delta_{ij} \\ 1 & u > \delta_{ij} \end{cases}$$

Si todos los δ_{ij} coinciden, la superposición de los eventos de tipos 1 y 2 es un proceso de inhibición simple tipo Strauss, y sus componentes también presentan regularidad, aunque menos estricta que la del modelo bivalente. Cuando $\delta_{12} = 0$, las componentes son procesos de inhibición simple tipo Strauss independientes, pero no hay inhibición entre eventos de distinto tipo y la superposición es menos regular que las componentes. Por último, si δ_{12} es grande con respecto a δ_{11} y δ_{22} , el diseño bivalente tiende a producir grupos separados de eventos de cada tipo. Cada componente presenta agrupación espacial, pero también, si $\delta_{ij} > 0$, regularidad a pequeña escala dentro de cada grupo.

ESTIMACIÓN DE LAS FUNCIONES NOTABLES

1. ESTIMACIÓN DE LAS PROPIEDADES DE SEGUNDO ORDEN.

1.1. Estimadores empíricos para procesos homogéneos.

Vamos a centrar nuestra atención en estimar la K-función. Pues, una vez obtenido $\hat{K}(t)$, si se tiene en cuenta la relación entre esta función y la intensidad de segundo orden, podemos construir fácilmente un estimador de $\lambda_2(t)$. Basta elegir un parámetro de suavizado $h > 0$, hacer la aproximación $\hat{K}'(t) \approx (\hat{K}(t+h) - \hat{K}(t))/h$ y definir el estimador

$$\hat{\lambda}_2(t) = \hat{\lambda}^2 (2\pi t)^{-1} \hat{K}'(t)$$

Este procedimiento genera un estimador de $\lambda_2(t)$ tipo histograma con intervalos de longitud h centrados en t .

Se ha desarrollado un estimador de $K(t)$ a partir de la definición:

$$\lambda K(t) = E[N_0(t)]$$

es decir, el número esperado de eventos a una distancia menor que t de un evento arbitrario. La intensidad λ es el número esperado de eventos por unidad de área, un estimador obvio de este parámetro es el número observado de eventos por unidad de área

$$\hat{\lambda} = n/|A|$$

Como $E(t) = E[N_0(t)]$ es el número esperado de eventos a una distancia menor que t de un evento arbitrario, podemos construir el siguiente estimador de $E(t)$. Sea $u_{ij} = \|x_i - x_j\|$, la distancia entre los eventos x_i y x_j . Se define

$$\hat{E}(t) = n^{-1} \sum_{i=1}^n \sum_{j \neq i} I(u_{ij} \leq t) \quad (5.1)$$

donde $I(\cdot)$ denota la función indicador. Este estimador es simplemente la sustitución de la esperanza por la correspondiente media muestral. La forma de $\hat{E}(t)$ pone de manifiesto la estrecha relación entre $K(t)$ y las distancias entre eventos. En la práctica este estimador tiene sesgo negativo debido al efecto frontera, para un evento situado a una distancia menor que t de la frontera de A , al contar los eventos que distan menos que t de él, se excluyen aquellos situados a una distancia inferior a t que están fuera de la región. Para corregir este efecto

frontera se han desarrollado varios métodos, a continuación se presenta el propuesto por Ripley (1976).

Sea $w(x, u)$ la proporción de circunferencia de radio u centrada en x que cae dentro de A . Sea $w_{ij} = w(x_i, u_{ij})$. Entonces, para cualquier proceso estacionario e isotrópico, w_{ij} es la probabilidad condicional de que un evento sea observado, dado que está a una distancia inferior a u_{ij} del evento x_i , en general $w_{ij} \neq w_{ji}$. Así, un estimador insesgado para $E(t)$ será:

$$\hat{E}(t) = n^{-1} \sum_{i=1}^n \sum_{j \neq i} w_{ij}^{-1} I(u_{ij} \leq t)$$

Por último, si se divide por $\hat{\lambda} = n/|A|$ tenemos el estimador de $K(t) = E(t)/\lambda$:

$$\hat{K}(t) = n^{-2} |A| \sum_{i=1}^n \sum_{j \neq i} w_{ij}^{-1} I(u_{ij} \leq t) \quad (5.2)$$

Por razones técnicas, Diggle (2003), toma como estimador de la intensidad $\hat{\lambda} = (n-1)/|A|$, así el estimador de $K(t)$ será:

$$\hat{K}(t) = (n(n-1))^{-1} |A| \sum_{i=1}^n \sum_{j \neq i} w_{ij}^{-1} I(u_{ij} \leq t) \quad (5.3)$$

Cuando el número de eventos, n , es grande, ambas versiones son prácticamente iguales.

El estimador de Ripley es aproximadamente insesgado para t suficientemente pequeño. Esta restricción en t es necesaria porque puede haber pesos w_{ij}^{-1} no acotados para t grande. Por ejemplo, si A es el cuadrado de área unidad, el límite superior de t es $\sqrt{2}/2 \approx 0.7$, pero no es habitual calcular $K(t)$ para valores tan grandes porque las fluctuaciones en $\hat{K}(t)$ aumentan con t . Por tanto, como no es realista estudiar los efectos de modelos que operan a la misma escala que las dimensiones de la región de estudio, la condición impuesta para t , no es restrictiva a efectos prácticos.

El software Splancs, incorpora un algoritmo para calcular los pesos cuando A es un polígono arbitrario. Se puede escribir la fórmula explícita de $w(x, u)$ para regiones simples, como círculos o rectángulos. Sea A el rectángulo $(0, a) \times (0, b)$, denotamos $x = (x_1, x_2)$ y sean $d_1 = \min(x_1, a - x_1)$, $d_2 = \min(x_2, b - x_2)$, así d_1 y d_2 son las distancias del punto x a los bordes vertical y horizontal más próximos a él. Para calcular $w(x, u)$ es necesario distinguir dos casos:

1. si $u^2 \leq d_1^2 + d_2^2$, entonces:

$$w(x, u) = 1 - \pi^{-1} \left[\cos^{-1}(\min(d_1, u)/u) + \cos^{-1}(\min(d_2, u)/u) \right]$$

2. si $u^2 > d_1^2 + d_2^2$, entonces:

$$w(x, u) = 0.75 - (2\pi)^{-1} [\cos^{-1}(d_1/u) + \cos^{-1}(d_2/u)]$$

Nótese que $w(x, u) = 1$ si $u \leq \min(d_1, d_2)$. Esta fórmula se aplica para valores de u en el rango $0 \leq u \leq 0.5 \min(a, b)$ que, como se ha visto anteriormente es suficiente a efectos prácticos.

Supongamos ahora que A es el disco de centro el origen y radio α . Sea $r = \sqrt{x_1^2 + x_2^2}$ la distancia de x al centro del disco. De nuevo, se distinguen dos casos:

1. si $u \leq a - r$, entonces $w(x, u) = 1$
2. se $u > a - r$, entonces $w(x, u) = 1 - \pi^{-1} \cos^{-1} \left(\frac{(a^2 - r^2 - u^2)}{(2ru)} \right)$

Esta fórmula se aplica a valores de u entre 0 y α .

La distribución en el muestreo de $\hat{K}(t)$ es analíticamente intratable, salvo cuando se tiene un proceso homogéneo de Poisson. Dado un modelo específico y una región A , se puede estimar la distribución en el muestreo mediante simulación. La expresión teórica de la varianza de $\hat{K}(t)$ para un proceso homogéneo de Poisson proporciona un punto de partida útil para un primer análisis del gráfica de $\hat{K}(t)$.

Para un proceso homogéneo de Poisson, en el que suponemos que el número de eventos observados en la región A es constante, Ripley (1988) desarrolla una aproximación asintótica de la varianza muestral de $\hat{K}(t)$. Lotwick y Silverman (1982) dan una fórmula exacta, pero en general su evaluación requiere integración numérica. Chetwynd y Diggle (1998) dan una nueva aproximación basada en refinamientos de fácil computación para cualquier forma de A .

La aproximación asintótica de Ripley, modificada teniendo en cuenta la expresión de $\hat{K}(t)$ dada por Diggle (5.3), es:

$$v_R(t) = 2 \left\{ |A| / (n-1) \right\}^2 \left\{ \pi t^2 / |A| + 0.96 P t^3 / |A|^2 + 0.13 (n / |A|) P t^5 / |A|^2 \right\} \quad (5.4)$$

donde P denota el perímetro de A . Esta expresión es adecuada para valores relativamente pequeños de t .

El resultado de Lotwick y Silverman, con la misma modificación, para un rectángulo A nos da la siguiente varianza para $\hat{K}(t)$

$$v_{LS}(t) = (n(n-1))^{-1} \left\{ 2b(t) - a_1(t) + (n-2)a_2(t) \right\} \quad (5.5)$$

donde, para una región rectangular A con perímetro P :

$$b(t) = \pi t^2 |A|^{-1} (1 - \pi t^2 / |A|) + |A|^{-2} (1.0716 P t^2 + 2.2375 t^4)$$

$$a_1(t) = |A|^{-2} (0.21 P t^3 + 1.3 t^4)$$

$$a_2(t) = |A|^{-3} (0.24 P t^5 + 2.62 t^6)$$

Estas expresiones son válidas para t menor o igual que el lado más corto del rectángulo. La aproximación de Chetwynd y Diggle incluye sumas de funciones de los pesos de corrección del efecto frontera w_{ij} . Fijado t , se define:

$$\phi_{ij} = 0.5(w_{ij} + w_{ji}) I(\|x_i - x_j\| \leq t)$$

Además se definen:

$$W_n = \sum_{i=1}^n \sum_{j \neq i} \phi_{ij}$$

$$X_n = \sum_{i=1}^n \sum_{j \neq i} \phi_{ij}^2$$

$$Z_n = \sum_{i=1}^n \left(\sum_{j \neq i} \phi_{ij} \right)^2$$

Sea $n^{(k)} = n(n-1)\dots(n-k+1)$, definimos $m_2(t) = X_n / n^{(2)}$, $m_3(t) = (Z_n - X_n) / n^{(3)}$ y $m_4(t) = (W_n^2 - 4Z_n + 2X_n) / n^{(4)}$. Entonces, la varianza estimada de $\hat{K}(t)$ es:

$$v_{CD}(t) = \left(2|A|^2 / n^{(2)} \right) \{ (3-2n)m_4(t) + 2(n-2)m_3(t) + m_2(t) \} \quad (5.6)$$

Esta expresión permite determinar la precisión de un estimador de $K(t)$ sin necesidad de asumir ningún modelo específico. Un método sencillo para esto consiste en dividir A en subregiones iguales, estimar $K(t)$ en cada una de ellas y calcular la varianza empírica de cada estimador. Así, si para cada t , k_i denota el estimador de $K(t)$ para la i -ésima subregión, el estimador global será:

$$\tilde{K}(t) = m^{-1} \sum_{i=1}^m k_i \quad (5.7)$$

con varianza aproximada

$$Var(\tilde{K}(t)) = (m(m-1))^{-1} \sum_{i=1}^m (k_i - \tilde{K}(t))^2 \quad (5.8)$$

Hay dos razones que nos conducen a esta aproximación. En primer lugar, se usa la varianza muestral de k_i como aproximación de su varianza real; en segundo lugar, al dividir la varianza de k_i por m se hace una suposición implícita de independencia entre los estimadores de $K(t)$ para subregiones disjuntas, esto es correcto para procesos homogéneos de Poisson, pero no para modelos más generales. Además, se espera que $\tilde{K}(t)$ sea menos eficiente que $\hat{K}(t)$ porque no usa la información aportada por pares de eventos situados en distintas subregiones, esta es una consideración importante salvo que n sea muy grande. Las reflexiones anteriores indican que el estimador $\tilde{K}(t)$ y su varianza deben utilizarse para valores relativamente pequeños de t o cuando las subregiones de A son sustituidas por réplicas del diseño.

1.2. Estimadores empíricos para procesos inhomogéneos.

Cuando la intensidad del proceso no es constante, es decir cuando el proceso no es homogéneo, el procedimiento descrito en la sección anterior no es válido.

Para definir la K-función inhomogénea, se necesitan ciertas condiciones previas. En primer lugar, si $\lambda(x), x \in \mathbb{R}^2$ es la intensidad de primer orden del proceso, se define:

$$M(A, B) = E \left[\sum_{x_i \in X \cap A} \sum_{x_j \in X \cap B} \frac{1}{\lambda(x_i) \lambda(x_j)} \right]$$

y se supone finita para todo par de conjuntos de Borel, entonces M es el momento de segundo orden de la medida aleatoria \mathcal{F} que asocia el peso $1/\lambda(x_i)$ a cada evento, es decir:

$$\mathcal{F} = \sum_{x_i \in X} \frac{1}{\lambda(x_i)}$$

Se dice que un proceso puntual es estacionario reponderado de segundo orden, cuando la medida \mathcal{F} es estacionaria de segundo orden. Esto equivale a decir que M es invariante por traslaciones, $M(A, B) = M(A+x, B+x)$ para todo $x \in \mathbb{R}^2$.

Un proceso estacionario de segundo orden es también estacionario reponderado de segundo orden. Estos procesos tienen intensidad nula o positiva en casi todo punto. Los procesos inhomogéneos de Poisson o cualquier refinamiento independiente de un proceso estacionario son ejemplos de procesos estacionarios reponderados de segundo orden.

Sea X un proceso puntual espacial estacionario reponderado de segundo orden, se define la K-función inhomogénea de X mediante:

$$K_{in\text{hom}}(t) = \frac{1}{|A|} E \left(\sum_{x_i \in X \cap A} \sum_{x_j \in (X \cap A) \setminus \{x_i\}} \frac{I(\|x_i - x_j\| \leq t)}{\lambda(x_i) \lambda(x_j)} \right) \quad (5.9)$$

donde A es un conjunto de Borel acotado en \mathbb{R}^2 , $I()$ es la función indicador, X es el diseño observado y r la máxima distancia entre pares de eventos x_i, x_j . Como se puede ver en la

definición, la K-función inhomogénea es una generalización al caso de intensidad variable de la K-función de un proceso homogéneo de Poisson.

Si tenemos un proceso estacionario reponderado de segundo orden con intensidad de primer orden conocida, $\lambda(x)$, esto último no es realista, podemos estimar $K_{inhom}(t)$ mediante una sencilla modificación en el estimador para un proceso homogéneo, $\hat{K}(t)$. Esta modificación consiste en reescalar la distancia entre eventos dividiéndola por el producto de la intensidad de primer orden en las localizaciones correspondientes, dando lugar al estimador

$$\hat{K}_{inhom,1}(t, \lambda) = |A|^{-1} \sum_{i=1}^n \sum_{j \neq i} w_{ij}^{-1} I(u_{ij} \leq t) / \{\lambda(x_i) \lambda(x_j)\}$$

Donde w_{ij}^{-1} es el corrector del efecto frontera propuesto por Ripley. Este estimador fue propuesto por Baddeley *et al.* (2000) que también discutieron la necesidad y las consecuencias de sustituir la intensidad de primer orden por un estimador, $\hat{\lambda}(x)$, dando lugar a la nueva aproximación de la K-función inhomogénea:

$$\hat{K}_{inhom}(t) = \frac{1}{|A|} \sum_{x_i \in X \cap A} \sum_{x_j \in (X \cap A) \setminus \{x_i\}} \frac{I(\|x_i - x_j\| \leq t)}{\hat{\lambda}(x_i) \hat{\lambda}(x_j) w_{ij}^{-1}} \quad (5.10)$$

En general, $\hat{\lambda}(x)$ es un estimador no paramétrico de la intensidad de primer orden, que se presenta en la sección 2 de este capítulo. Aunque también es posible estimar la intensidad variable mediante un ajuste paramétrico de un modelo sin interacciones, como veremos en próximos capítulos.

El estimador no paramétrico de la intensidad tiene la siguiente expresión:

$$\hat{\lambda}_h(x) = \sum_{x_i \in X \cap A} k_h(x - x_i) / \int_A k_h(x - x_i) dx \quad (5.11)$$

siendo $k_h(u) = h^{-2}k(u/h)$ es una función núcleo con parámetro de suavizado h .

Sin embargo, el estimador de la K-función inhomogénea que se obtiene cuando se utiliza $\hat{\lambda}_h(x)$ en (5.10) presenta un fuerte sesgo negativo, como consecuencia del sesgo positivo de $\hat{\lambda}_h(x)$. Baddeley *et al.* (2000) estudiaron este sesgo y propusieron la siguiente modificación del estimador de la intensidad para corregirlo.

$$\bar{\lambda}_h(x) = \sum_{x_i \in X \cap A \setminus \{x\}} k_h(x - x_i) / \int_A k_h(x - x_i) dx \quad (5.12)$$

donde se elimina el efecto de cada evento en el estimador de la intensidad en ese punto. Para un proceso de Poisson con clústers, con ventana suficientemente pequeña, también se verifica que el sesgo de $\bar{\lambda}_h(x)$ es menor que el de $\hat{\lambda}_h(x)$. Nótese que $\hat{\lambda}_h(x) = \bar{\lambda}_h(x)$, $x \notin X \cap A$, por tanto si $\hat{\lambda}_h(x)$ y $\bar{\lambda}_h(x)$ son aproximados por sus valores evaluados en una red de puntos fija, los dos estimadores de la intensidad serán iguales con probabilidad 1.

1.3. Estimadores empíricos para procesos multivariantes.

Para estimar $K_{12}(s)$ en un diseño bivalente se parte de la misma idea que en el caso univariante pero se miden distancias entre pares de eventos de distinto tipo. Así, si u_{ij} es la distancia entre el i -ésimo evento de tipo 1 y el j -ésimo de tipo 2, w_{ij} como antes y n_1, n_2 el número de eventos de tipo 1 y 2 respectivamente, se pueden construir dos estimadores de $K_{12}(s)$

$$\hat{\lambda}_2 \hat{K}_{12}(s) = n_1^{-1} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij}^{-1} I(u_{ij} \leq s)$$

$$\hat{\lambda}_1 \hat{K}_{21}(s) = n_2^{-1} \sum_{i=1}^{n_2} \sum_{j=1}^{n_1} w_{ji}^{-1} I(u_{ij} \leq s)$$

Calculando la media ponderada de estos estimadores, se tiene:

$$\begin{aligned} \hat{K}_{12}(s) &= (n_1 n_2)^{-1} |A| \left\{ n_1 \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij}^{-1} I(u_{ij} \leq s) + n_2 \sum_{j=1}^{n_2} \sum_{i=1}^{n_1} w_{ji}^{-1} I(u_{ij} \leq s) \right\} / (n_1 + n_2) = \\ &= (n_1 n_2)^{-1} |A| \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij}^* I(u_{ij} \leq s) \end{aligned} \quad (5.13)$$

donde

$$w_{ij}^* = (n_1 w_{ij}^{-1} + n_2 w_{ji}^{-1}) / (n_1 + n_2)$$

Lotwick y Silverman (1982) desarrollaron la fórmula de la varianza para el caso bivalente. Sean las funciones $b(t), a_1(t)$ e $a_2(t)$ definidas para el caso univariante y $c = n_2 / (n_1 + n_2)$. Entonces, cuando las componentes del proceso son procesos homogéneos de Poisson independientes, se tiene:

$$\begin{aligned} Var(\hat{K}_{12}(t)) &= \\ &= (n_1 n_2)^{-1} |A|^2 \left\{ b(t) - 2c(1-c)a_1(t) + [(n_1 - 1)c^2 + (n_2 - 1)(1-c)^2] a_2(t) \right\} \end{aligned} \quad (5.14)$$

1.4. Estimación de la intensidad de segundo orden con integrales ponderadas.

Recordemos que la K-función de un proceso estacionario está definida por

$$K(t) = 2\pi\lambda^{-2} \int_0^t \lambda_2(s) s ds \quad (5.15)$$

donde λ y $\lambda_2(s)$ son, respectivamente, las intensidades de primer y segundo orden. Por tanto, muchos problemas de inferencia no paramétrica para procesos puntuales se pueden resolver estimando la integral ponderada

$$K_\phi(t) = 2\pi\lambda^{-2} \int_0^t \phi(s)\lambda_2(s)sd s \quad (5.16)$$

donde $\phi(s)$ es una función adecuada para cada problema específico. En lugar de centrarnos en un problema concreto, vamos a ver el resultado general obtenido por Berman y Diggle (1989).

Expresamos $K_\phi(t)$ como

$$K_\phi(t) = 2\pi\lambda^{-2}J(t)$$

con

$$J(t) = \int_0^t \phi(s)\lambda_2(s)sd s \quad (5.17)$$

Usando integración por partes para evaluar $J(t)$ y sustituyendo por la definición de $K(t)$, se obtiene

$$J(t) = \lambda^2(2\pi)^{-1} \left(K(t)\phi(t) - \int_0^t K(s)\phi'(s)ds \right) \quad (5.18)$$

Ahora es inmediato construir un estimados de $J(t)$, pues $\phi(\cdot)$ es una función conocida y podemos sustituir λ y $K(t)$ por sus estimadores. En la práctica, la integral del lado derecho de la expresión (5.18) se resuelve por procedimientos numéricos. Este estimador es más estable que un estimador de $J(s)$ construido a partir de (5.17).

2. ESTIMACIÓN NO PARAMÉTRICA DE LA INTENSIDAD VARIABLE.

2.1. Estimador núcleo de la intensidad.

Supongamos que el conjunto de datos a estudiar son una realización parcial de un proceso de Cox y que queremos estimar una realización de $\Lambda(x)$, es decir la intensidad del proceso. Un estimador muy sencillo e intuitivo consiste en contar, para cada localización x , el número de eventos a una distancia menor que h de x y reescalar por πh^2 , el área del disco de radio h . En la práctica es necesario ajustar este estimador para evitar el efecto frontera. Sea $N(x, h)$ el número de eventos del proceso de Cox con distancia inferior a h de x , entonces el estimador no paramétrico de la realización de $\Lambda(x)$ viene dado por:

$$\hat{\lambda}(x) = N(x, h) / (\pi h^2) \quad (5.19)$$

La primera cuestión que nos planteamos en el momento de aplicar este estimador es qué valor de h se debe utilizar. Para elegir un h adecuado se calcula el mínimo error cuadrático medio (MSE) del estimador.

Diggle (1985b) obtiene la expresión del mínimo error cuadrático medio para un proceso de Cox estacionario e isotrópico. Si el proceso estocástico tiene media $E(\Lambda(x)) = \lambda$ y función de covarianzas $\gamma(u)$, entonces el proceso de Cox tendrá intensidad λ e intensidad de segundo orden $\lambda_2(u) = \gamma(u) - \lambda^2$. El error cuadrático medio de $\hat{\lambda}(x)$, será

$$MSE(h) = E \left[\left(\hat{\lambda}(x) - \Lambda(x) \right)^2 \right]$$

donde la media se calcula respecto a la distribución de Cox, es decir respecto a $\Lambda(\cdot)$ y a la realización del proceso condicionados a $\Lambda(\cdot)$. Debido a la estacionariedad, $MSE(h)$ es independiente de x . Tomando $x=0$ y usando argumentos condicionales estándar, tenemos que

$$\begin{aligned} MSE(h) &= E_{\Lambda} \left[E_N \left[\left(N/(\pi h^2) - \Lambda(0) \right) \middle| \Lambda(\cdot) \right] \right] = \\ &= E_{\Lambda} \left[Var_N \left[N/(\pi h^2) \middle| \Lambda(\cdot) \right] + \left\{ E_N \left[N/(\pi h^2) \middle| \Lambda(\cdot) \right] - \Lambda(0) \right\}^2 \right] \end{aligned} \quad (5.20)$$

donde $N = N(0, h)$. Dado que $N(0, h)$ condicionado a $\Lambda(\cdot)$ sigue una distribución de Poisson con media y varianza $\int_{B(0, h)} \Lambda(x) dx$

$$\begin{aligned} MSE(h) &= E_{\Lambda} \left[\int \Lambda(x) dx / (\pi h^2)^2 + \iint \Lambda(x) \Lambda(y) dy dx - 2 \int \Lambda(x) \Lambda(0) dx + \lambda(0)^2 \right] = \\ &= \lambda / (\pi h^2) + \iint \lambda_2(\|x - y\|) dy dx - 2 \int \lambda_2(\|x\|) dx + \lambda_2(0) \end{aligned}$$

Ahora, usando que $\int \lambda_2(\|x\|) dx = \lambda^2 K(h)$, se obtiene

$$MSE(h) = \lambda_2(0) + \lambda(1 - 2\lambda K(h)) / (\pi h^2) + (\pi h^2)^{-2} \iint \lambda_2(\|x - y\|) dy dx \quad (5.21)$$

como el primer término de la expresión anterior no depende de h , el parámetro de suavizado óptimo para $MSE(h)$ será el que minimice la expresión

$$M(h) = \lambda(1 - 2\lambda K(h)) / (\pi h^2) + (\pi h^2)^{-2} \iint \lambda_2(\|x - y\|) dy dx \quad (5.22)$$

El primer término de (5.22) es función de la K de Ripley, que puede ser sustituida por un estimador $\hat{K}(\cdot)$. Por otra parte, pasando a coordenadas polares, la integral doble del segundo

término puede transformarse en una integral análoga a (5.15) y puede ser estimada como se indica en la sección 1.4.

Diggle y Marron (1988) demuestran la equivalencia entre el método de validación cruzada para calcular la ventana optima en la estimación no paramétrica de la densidad y el MSE para la intensidad de un proceso de Cox. Esta equivalencia permite obtener y acotar la expresión asintótica del $MSE(h)$.

Hasta ahora no se ha tenido en cuenta el efecto frontera en el estimador $\hat{\lambda}(x)$. La forma más natural de solucionar el problema es cambiar el denominador πh^2 por el área de la intersección del disco con la región de estudio. Así, si se observa una realización del proceso en la región A y $B(x, h)$ es el disco de centro x y radio h , el estimador corregido será

$$\hat{\lambda}(x) = N(x, h) / |A \cap B(x, h)| \quad (5.23)$$

Se puede interpretar $\hat{\lambda}(x)$ como un estimador tipo núcleo. Dadas las observaciones $\{x_i \in A, i = 1, \dots, n\}$, tenemos

$$\hat{\lambda}(x) = h^{-2} \sum_{i=1}^n k(h^{-1}(x - x_i))$$

donde la función núcleo

$$k(u) = \begin{cases} (\pi u^2)^{-1}, & \|u\| \leq 1 \\ 0, & \|u\| > 1 \end{cases} \quad (5.24)$$

es una densidad bivalente, radialmente simétrica. Si queremos construir un estimador más suave, se puede utilizar otro núcleo, por ejemplo:

$$k(u) = 3\pi^{-1} (1 - \|u\|^2)^2$$

Debemos tener en cuenta que cuando se cambia el núcleo también hay que cambiar el valor de h . Para tener estimadores comparables, se toma un parámetro de suavizado que haga coincidir la varianza de ambas densidades.

Una vez aplicada la corrección del efecto frontera al estimador núcleo, se obtiene la siguiente expresión:

$$\hat{\lambda}(x) = \sum_{i=1}^n k_h(u) / \int_A k_h(\|x\|) dx$$

siendo $k_h(u) = h^{-2} k(u/h)$.

Cowling, Hall y Phillips (1996) aplican estimación tipo núcleo y desarrollan teoría asintótica para el MSE del estimador de la intensidad y su versión bootstrap para procesos inhomogéneos de Poisson, en lugar de procesos de Cox.

2.2. Estimación bootstrap de la intensidad de un proceso puntual inhomogéneo en \mathbb{R}

En esta sección se introduce la teoría asintótica para el MSE del estimador de la intensidad y su versión bootstrap para procesos inhomogéneos de Poisson en \mathbb{R} desarrollada por Cowling, Hall y Phillips (1996).

Dada una realización (x_1, \dots, x_n) de un proceso de Poisson no necesariamente estacionario con intensidad $\lambda(x)$ en el intervalo $[0,1]$, el estimador núcleo de la intensidad es:

$$\hat{\lambda}_h(x) = \sum_{i=1}^n k_h(x - x_i) = h^{-1} \sum_{i=1}^n k((x - x_i)/h)$$

k ahora es una función núcleo unidimensional.

Se establece la siguiente condición de consistencia para la intensidad: si $\lambda = l\mu$ donde μ es una función fija y l un escalar, entonces:

$$\hat{\lambda}/l \rightarrow \mu \Leftrightarrow \begin{cases} \lim_{l \rightarrow \infty} h = 0 \\ \lim_{l \rightarrow \infty} hl = \infty \end{cases}$$

Bajo la condición anterior, para un núcleo de orden 2, si μ'' existe y es continua, tenemos las siguientes aproximaciones para sesgo y varianza:

$$B(\hat{\lambda}) \approx \frac{h^2}{2} \left(\int u^2 k(u) du \right) \lambda''(x) \Rightarrow B(\hat{\lambda}) = O(h^2 l)$$

$$V(\hat{\lambda}) \approx h^{-1} \lambda(x) \int k^2(u) du \Rightarrow V(\hat{\lambda}) = O(h^{-1} l)$$

Por tanto, el MSE, será:

$$MSE = E(\hat{\lambda} - \lambda)^2 \sim \frac{h^4}{4} d_k^2 \lambda''^2 + h^{-1} \lambda \int k^2$$

donde $d_k = \int t^2 k(t) dt$. La ventana óptima correspondiente:

$$h_{MSE} = \left(\lambda \int k^2 / \lambda''^2 d_k \right)^{1/5} = l^{-1/5} \left(\mu \int k^2 / \mu''^2 d_k^2 \right)^{1/5}$$

que es análoga a la ventana óptima en densidades.

Integrando la expresión del MSE, se obtienen el MISE y su ventana óptima:

$$MISE = \int E(\hat{\lambda} - \lambda)^2 \sim \frac{h^4}{4} d_k^2 \int \lambda^{n^2} + h^{-1} (\int \lambda) (\int k^2)$$

$$h_{MISE} = \left((\int \lambda) (\int k^2) / (\int \lambda^{n^2}) d_k^2 \right)^{1/5} = l^{-1/5} \left((\int \mu) (\int k^2) / (\int \mu^{n^2}) d_k^2 \right)^{1/5}$$

Cowling, Hall y Phillips (1996) proponen utilizar bootstrap suavizado para heredar las propiedades de regularidad del estimador núcleo y poder aproximar el MSE y su ventana óptima. El método de remuestreo propuesto es el siguiente:

1. Generar $X_i^* \sim Poisson(\hat{\lambda})$.
2. El proceso de Poisson bootstrap será $\{x_1^*, \dots, x_{N^*}^*\}$, donde $N^* \sim Poisson(\Lambda(1))$, siendo $\Lambda(x) = \int_0^x \hat{\lambda}(y) dy$.

También se puede expresar como sigue:

1. $N^* \sim Poisson(\Lambda(1))$
2. $X_i^* = Y_i^* + hZ_i^*$, $i = 1, \dots, n$, donde:
 - 2.1. Y_i^* se eligen por muestreo con reemplazamiento en (x_1, \dots, x_n)
 - 2.2. Z_i^* i.i.d. k .

El estimador bootstrap de la intensidad, será:

$$\hat{\lambda}_h^*(x) = \sum_{i=1}^n k_h(x - x_i^*) = h^{-1} \sum_{i=1}^n k((x - x_i^*)/h)$$

y su esperanza es: $E_*[\hat{\lambda}_h^* | X] = \int \hat{\lambda}(x - hy) \delta(y) dy$.

A menos que los datos estén próximos entre sí, no se puede construir un estimador de λ utilizando solo información local. Por tanto, se adopta un modelo asintótico que, como se ha visto, asegura la consistencia del estimador $\lambda(x) = l\mu(x)$. Cuando $l \rightarrow \infty$ estamos aumentando la intensidad del proceso, es decir, el número esperado de eventos por unidad de área. Recordemos que en el caso de densidad se aumentaba el número de observaciones.

Este modelo asintótico permite construir una aproximación Gaussiana de $\hat{\lambda}$ suficientemente precisa para desarrollar teoría límite de primer orden. Así se demuestra que $\hat{\lambda} - E(\hat{\lambda}) \rightarrow N\left(0, \left(h^{-1} \lambda \int \delta^2\right)^{1/2}\right)$ y que su versión bootstrap estima de forma consistente esta distribución. También nos permite desarrollar las expresiones asintóticas de sesgo y varianza correspondientes a ambos estimadores. Así, se demuestra que para un núcleo de orden 2 y si

μ'' existe y es continua, el MSE bootstrap aproxima el MSE del estimador con probabilidad 1, dando lugar a la expresión

$$MSE^* = E_* \left(\hat{\lambda}^* - \hat{\lambda} | X \right)^2 = \frac{h^4}{4} d_k^2 l^2 \mu''^2(x) + h^{-1} l \mu(x) \int k^2 + o(h^4 l^2)$$

2.3. ESTIMACIÓN DE LA INTENSIDAD PARA LOS DATOS DE FUEGOS.

Una vez que los contrastes de CSR nos han llevado a rechazar la hipótesis de que nuestros datos se ajusten a un proceso homogéneo de Poisson, el siguiente paso del análisis será estimar la intensidad variable del conjunto de datos que, como se ha visto anteriormente permitirá estimar la K-función inhomogénea.

En este caso se ha aplicado el método no paramétrico implementado en Spatstat, que utiliza el núcleo Gaussiano:

$$k_\sigma(x) = \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{x_1^2 + x_2^2}{2\pi\sigma^2} \right\}$$

Donde la desviación típica del núcleo, σ , actúa como parámetro de suavizado.

Para seleccionar la ventana óptima, se ha tomado como referencia el método bootstrap propuesto por Cowling, Halli y Philips (1996) que permite aproximar el MSE del estimador aplicando bootstrap suavizado. En esta aplicación se ha obtenido la ventana óptima a partir de la aproximación bootstrap del MISE mediante el siguiente procedimiento de remuestreo:

1. Calcular el estimador de la intensidad con núcleo Gaussiano:

$$\hat{\lambda}_{\sigma_1}(x) = \sum_{i=1}^n k_\sigma(u) / \int_A k_\sigma(\|x\|) dx$$

tomando como ventana piloto la utilizada por defecto en Spatstat, esto es:

$$\sigma_1 = \frac{1}{8} \min \{ rango(X_1), rango(X_2) \}$$

2. Para cada posible valor de la ventana σ .
 - a. Fijado B, generar B realizaciones X_b^* del proceso inhomogéneo de Poisson con intensidad $\hat{\lambda}_{\sigma_1}(x)$. El método utilizado para realizar las simulaciones consiste en simular un proceso de Poisson con intensidad $\lambda_0 = \max_{x \in A} \hat{\lambda}_{\sigma_1}(x)$, y mantener cada evento en X con probabilidad $\lambda(x)/\lambda_0$.

- b. Calcular el estimador no paramétrico de la intensidad para cada réplica, $\hat{\lambda}_{\sigma,b}^*(x)$
- c. Calcular el MISE bootstrap:

$$MISE^*(\sigma) = \frac{1}{B} \sum_{b=1}^B \int_A \left(\hat{\lambda}_{\sigma,b}^*(x) - \hat{\lambda}_{\sigma_1}(x) \right)^2 dx$$

- 3. La ventana óptima será la que minimice el MISE*.

$$\sigma_{opt} = \min_{\sigma > 0} MISE^*(\sigma)$$

Debemos recordar que en esta aplicación se está extendiendo teoría desarrollada en \mathbb{R} a procesos puntuales espaciales en \mathbb{R}^2 . Por otra parte, como se puede ver en la explicación del método bootstrap, el algoritmo utilizado para generar las réplicas se basa en un procedimiento de aceptación-rechazo. Aunque en este caso se han obtenido resultados razonables con el procedimiento aplicado, en el futuro sería conveniente desarrollar un método bootstrap suavizado para procesos puntuales espaciales y justificar formalmente la convergencia del MISE* al MISE.

Se ha aplicado el algoritmo anterior a cada uno de los conjuntos de datos que estamos estudiando. En este caso se han realizado $B = 100$ simulaciones para cada una de las 100 ventanas equidistantes propuestas entre 1 y 12km, el valor de la ventana piloto es aproximadamente 5.5 km. Para cada diseño se ha obtenido el gráfico de σ frente al MISE* y el estimador no paramétrico de la intensidad correspondiente a la ventana óptima obtenida mediante el algoritmo propuesto.

Observando el estimador de la intensidad obtenido para todos los fuegos registrados en el periodo 1991-2008, se tiene que en la mayor parte del distrito la intensidad del proceso es superior a 1.5×10^{-6} fuegos/m², es decir que se han registrado al menos 1.5 fuegos por km². Debemos destacar que la zona sureste del distrito, formada por los municipios de Navia de Suarna y Pedrafita do Cebreiro, es la más conflictiva, llegando a alcanzarse los 6 fuegos/km². Las causas de estos fuegos se encuentran en la conflictividad existente a causa de las disputas de propiedad y gestión de la tierra. En esta zona se han catalogado gran cantidad de superficie de monte como Monte Vecinal en Mano Común (mvmc) lo cual es incorrecto, pues no se trata de montes vecinales sino de montes de baras o voces que es un tipo de monte de proindiviso no reconocido hasta fechas recientes por la legislación. También hay otro pico en el centro del distrito, que corresponde a los montes que limitan entre los municipios de Baleira, Baralla, Becerreá y Navia. En esta zona la causalidad es en gran medida similar al caso anterior, con la incorporación de un factor como es la ganadería extensiva que aprovecha el monte durante algunas épocas del año, sobre todo en verano, provocando los incendios de primavera.

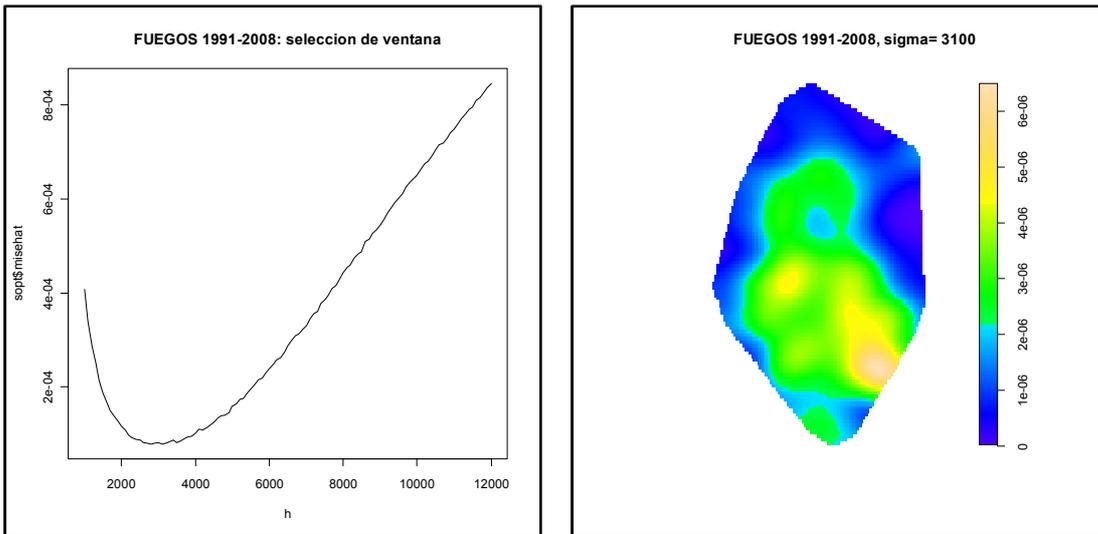


Gráfico 1: Selección de ventana y estimador no paramétrico de la intensidad para los fuegos registrados en el periodo 1991-2008.

El análisis de los fuegos registrados según tipo, gráficos 2-5 confirma la conflictividad del sureste del distrito al ser la zona con mayor probabilidad tanto de incendios como de quemas, en cuanto a los conatos, la región donde se ha registrado una mayor actividad es más amplia, pues abarca tanto el centro como el sureste del distrito. Las causas de este comportamiento de los diferentes tipos de eventos hay que buscarlas en las diferentes condiciones orográficas, poblacionales y de los usos del suelo. Todo conato es potencialmente un incendio a menos que se extinga. Por tanto, la facilidad y rapidez en la extinción juega un papel clave en este proceso. La zona central del distrito, presenta conflictividad, pero el hecho de ser una zona más llana, con más núcleos habitados y con mayor fragmentación de los usos del suelo contribuye a que los conatos no pasen a incendios.

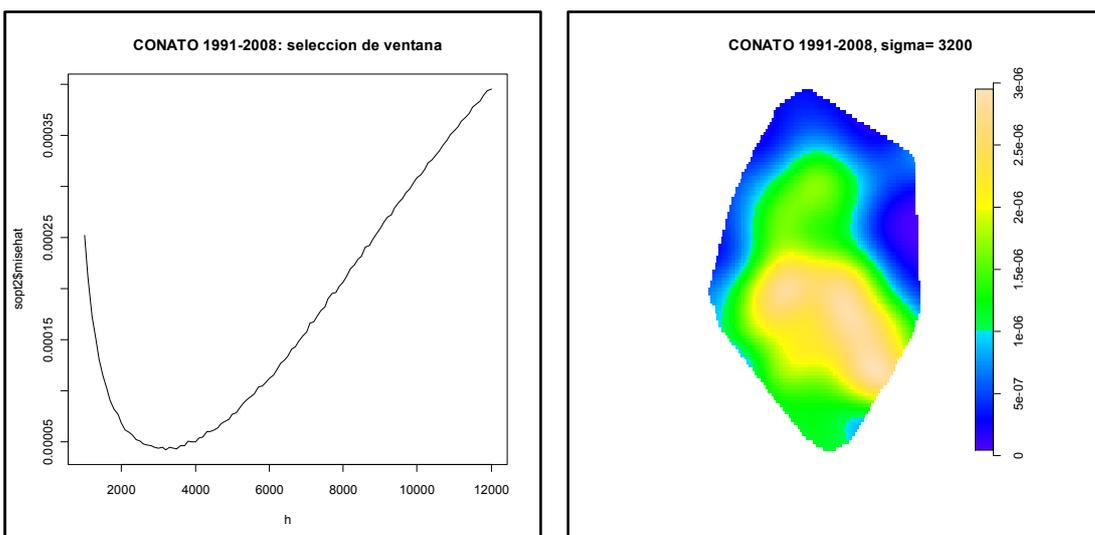


Gráfico 2: Selección de ventana y estimador no paramétrico de la intensidad para los conatos registrados en el periodo 1991-2008.

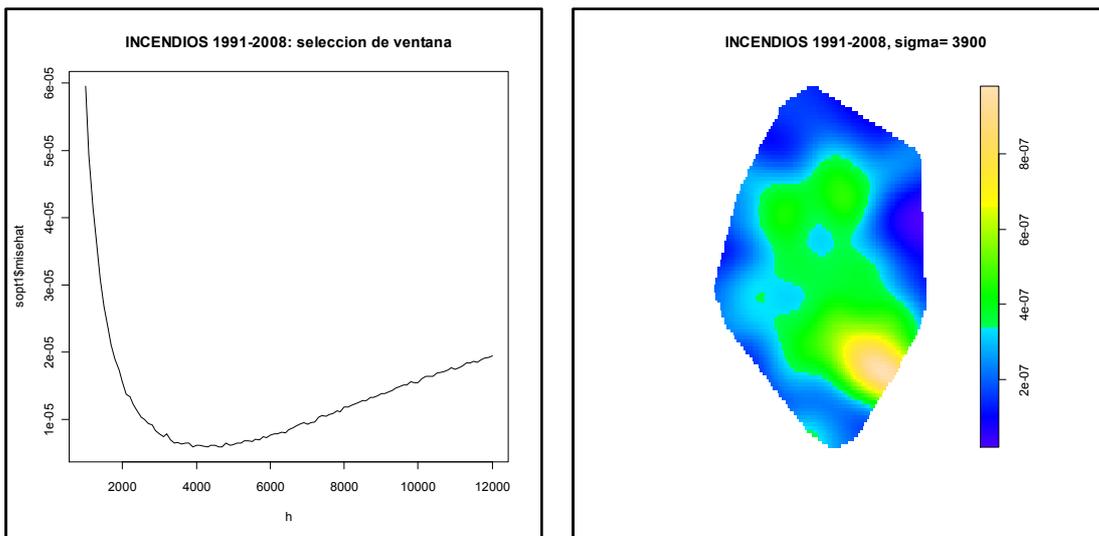


Gráfico 3: Selección de ventana y estimador no paramétrico de la intensidad para los incendios registrados en el periodo 1991-2008.

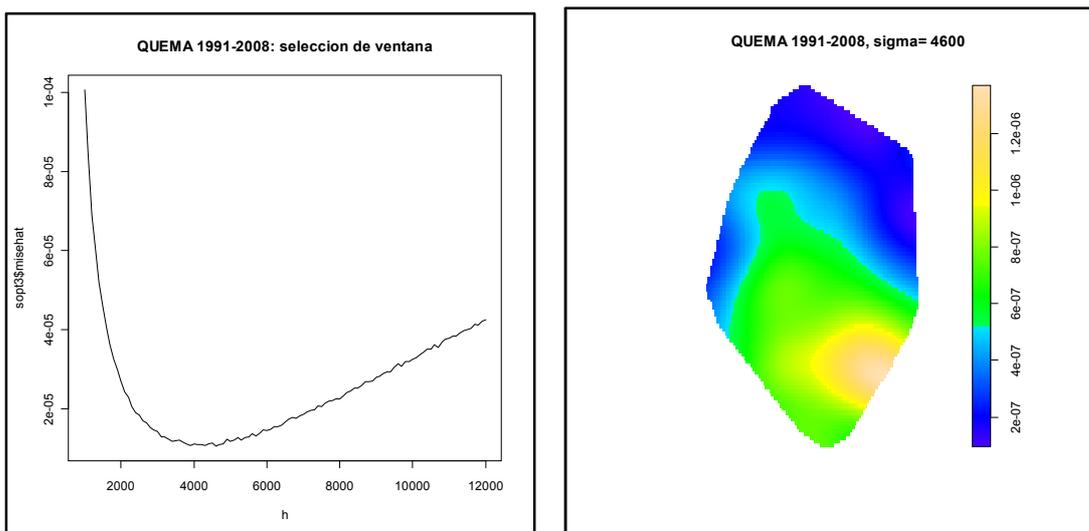


Gráfico 4: Selección de ventana y estimador no paramétrico de la intensidad para las quemas registrados en el periodo 1991-2008.

Otro aspecto interesante en este estudio es el análisis de la distribución espacial de los fuegos clasificados por causa.

A continuación se muestran las intensidades estimadas, gráficos 5-7, correspondientes a fuegos intencionados (4153), con causa natural (279) y con causa desconocida (197). Destacar que la gran mayoría de los fuegos han sido intencionados, asimismo la mayoría de los fuegos catalogados como debidos a causa desconocida son intencionados. Estos factores hay que tenerlos muy en cuenta para explicar posteriormente el comportamiento en la zona del fuego. Los fuegos naturales, presentan un carácter aleatorio y afectan principalmente a aquellas zonas en las que no se producen incendios intencionados, esto es debido a que el

monte se encuentra en excelentes condiciones para arder a causa de la gran cantidad de materia vegetal que acumula porque antes no ha ardido. Por el contrario en los montes en que existen causas antrópicas los pirómanos seleccionan las zonas que desean quemar reiteradamente, lo cual supone un determinismo en la distribución de los incendios. Observando las intensidades estimadas, se tiene que la intensidad de los fuegos intencionados es muy similar a la del total de fuegos, observándose los mismos puntos conflictivos en el sureste y en el centro del distrito. Los fuegos debidos a causas naturales muestran menos variabilidad a lo largo de la región si bien su intensidad no llega a ser homogénea, recordemos que el test Chi-cuadrado aceptaba la hipótesis nula de intensidad constante para este conjunto de datos. En este caso la región con mayor intensidad es el norte de distrito. Por último, los fuegos debidos a causas desconocidas, al igual que los que han sido provocado por causas naturales, no presentan mucha variabilidad, observándose en este caso dos focos más activos situados en el sureste y en el norte de la región.

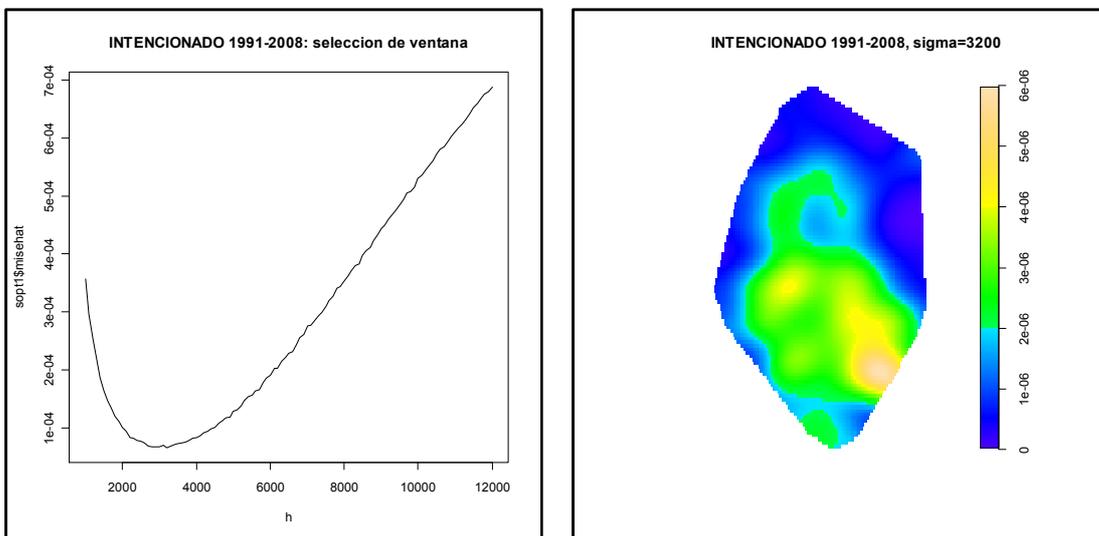


Gráfico 5: Selección de ventana y estimador no paramétrico de la intensidad para fuegos intencionados registrados en el periodo 1991-2008

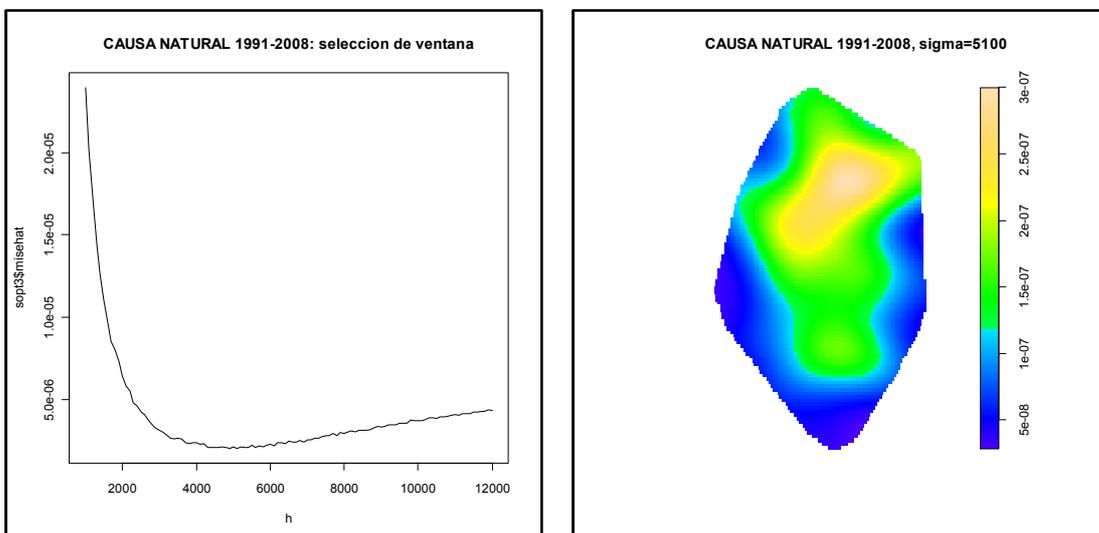


Gráfico 6: Selección de ventana y estimador no paramétrico de la intensidad para los fuegos debidos a causas naturales registrados en el periodo 1991-2008

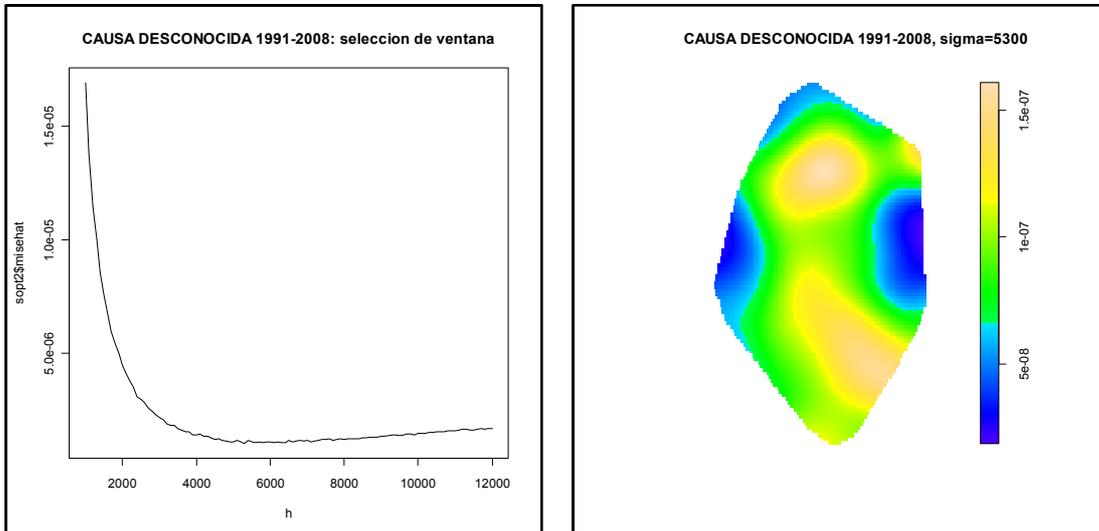


Gráfico 7: Selección de ventana y estimador no paramétrico de la intensidad para fuegos con causa desconocida registrados en el periodo 1991-2008

Por último, se han analizado los fuegos registrados en un solo año, 1992. Así, además de reducir el coste computacional y el número de eventos repetidos, se facilita la interpretación de los patrones desde el punto de vista de práctico. Se ha elegido el año 1992 por ser el más conflictivo con 879 fuegos registrados.

Los gráficos 8-11, en los que se presentan el MISE bootstrap para las ventanas propuestas y la intensidad obtenida con la ventana óptima, muestran que la distribución de los fuegos en 1992 es similar a la observada en todo el intervalo 1991-2008. Así, el sureste de la región aparece de nuevo como la zona con más actividad, al ser la que registra mayor número de fuegos y donde la intensidad de los incendios es también notablemente superior a la observada en el resto de la región. . Estos resultados explican una vez más la importancia de la causa humana o provocada para la puesta en marcha de incendios y cómo esta situación es endémica en ciertos municipios y parroquias.

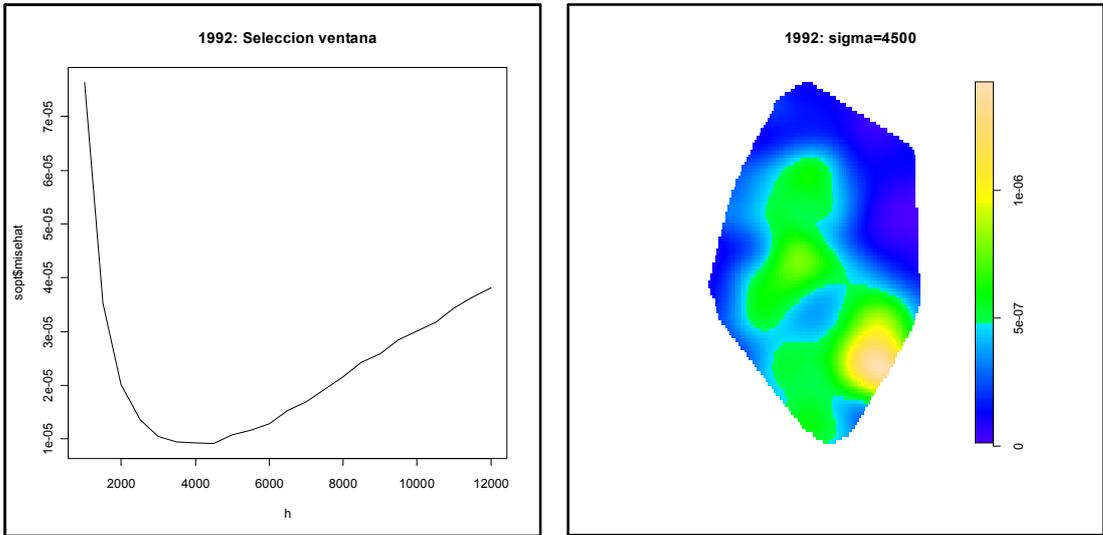


Gráfico 8: Selección de ventana y estimador no paramétrico de la intensidad para los fuegos registrados en 1992

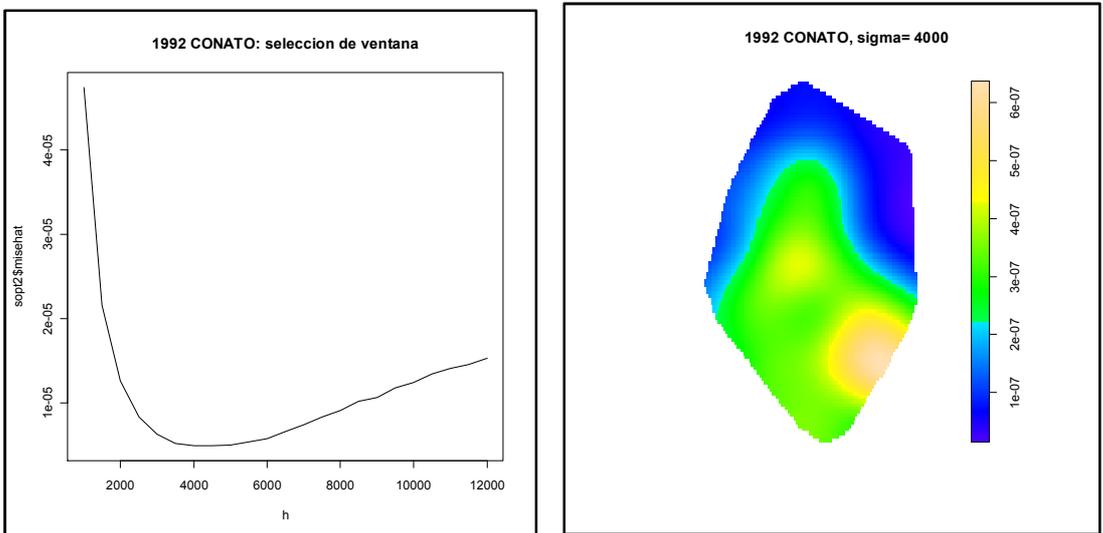


Gráfico 9: Selección de ventana y estimador no paramétrico de la intensidad para los conatos registrados en 1992

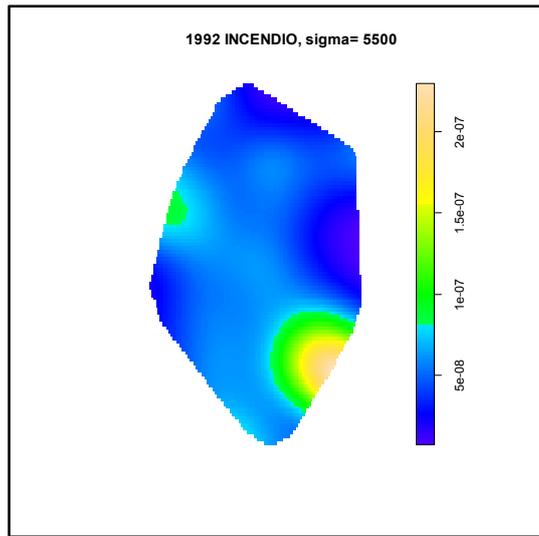
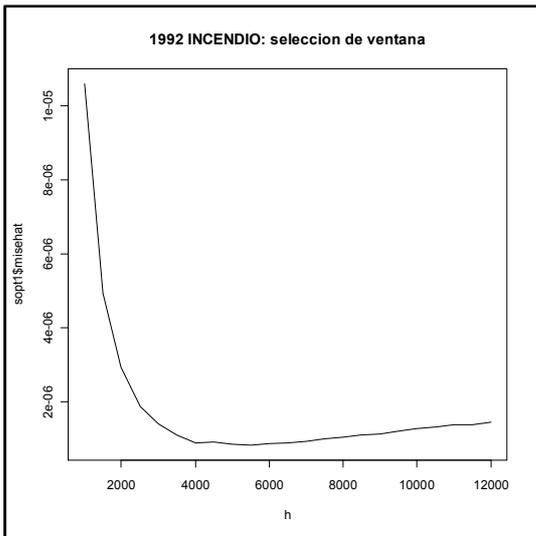


Gráfico 10: Selección de ventana y estimador no paramétrico de la intensidad para los incendios registrados en 1992

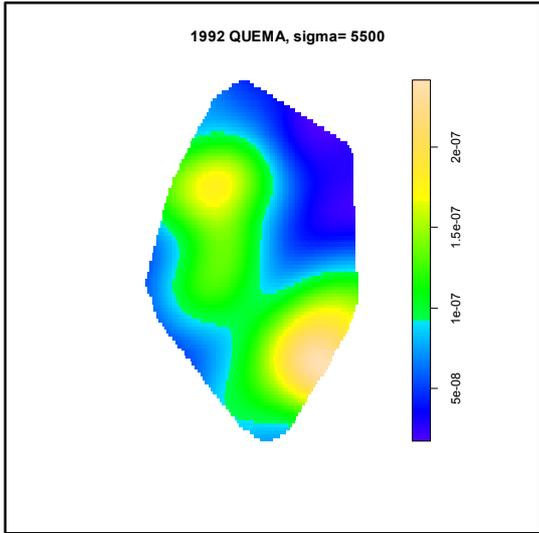
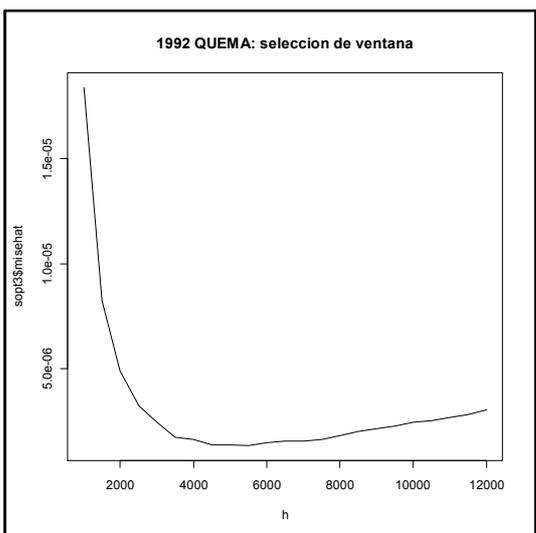


Gráfico 11: Selección de ventana y estimador no paramétrico de la intensidad para quemas registradas en 1992

PROCESOS INHOMOGÉNEOS

1. K- FUNCIÓN PARA PROCESOS INHOMOGÉNEOS.

En el capítulo 3 se ha contrastado la aleatoriedad espacial completa, tanto para el total de fuegos registrados como para los fuegos registrados según tipo y causa en el periodo 1991-2008. Tras aplicar un contraste Chi-cuadrado y tests basados en la distribución de espacio vacío, distribución de vecino más próximo y K-función, se ha concluido que no podemos suponer CSR en ninguno de los diseños estudiados. En concreto el test chi-cuadrado indica que la intensidad de estos procesos no es constante. Por tanto, si queremos analizar la existencia de interacciones entre eventos, debemos utilizar técnicas alternativas a los contrastes de CSR para procesos homogéneos.

Para extender el análisis de segundo orden del proceso al caso no homogéneo, se introduce la K-función inhomogénea. Así se elimina la hipótesis de intensidad constante, aunque pueden seguir manteniéndose estacionariedad e isotropía, en particular el proceso debe ser estacionario reponderado de segundo orden. La interpretación de la K-función inhomogénea es la misma que se tenía en el caso homogéneo, salvo que ahora la intensidad no es constante, sino que depende de la localización de los eventos, $\lambda(x_i)$ es el valor de la intensidad variable en x_i . Como hemos visto en el capítulo anterior, la K-función inhomogénea se define como:

$$K_{inhom}(t) = \frac{1}{|A|} E \left(\sum_{x_i \in X \cap A} \sum_{x_j \in (X \cap A) \setminus \{x_i\}} \frac{I(\|x_i - x_j\| \leq t)}{\lambda(x_i)\lambda(x_j)} \right) \quad (6.1)$$

Donde A es un conjunto de Borel acotado en \mathbb{R}^2 , $I()$ es la función indicador, X es el proceso puntual y r la máxima distancia entre pares de eventos x_i, x_j .

Baddeley et al (2000) propusieron el siguiente estimador puntual insesgado de la K-función inhomogénea:

$$\hat{K}_{inhom}(t) = \frac{1}{|A|} \sum_{x_i \in X \cap A} \sum_{x_j \in (X \cap A) \setminus \{x_i\}} \frac{I(\|x_i - x_j\| \leq t)}{\hat{\lambda}(x_i)\hat{\lambda}(x_j)w_{ij}} \quad (6.2)$$

donde w_{ij} es el corrector del efecto frontera.

Al igual que en el caso homogéneo también se puede transformar la K-función mediante:

$$L_{inhom}(t) = \sqrt{K_{inhom}(t)/\pi}$$

Esta función verifica que, bajo CSR $L_{inhom}(t) = t$, para la K-función teníamos $K_{inhom}(t) = \pi t^2$.

Una vez estimada la K inhomogénea para el proceso observado, $\hat{K}_{inhom,1}(t)$ se puede aplicar un contraste de CSR basado en esta función. Se calcula la K-función $\hat{K}_i(t)$, $i = 2, \dots, s$, para $s-1$ simulaciones independientes de un proceso con intensidad estimada $\hat{\lambda}(x)$ y se definen las envolturas superior e inferior:

$$U(t) = \max_{i=2, \dots, s} \{ \hat{K}_{inhom,i}(t) \}$$

$$L(t) = \min_{i=2, \dots, s} \{ \hat{K}_{inhom,i}(t) \}$$

Representar gráficamente la K-función de los datos observados $\hat{K}_{inhom,1}(t)$ junto con las envolturas superior e inferior, proporciona un test gráfico de CSR. Así, valores de la K-función empírica mayores que la envoltura superior o menores que la inferior indican, respectivamente, diseño agregado o regular. De modo análogo se puede plantear el test basado en la L-función inhomogénea.

Para calcular K_{inhom} se necesita estimar previamente la intensidad en cada evento. Un estimador adecuado de la intensidad variable del proceso se puede obtener mediante el ajustar de un modelo paramétrico o bien aplicando estimación no paramétrica.

El primer método consiste en ajustar un modelo que explique tanto la tendencia espacial como las interacciones entre eventos. Cuando no hay interacción entre eventos se puede ver que los modelos paramétricos donde el logaritmo de la intensidad es un polinomio pueden explicar convenientemente la intensidad de los procesos correspondientes a todos los fuegos y también los relativos a cada tipo o causa. La expresión formal de la intensidad asociada a estos modelos es:

$$\lambda(x) = \exp\{\theta^T p_{x,m}\}$$

Donde $p_{x,m}$ es un polinomio de orden m en \mathbb{R}^2 y θ es el vector de coeficientes asociado al polinomio. Estos coeficientes se obtendrán mediante el método de ajuste de modelos por máxima pseudoverosimilitud implementado en Spatstat.

El segundo método de estimación de la intensidad variable es la construcción de un estimador tipo núcleo de λ , su expresión es:

$$\hat{\lambda}_h(x) = \frac{1}{p_h(x)} \sum_{i=1}^n k_h(x - X_i) = \frac{1}{p_h(x)h^2} \sum_{i=1}^n k((x - X_i)/h)$$

Donde k es la función núcleo, h el parámetro de suavizado y $p_h(x) = \int_A h^2 k((x-u)/h) du$ es el corrector de efecto frontera. Spatstat utiliza un núcleo Gaussiano donde σ actúa como parámetro ventana. Así valores de σ grandes nos llevan a un sobresuavizado, aproximando la

intensidad a una constante, mientras que valores demasiado pequeños introducen demasiada variabilidad reflejando tendencia local en lugar de global.

2. APLICACIÓN A FUEGOS.

Una vez que el test Chi-cuadrado, aplicado a los distintos conjuntos de datos objeto de estudio, nos lleva a rechazar la hipótesis de intensidad constante, es recomendable utilizar la K- función inhomogénea para contrastar la hipótesis de independencia entre eventos.

Para aplicar el contraste de CSR, se ha calculado la K-función inhomogénea utilizando el estimador no paramétrico de la intensidad propuesto en el capítulo anterior. Junto con los gráficos obtenidos para la K de Ripley inhomogénea vuelven a presentarse los correspondientes al K-test bajo la hipótesis de homogeneidad para ver si nos conducen a conclusiones similares.

En primer lugar, se han analizado todos los fuegos registrados entre 1991 y 2008. El test basado en la K-función inhomogénea muestra indicios de agregación para eventos situados a menos de 4 km. Si comparamos este resultado con el obtenido al suponer intensidad constante, el K-test homogéneo daba una mayor evidencia de diseño agregado al ser mayor la distancia entre la K-función observada y envoltura superior.

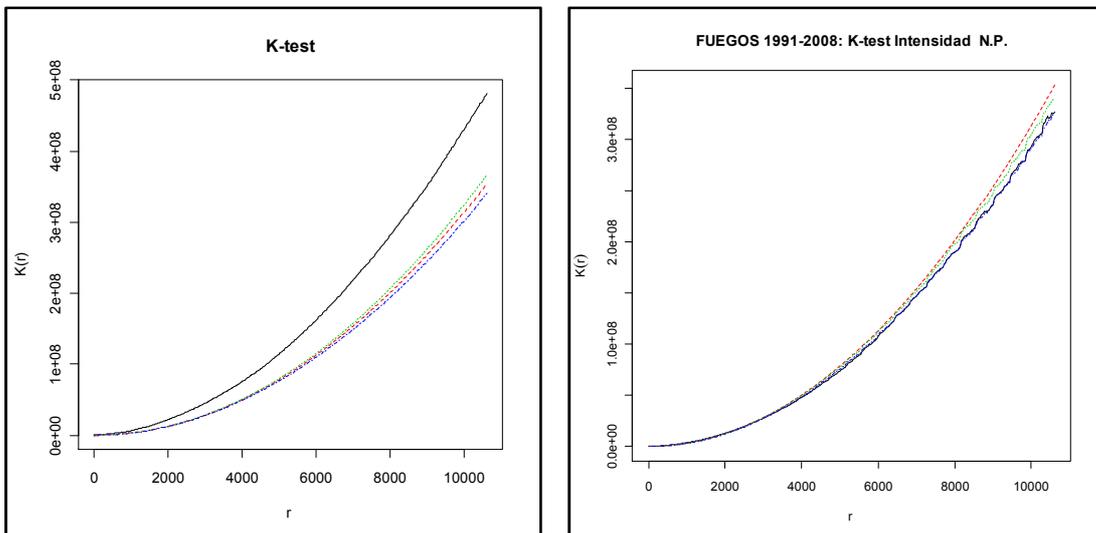


Gráfico 1: K test homogéneo e inhomogéneo, para fuegos. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

En los gráficos 2-5 se presentan los tests para los datos clasificados en función de la causa del fuego. Si comparamos los tests homogéneos e inhomogéneos para cada diseño, vemos que la evidencia en contra de CSR es mucho mayor cuando se supone intensidad constante. Respecto al test basado en la K-función inhomogénea, para las tres causas posibles de fuegos vemos un ligero indicio de diseño agregado con radio de interacción próximo a 2 km

e independencia a partir de esa distancia. El hecho de que para distancias grandes la K-función empírica esté por debajo de las bandas de confianza, que se interpretaría como inhibición entre fuegos a esas distancias, debe observarse con cautela, ya que con t aumentan las fluctuaciones de la K-función y además no es realista estudiar los efectos de modelos que operan a la misma escala que las dimensiones de la región de estudio.

Gráfico 2: K test homogéneo y no homogéneo, no paramétrico (centro) y paramétrico (derecha) para fuegos intencionados. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

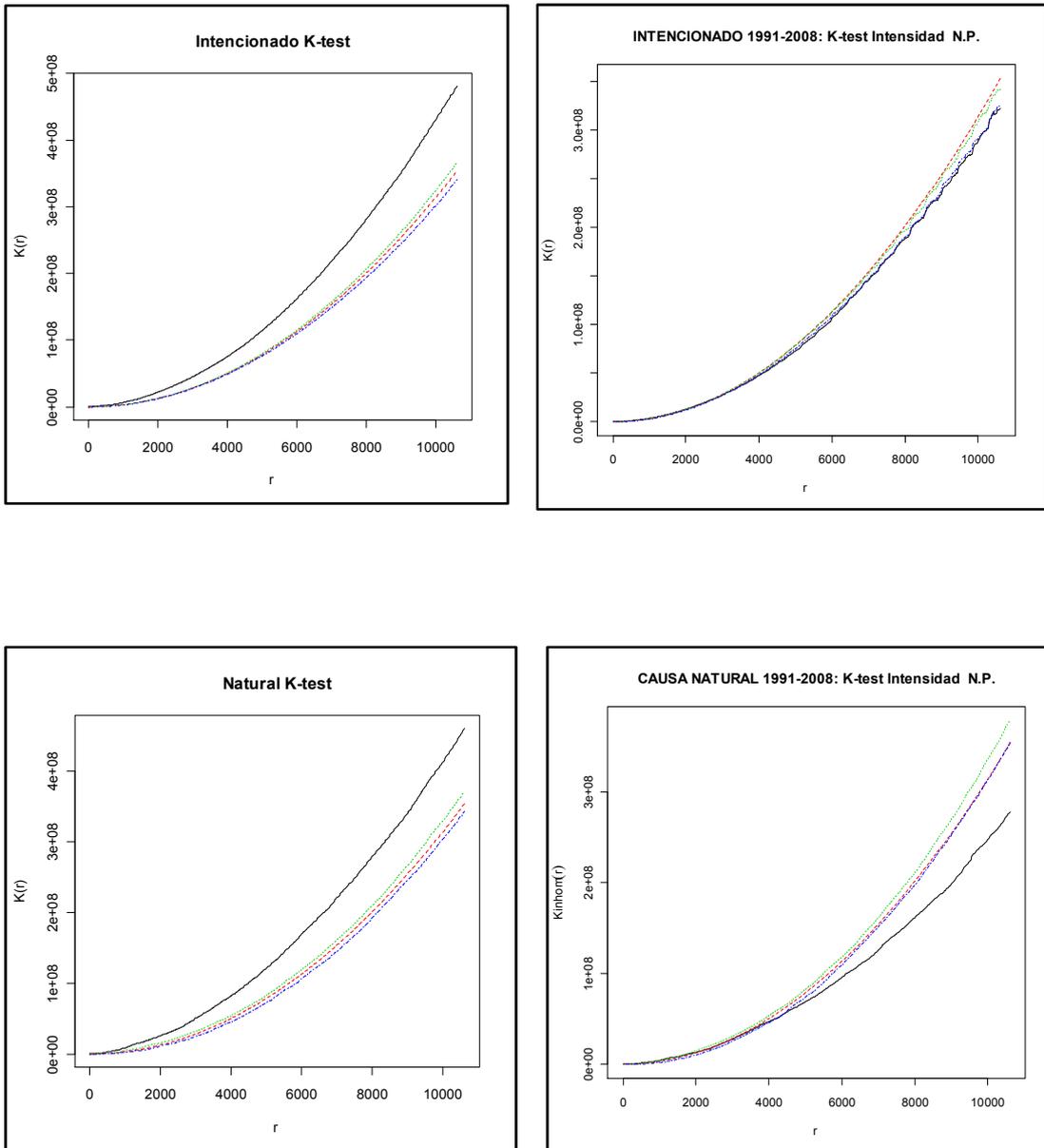


Gráfico 3: K test homogéneo y no homogéneo para fuegos debidos a causas naturales. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

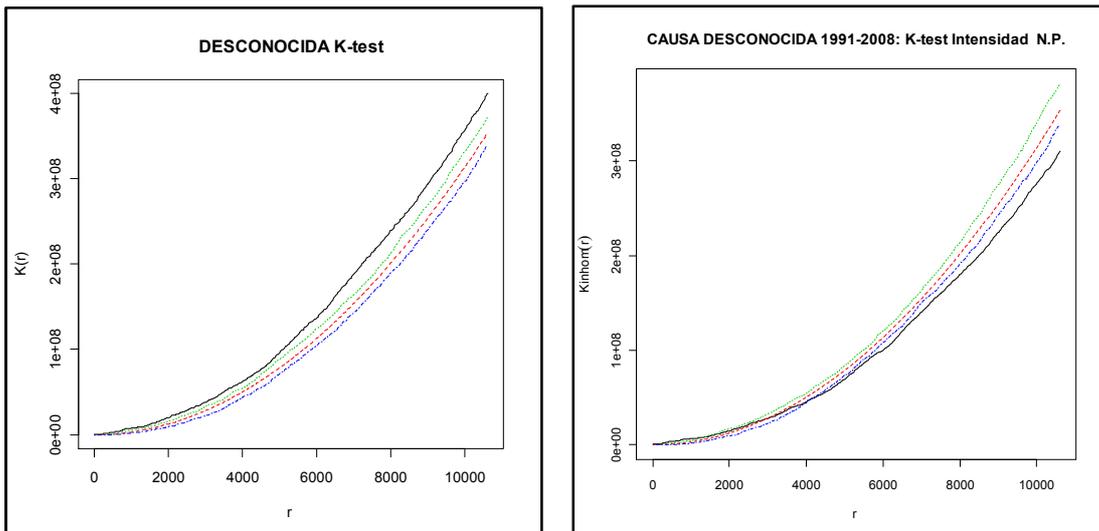


Gráfico 4: K test homogéneo y no homogéneo para fuegos con causa desconocida. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

Por último, en los gráficos 5-7, se muestran los resultados de los tests aplicados a los diseños correspondientes a fuegos clasificados en función de la superficie quemada. Al comparar los test, se observa de nuevo una mayor evidencia en contra de CSR cuando suponemos intensidad homogénea. Si nos fijamos en los gráficos correspondientes al K-test inhomogéneo, la K-función para conatos se mantiene dentro de las envolturas, por lo que podemos decir que el proceso inhomogéneo refleja el comportamiento de estos fuegos. Para incendios y quemas la K-función empírica se mantiene ligeramente por encima de la envoltura superior correspondiente, indicando diseño agregado para estos dos tipos de fuegos, los radios de interacción correspondientes son 2 km para incendios y 8 km para quemas.

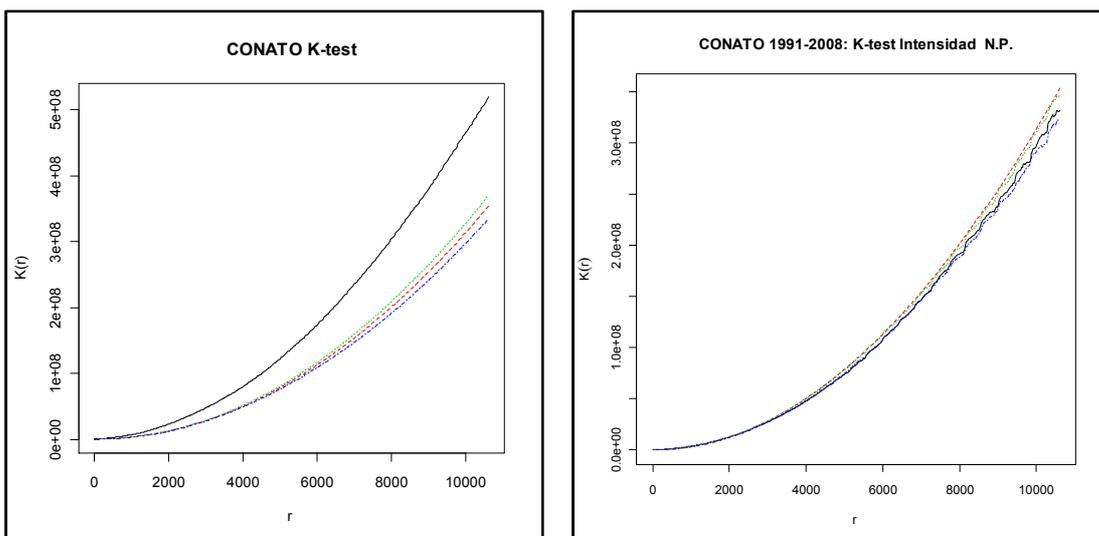


Gráfico 5: K test homogéneo e inhomogéneo para conatos. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

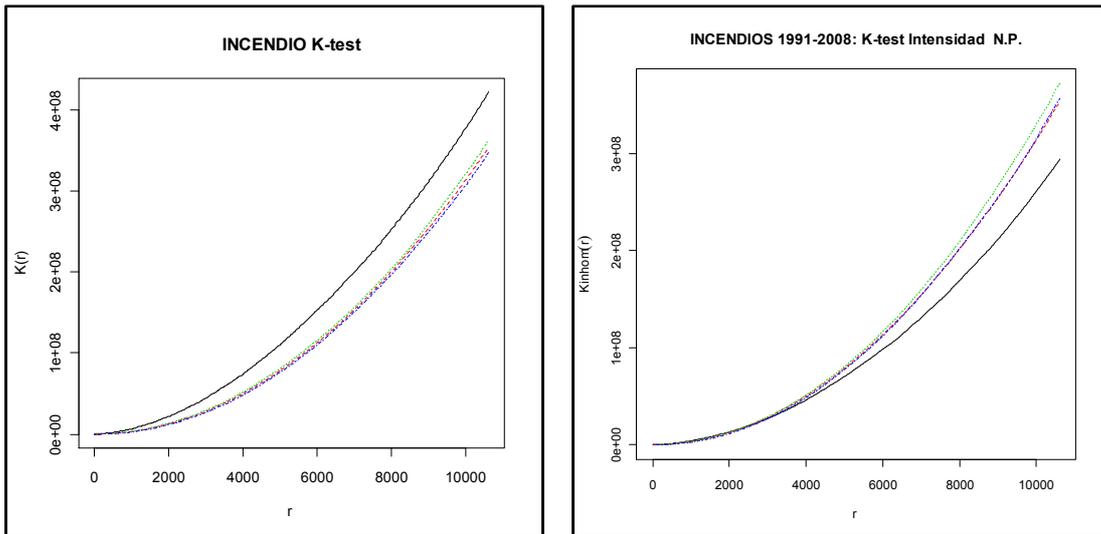


Gráfico 6: K test homogéneo y no homogéneo para fuegos. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

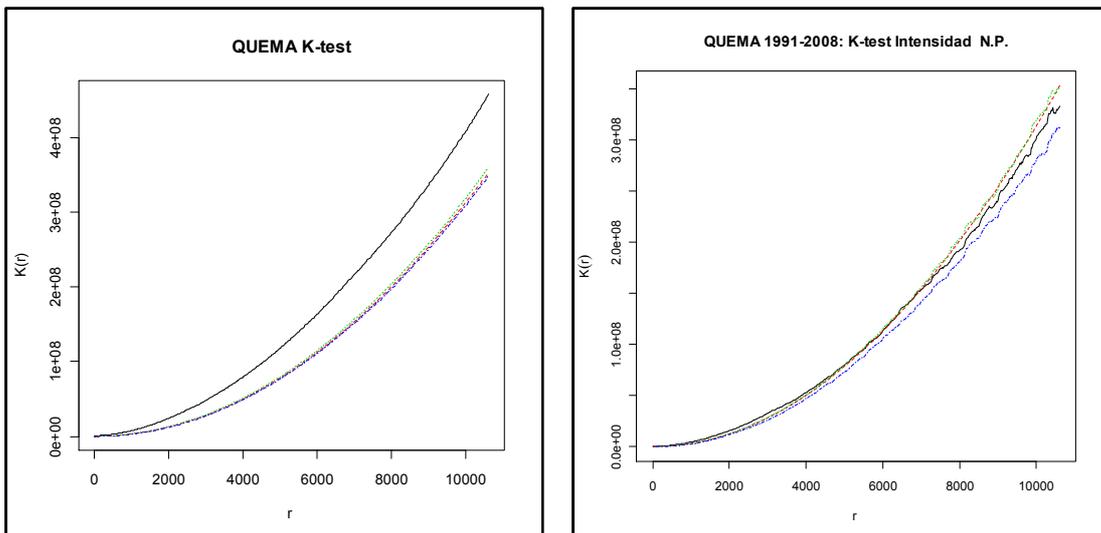


Gráfico 7: K test homogéneo y no homogéneo para quemas. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

Al comparar los resultados obtenidos con el K-test homogéneo y el K-test inhomogéneo, se ha visto que cuando suponemos intensidad constante la evidencia en contra de CSR es mucho mayor. Esto sucede porque en el caso homogéneo la diferencia respecto a CSR puede ser debida tanto a la interacción entre datos como a la tendencia espacial. Por tanto, cuando se rechaza la hipótesis de CSR para un proceso homogéneo es recomendable repetir el análisis asumiendo intensidad no constante antes de plantear un modelo con interacciones.

A continuación se muestran los resultados obtenidos para el análisis de datos correspondiente a los fuegos registrados en 1992, al limitarnos a un único año se ha eliminado el problema del coste computacional, se ha reducido considerablemente el número de eventos

repetidos y se mejora la interpretación práctica de los resultados. La K-función inhomogénea de cada patrón se ha obtenido a partir del estimador no paramétrico de la intensidad variable.

En el gráfico 8 se muestran los K-test correspondientes a los 879 fuegos registrados en 1992. El K-test homogéneo indica claramente diseño agregado, mientras que cuando no se supone intensidad constante tenemos solo ligeros indicios de clustering, por la proximidad de la distribución observada a la envoltura superior, para radios inferiores a 3 km.

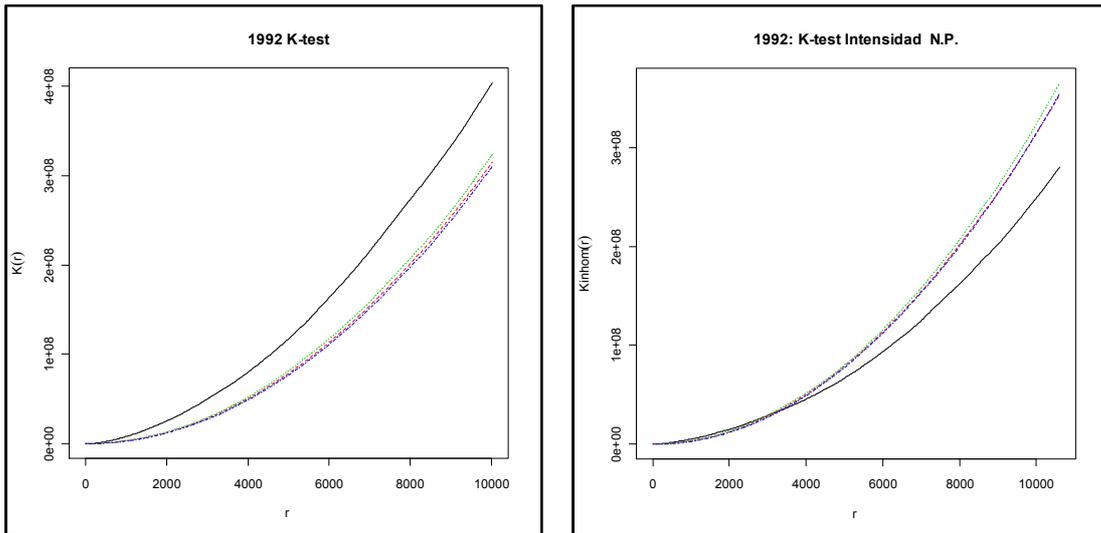


Gráfico 8: K test homogéneo e inhomogéneo para los fuegos registrados en 1992. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

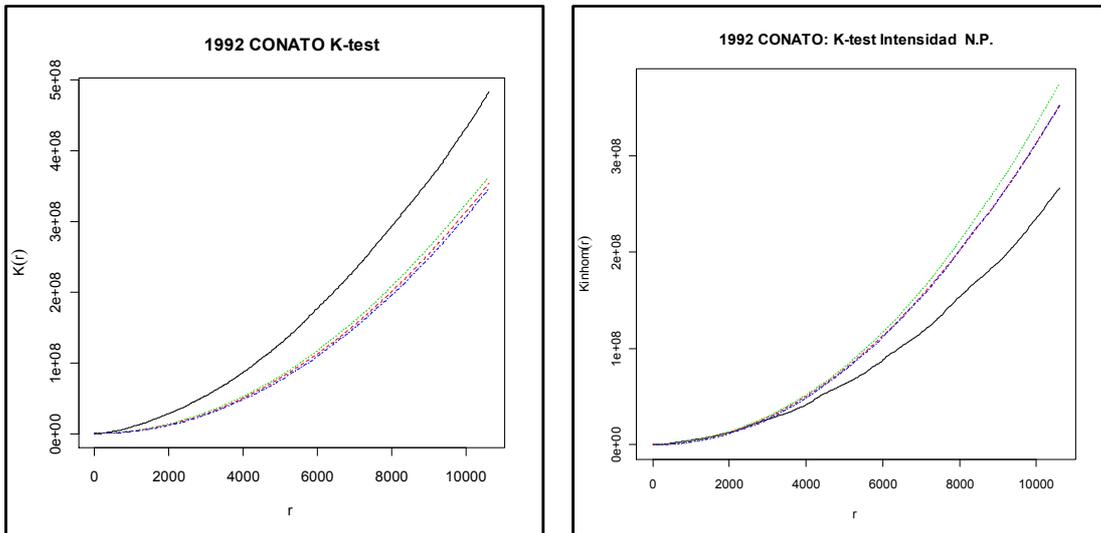


Gráfico 9: K test homogéneo e inhomogéneo para los conatos registrados en 1992. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

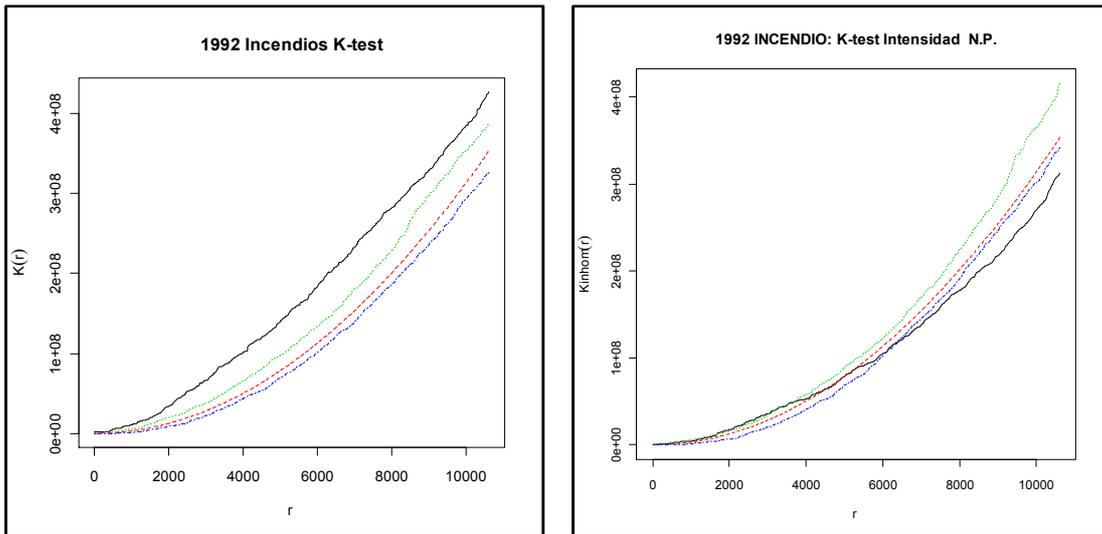


Gráfico 10: K test homogéneo e inhomogéneo para los incendios registrados en 1992. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

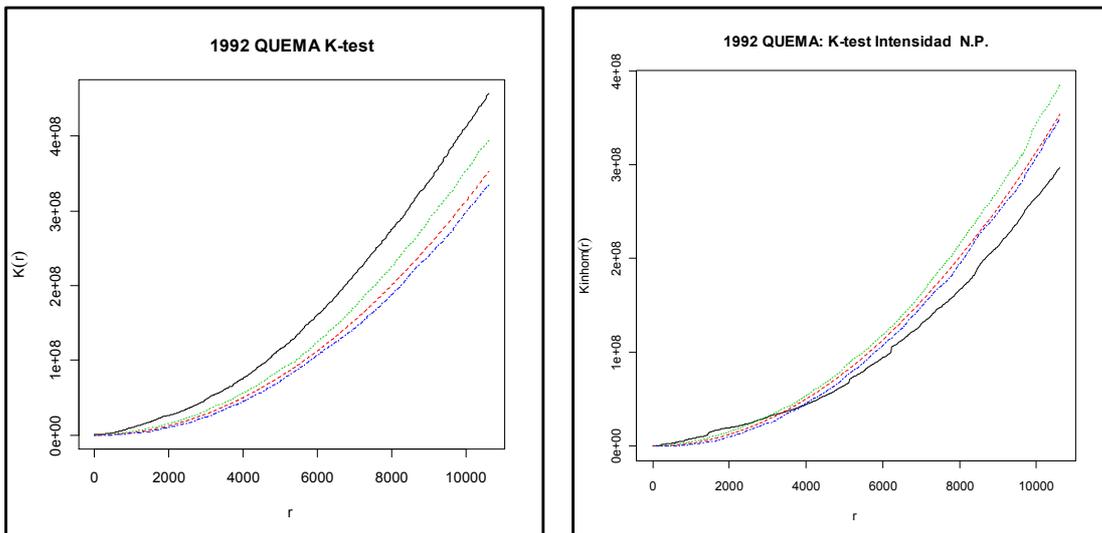


Gráfico 11: K test homogéneo e inhomogéneo para las quemas registradas en 1992. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

Si analizamos los datos en función de la superficie quemada. En el caso homogéneo vuelven a verse evidencias claras de diseño agregado para cada uno de los tipos de fuego. Según el contraste inhomogéneo podría aceptarse que la distribución de los conatos se corresponde con un proceso inhomogéneo de Poisson, mientras que tanto incendios como quemas muestran indicios de diseño agregado con radios de interacción de 4 km y 3 km respectivamente.

Respecto a los fuegos clasificados según causa se ha estudiado únicamente la distribución de los incendios intencionados al contar con muy pocos datos para causa natural (22) o desconocida (21). Dado que los fuegos intencionados en este caso representan el 95% del total de fuegos registrados en 1992, los resultados de los test son prácticamente iguales a los vistos en el gráfico 9 para el conjunto total. Así, vemos de nuevo que si suponemos

homogeneidad se sobreestima la agregación y en el caso inhomogéneo se ven indicios de clustering con radio de interacción de 2 km.

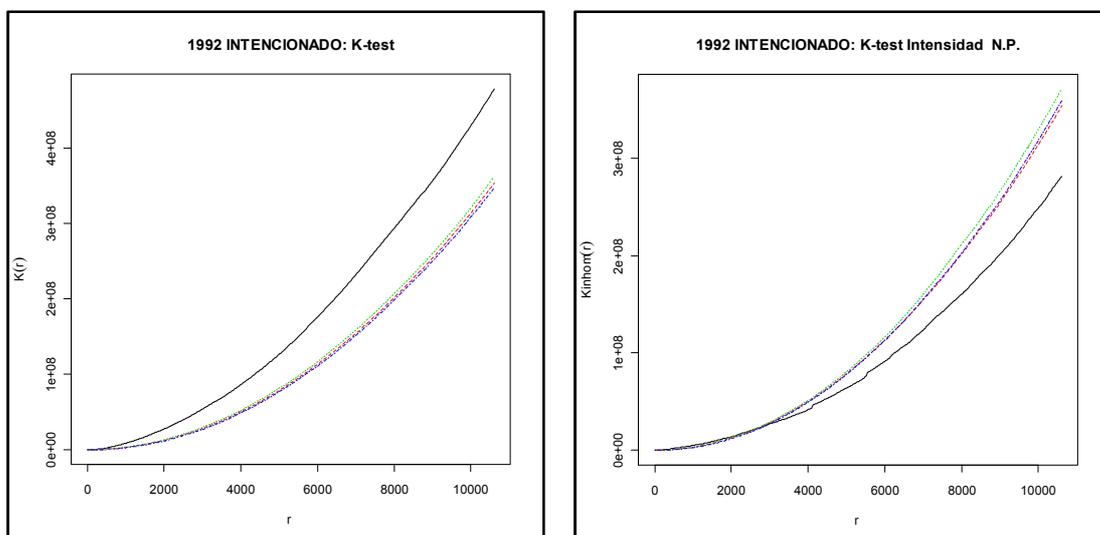


Gráfico 12: K test homogéneo e inhomogéneo para los fuegos intencionados registradas en 1992. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

3. PROCESOS PUNTUALES MULTITIPO

La base de datos de incendios es un proceso puntual espacial marcado por, entre otras variables, año, causa y tipo de fuego. La interacción espacial entre dos tipos de eventos sucede cuando los eventos de distintos tipos aparecen más próximos o alejados de lo que se espera cuando se asume que ambos procesos son independientes. Cuestiones como “¿están los fuegos registrados en 1995 cerca o lejos de los de 1994?”, “¿se han registrado incendios cerca de los conatos?” pueden responderse analizando la función K-cross que mide el número esperado de eventos de tipo 2 en un entorno de un evento arbitrario de tipo 1. A continuación se recuerda la definición formal de esta función y se presentan los contrastes de independencia para procesos multivariantes que hemos utilizado, junto con los resultados obtenidos al aplicar estos tests de independencia a los datos.

3.1. Función K-cross

Como ya se ha visto anteriormente, la función K-cross es la extensión de la K-función a procesos puntuales con marcas. Para un proceso puntual con intensidad constante para cada tipo de eventos, la función K-cross se define como:

$$K_{ij}(t) = \lambda_j^{-1} E \left[\begin{array}{c} \text{eventos de tipo j a una distancia menor que t} \\ \text{de un evento de tipo i arbitrario} \end{array} \right] \quad (6.3)$$

donde λ_j es la intensidad (constante) del proceso marginal formado por los eventos de tipo j. Si el proceso bivalente es estacionario, además de homogéneo, entonces $K_{ij}(t) = K_{ji}(t)$

En la sección 1.3 del capítulo anterior se introduce el estimador de K-cross para procesos puntuales homogéneos. Adoptando la notación de Schabenger & Gotway (2006), el estimador de K_{ij} para un conjunto de eventos de tipos i y j, con intensidades constantes respectivas λ_i, λ_j observadas en la región A, viene dado por la expresión:

$$\hat{K}_{ij}(t) = \frac{1}{\hat{\lambda}_j} \frac{1}{\hat{\lambda}_i |A|} \sum_k \sum_l w_{x_k, y_l}^{-1} I(\|x_k - y_l\| \leq t) \quad (6.4)$$

donde w_{x_k, y_l} es el corrector del efecto frontera y $\hat{\lambda}_i, \hat{\lambda}_j$ los estimadores empíricos de la intensidad constante de cada uno de los procesos marginales. Este es el estimador de máxima verosimilitud para un proceso homogéneo de Poisson.

Si los procesos puntuales univariantes no son homogéneos se puede obtener un estimador de la función K-cross inhomogénea sustituyendo en (6.4) $\hat{\lambda}_i$ y $\hat{\lambda}_j$ por estimadores no paramétricos de las intensidades marginales.

Cuando hay independencia entre los dos tipos de eventos se verifica que $K_{ij}(t) = \pi t^2$ independientemente del diseño marginal de cada uno de los procesos. Esta propiedad sugiere la elaboración de un contraste de independencia basado en la función K-cross análogo al K-test de CSR para procesos univariantes. En lugar de aplicar el test basado en la K-cross es habitual utilizar la L-cross

$$L_{ij}(t) = \sqrt{\frac{K_{ij}(t)}{\pi}}$$

Esta función verifica que bajo independencia de los dos tipos de eventos $L_{ij}(t) = t$, proporcionando un test gráfico de interpretación más sencilla que el basado en la K-función.

3.2. Análisis de interacción de los datos.

La experiencia nos dice que los pirómanos intentan provocar un incendio reiteradamente hasta que lo consiguen. Un modo de ver si esto se refleja en nuestros datos es el análisis de interacciones entre distintos tipos de fuegos, donde los conatos o quemas serían los intentos previos al incendio. Con este fin se han aplicado dos tests gráficos basados en la K-cross. En primer lugar, se ha supuesto homogeneidad en los procesos unidimensionales y, a continuación, se ha aplicado el test inhomogéneo tomando como estimadores de las

intensidades marginales obtenidas en el capítulo anterior. A continuación se presentan, gráficos 13-15, los resultados de la aplicación de ambos test a cada uno de los diseños.

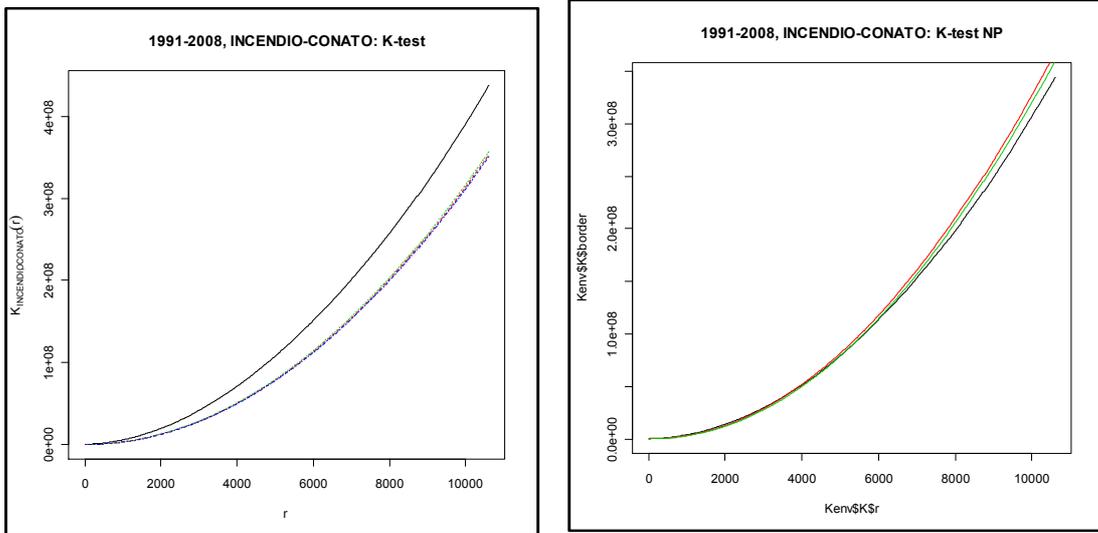


Gráfico 13: K-cross test para el par incendios-conatos. Izquierda: K-función homogénea, distribución empírica (negro), L(t) (azul), U(t) (verde), Media de las simulaciones (rojo). Derecha: k-función inhomogénea, distribución empírica (negro), L(t) (verde), U(t) (rojo).

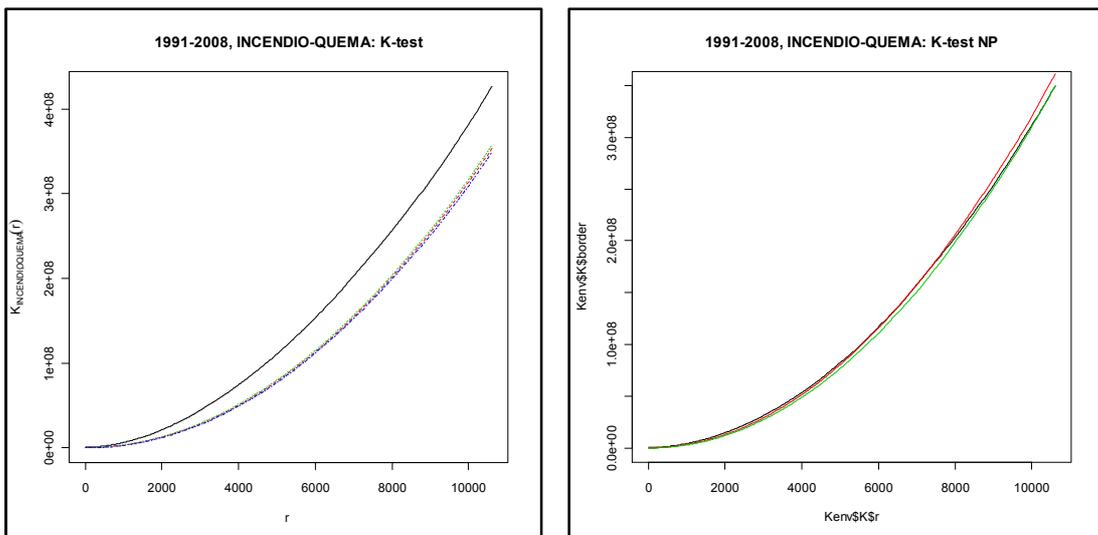


Gráfico 14: K-cross test para el par incendios-quemas. Izquierda: K-función homogénea, distribución empírica (negro), L(t) (azul), U(t) (verde), Media de las simulaciones (rojo). Derecha: k-función inhomogénea, distribución empírica (negro), L(t) (verde), U(t) (rojo).

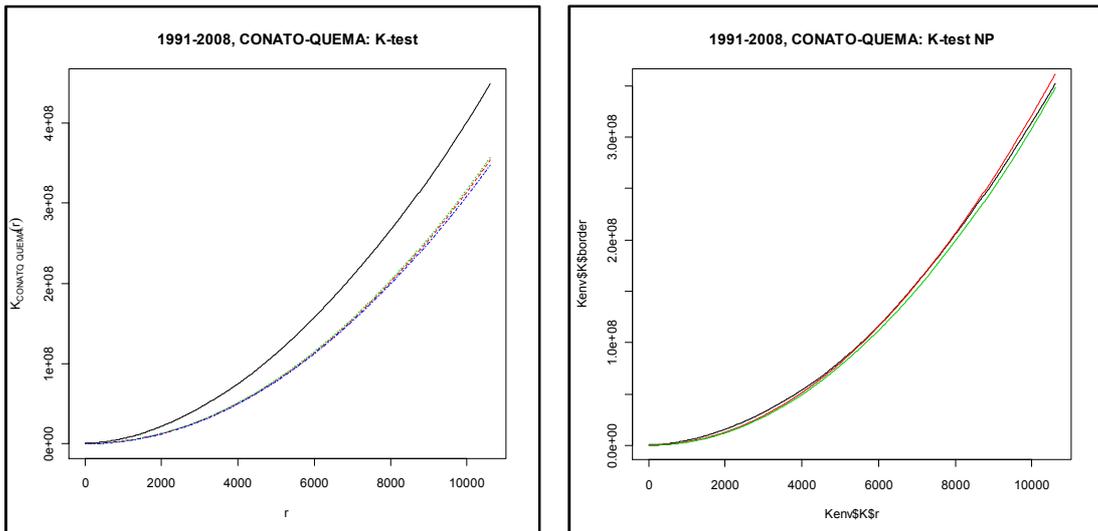


Gráfico 15 : K-cross para el par conato-quema. Izquierda: K-función homogénea, distribución empírica (negro), L(t) (azul), U(t) (verde), Media de las simulaciones (rojo). Derecha: k-función inhomogénea, distribución empírica (negro), L(t) (verde), U(t) (rojo).

Lo primero que se observa en estos gráficos es que, al igual que en el caso unidimensional, si suponemos homogeneidad cuando no la hay, el test basado en la función K-cross interpreta como interacción lo que en realidad es tendencia. Así, mientras para los gráficos correspondientes al test homogéneo la distancia entre la K-cross observada y la envolura superior es muy grande, en el test inhomogéneo esta distancia se reduce considerablemente, aun así siguen viéndose indicios de atracción entre los distintos tipos de fuegos. Los radios de interacción observados son 3 km para incendio- conato y 5 tanto para incendio-quema como para conato-quema.

Se ha realizado el mismo análisis para los fuegos registrados en 1992 ya que la interpretación de los resultados es más sencilla si consideramos fuegos de un único año, al estar midiendo interacción entre eventos próximos en el tiempo.

En los gráficos 16-18 volvemos a ver como al suponer homogeneidad el contraste identifica como interacción lo que en realidad es tendencia. En los contrastes basados en la K-cross inhomogénea la distribución observada es mayor que la envolura superior, indicando interacción positiva entre los distintos pares de procesos. Los radios de interacción observados son 3 km para incendio-conato y 4 tanto para incendio-quema como para conato-quema. Ahora, al contar con datos de un único año y dado que un alto porcentaje de los fuegos son provocados, se puede decir que este resultado avala la teoría de que los pirómanos intentan provocar un incendio varias veces, quemar y conatos, en la misma zona hasta que finalmente lo consiguen.

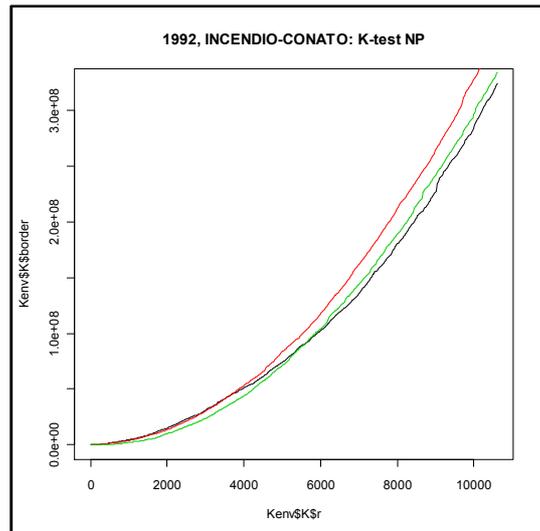
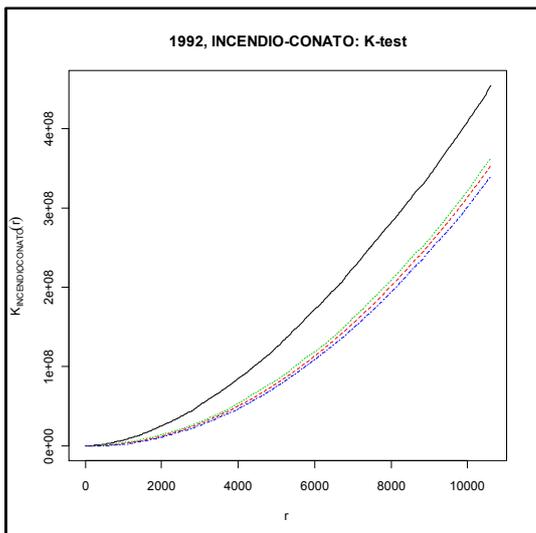


Gráfico 16: K-cross test homogéneo e inhomogéneo para el par incendios-conatos registrados en 1992. Distribución empírica (negro), L(t) (rojo), U(t) (verde)

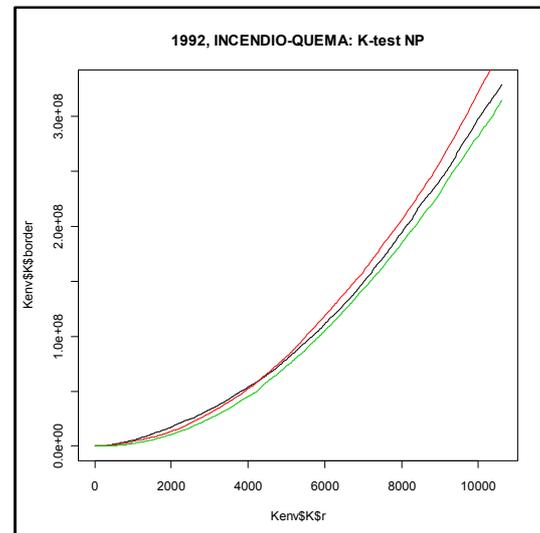
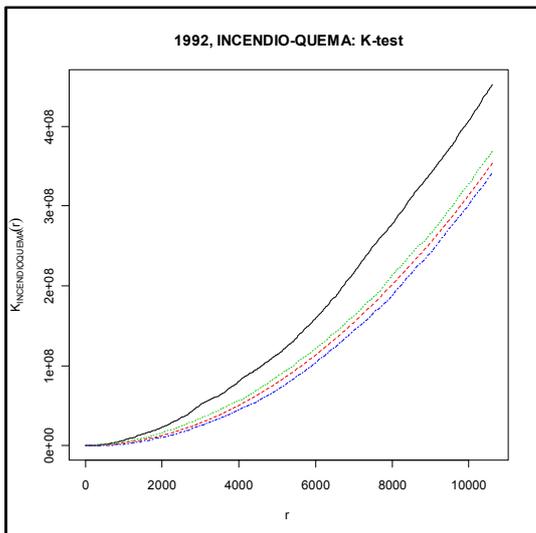


Gráfico 17: K-cross test homogéneo e inhomogéneo para el par incendios-quemas registrados en 1992. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

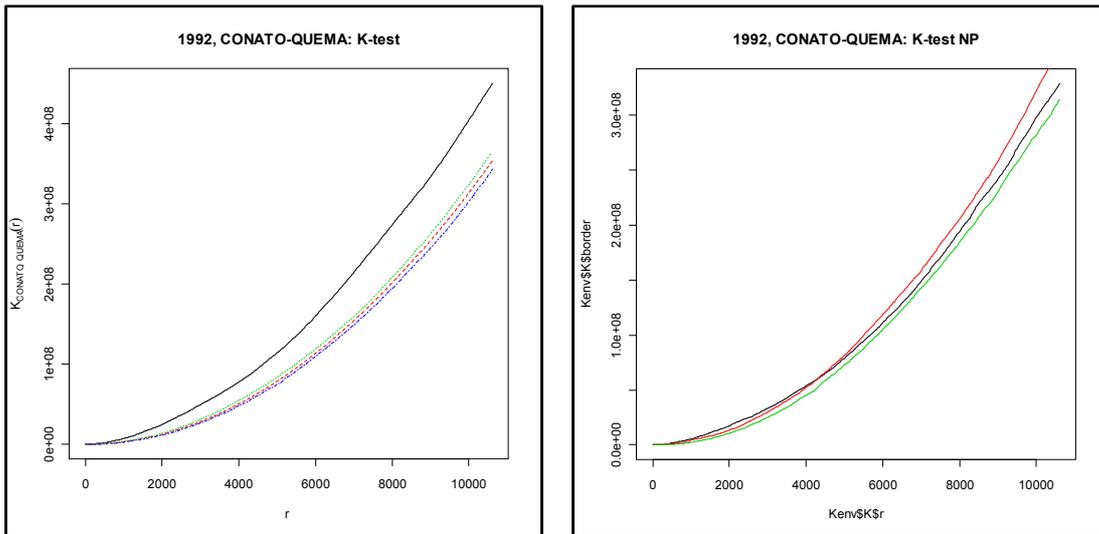


Gráfico 18: K-cross test homogéneo e inhomogéneo para el par conatos-queamas registrados en 1992. Distribución empírica (negro), L(t) (rojo), U(t) (verde)

Otro de los aspectos de gran interés en este estudio es analizar la relación entre las localizaciones de los fuegos entre años, tanto consecutivos como a distancias mayores. En esta ocasión, para analizar la dependencia entre eventos en función del año, se ha utilizado el L-índice, definido por Genton et al (2006), que nos permite presentar en un único gráfico el estudio para varios años. Para un proceso homogéneo de Poisson, el L-índice es una aproximación del área entre $\hat{L}_{ij}(t)$ y la línea de referencia, esto es $L_{ij}(t) = t$ que se obtiene sumando $\hat{L}_{ij}(t) - t$, en este caso para valores de t menores que 11 km. Este índice resulta de gran utilidad para comparar la evolución del área entre $\hat{L}_{ij}(t)$ y t para una sucesión de pares de años situados a la misma distancia.

Igual que se analiza la relación entre distintos tipos de fuegos, también se va a estudiar la independencia suponiendo tanto homogeneidad como intensidad no constante en cada uno de los procesos marginales y calculando el L-índice correspondiente a cada una de las situaciones. Junto con los L-índices se han simulado las envolturas superior e inferior. Si suponemos que ambos procesos son homogéneos de Poisson, se estiman las intensidades marginales $\hat{\lambda}_i$ y $\hat{\lambda}_j$, se simula un proceso de Poisson para cada una de las intensidades. Los puntos simulados con $\hat{\lambda}_i$ se etiquetan con la marca i y los simulados con $\hat{\lambda}_j$ con la marca j . Estos dos conjuntos de puntos se combinan para generar un único proceso. Este procedimiento se repite 100 veces y se calcula $\hat{L}_{ij}(r)$ para cada simulación, el máximo y el mínimo de las 100 $\hat{L}_{ij}(r)$ serán, respectivamente, la envoltura superior e inferior. Si para el proceso observado $\hat{L}_{ij}(r)$ es mayor que la envoltura superior o menor que la inferior tendremos, respectivamente, atracción o inhibición entre los dos tipos de eventos.

Cuando el proceso no es homogéneo, se aplica el mismo algoritmo sustituyendo el estimador de la intensidad constante por un estimador de la intensidad variable tanto en el

cálculo de $\hat{L}_{ij}(r)$ como en las simulaciones para obtener las envolturas, aquí se ha utilizado estimación no paramétrica.

Observando los datos vemos que, en general, años con un gran número de fuegos van seguidos de años con pocos incendios y viceversa, por lo que es razonable analizar la evolución de los fuegos a lo largo de los años para detectar posibles interacciones. Se han comparado pares de años consecutivos (1991-1992, 1992-1993,...), alternos (1991-1993, 1992-1994,...), a distancia 3 (1991-1994, 1992-1995,...), a distancia 4 (1991-1995, 1992-1996,...), a distancia 5 (1991-1996, 1992-1997,...) y a distancia 6 (1991-1997, 1992-1998,...). En el gráfico 19, mediante el L-índice para las distintas sucesiones de datos se aporta información concisa de todos los contrastes realizados. La línea sólida negra y la discontinua azul, unen los L-índices obtenidos en cada par de años, para procesos homogéneos e inhomogéneos, respectivamente. Los segmentos verticales unen los L-índices correspondientes a las envolturas superior e inferior calculadas por simulación. Cuando los L-índices observados están dentro de las bandas, los procesos marginales son independientes, mientras que L-índices observados por encima o por debajo de los límites indican, respectivamente, agregación o inhibición.

Al observar las gráficas, lo primero que debemos comentar es que las bandas para los pares que incluyen los años 2007 y 2008 son mucho más anchas que las demás debido a que en estos dos años se han registrado muy pocos incendios. Como se ha visto en análisis previos, cuando suponemos procesos homogéneos se sobreestima la interacción positiva, sin embargo cuando se utiliza el estimador no paramétrico de la intensidad variable, el L-índice observado se mantiene dentro de la banda delimitada por las envolturas, esto indica que no hay interacción entre años. Por tanto, podemos decir que la distribución de los fuegos cada año no es homogénea y que conocer la localización de los fuegos un año no sería un buen predictor los de años posteriores.

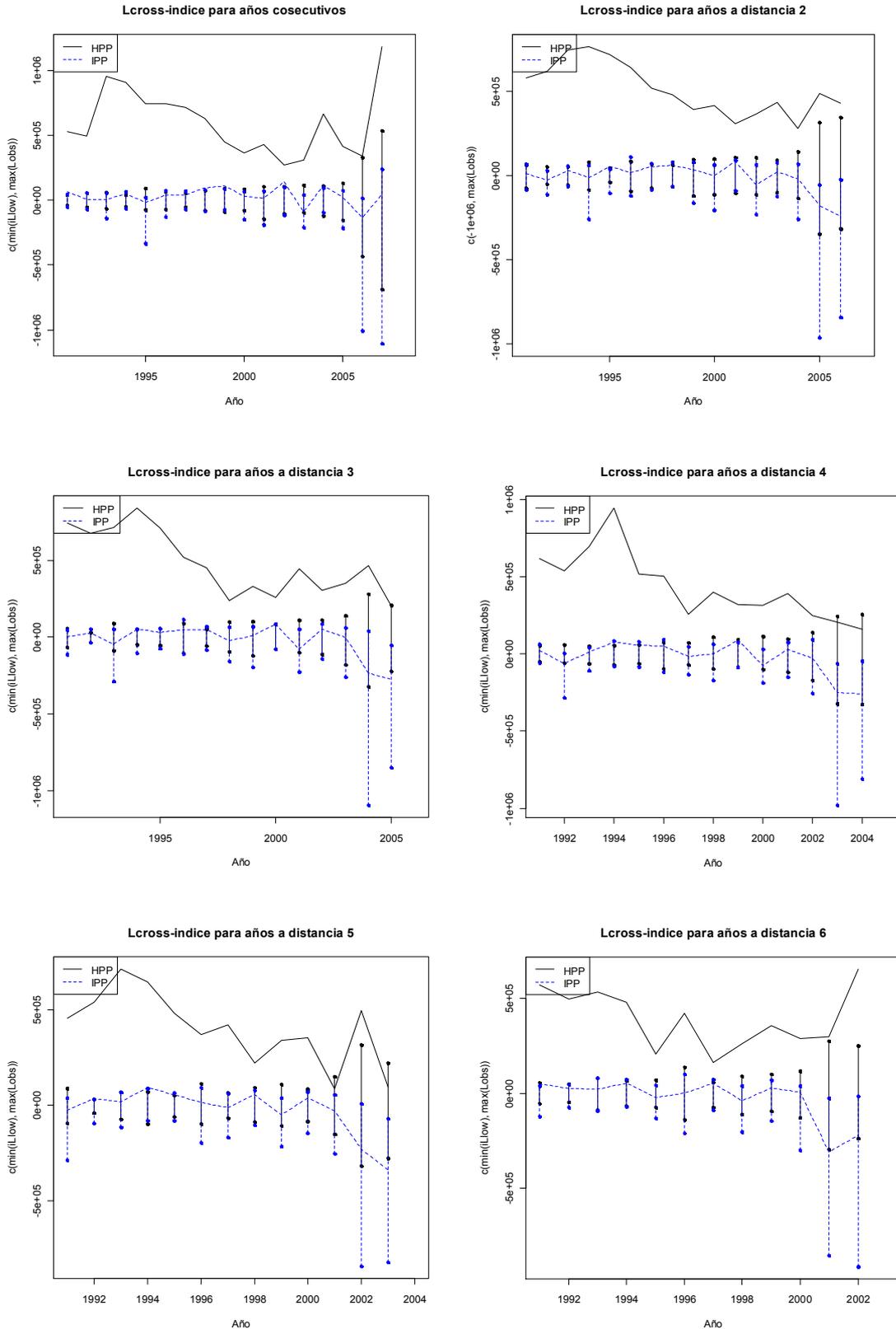


Gráfico 19: L-índice para procesos homogéneos e inhomogéneos con intensidad no paramétrica. Las líneas verticales representan el L-índice para las envolturas simuladas, cada año del eje X corresponde al primer año del par estudiado.

AJUSTE DE MODELOS

1. AJUSTE DE MODELOS POR MÁXIMA VEROSIMILITUD

La función de máxima verosimilitud juega un papel fundamental en la inferencia estadística, tanto clásica como bayesiana. El punto débil de este método es que para algunos modelos la función que proporciona este resulta intratable, motivo por el cual a veces se prefieren los métodos no paramétricos. Esta dificultad de los métodos de máxima verosimilitud se puede salvar aplicando métodos de Monte Carlo para estimar la función correspondiente.

Ogata y Tanemura (1981,1985,1985,1986) y Penttinen (1984) desarrollaron métodos de máxima verosimilitud para procesos puntuales que incorporan tanto inhomogeneidad como dependencia, clústering o regularidad. Los métodos de máxima verosimilitud tiene un coste computacional elevado y utilizan algoritmos de simulación específicos para cada posible modelo. El coste computacional es especialmente alto para procesos inhomogéneos de Poisson, debido a que la dimensión del vector de parámetros puede ser muy grande y a la complejidad de las simulaciones. Este problema dificulta la práctica de ajustar modelos alternativos para un conjunto de datos y la introducción de funciones de suavización como términos del modelo.

En este trabajo se ha aplicado el método de máxima verosimilitud desarrollado por Baddeley y Turner (2000), que es una extensión del propuesto por Berman y Turner (1992) a la familia de procesos puntuales de Gibbs de familia exponencial. En este método se propone un estimador basado en pseudoverosimilitud como alternativa al estimador de máxima verosimilitud (MLE).

1.1. Verosimilitud.

Sea $x = \{x_1, \dots, x_n\}$ una realización del proceso puntual espacial X en la región acotada $W \subset \mathbb{R}^2$. Como hemos visto anteriormente, los distintos modelos pueden definirse en función de su verosimilitud respecto al proceso homogéneo de Poisson con intensidad unitaria. Así, supongamos que $f(x; \theta)$ es la densidad de X respecto al proceso homogéneo de Poisson con intensidad 1 en W . Asumimos también que $f(x; \theta)$ es hereditaria, es decir $f(x; \theta) > 0 \Rightarrow f(y; \theta) > 0$ para todo $y \subseteq x$. Los procesos puntuales que verifican estas condiciones definen la clase de procesos de Gibbs en W y su distribución depende del parámetro θ perteneciente a una familia $\Theta \subset \mathbb{R}^d$.

El proceso homogéneo de Poisson con intensidad $\lambda > 0$ tiene densidad:

$$f(x, \lambda) = e^{-(\lambda-1)|W|} \lambda^{n(x)}$$

donde $n(x)$ es el número de eventos en x , de esta función se obtiene el estimador de máxima verosimilitud $\hat{\lambda} = n(x)/|W|$.

Consideremos ahora un proceso inhomogéneo de Poisson con función de intensidad $\lambda: W \rightarrow \mathbb{R}$. En un modelo estadístico la intensidad $\lambda_\theta(x)$ refleja la tendencia o su dependencia respecto a covariables a través del parámetro θ . La densidad de un proceso inhomogéneo de Poisson es:

$$f(x, \lambda) = \prod_{i=1}^{n(x)} \lambda_\theta(x_i) \exp\left(-\int_W (\lambda_\theta(u)-1) du\right) \quad (7.1)$$

Maximizar (7.1) requiere generalmente utilizar métodos de optimización iterativos.

Los procesos de interacción par en W con intensidad $b_\theta: W \rightarrow \mathbb{R}^+$ y función de interacción $h_\theta: W \times W \rightarrow \mathbb{R}$ tienen densidad:

$$f(x, \lambda) = \alpha(\theta) \prod_{i=1}^{n(x)} b_\theta(x_i) \prod_{i<j} h_\theta(x_i, x_j) \quad (7.2)$$

donde $\alpha(\theta)$ es la constante normalizadora. Cuando $h_\theta = 1$ el modelo se reduce a un proceso inhomogéneo de Poisson con intensidad $b_\theta(u)$.

La constante normalizadora $\alpha(\theta)$ suele ser una función intratable de θ , por lo que la evaluación y maximización de la función de verosimilitud resulta muy complicada.

1.2. Pseudoverosimilitud.

Como acabamos de ver, la intratabilidad de la constante normalizadora que aparece en los modelos exponenciales dificulta enormemente evaluar y maximizar la función de verosimilitud para estimar los parámetros del modelo. Una alternativa para evitar esta dificultad es la función de pseudoverosimilitud. Para construir esta función se necesita la intensidad condicional de Papangelou $\lambda(u, x)$ de X en la región W , que se interpreta como la probabilidad de tener un evento en u condicionado a que el resto del proceso es x . Para cualquier proceso de Gibbs, la intensidad condicionada en el punto $u \in W$ es:

$$\lambda(u, x) = \begin{cases} f(x \cup \{u\})/f(x); & u \notin X \\ f(x)/f(x - \{u\}); & u \in X \end{cases} \quad (7.3)$$

Por ejemplo, un proceso inhomogéneo de Poisson con intensidad $\lambda(\cdot)$ tiene intensidad condicionada $\lambda(u; x) = \lambda(u)$ en todo punto. Para un proceso de Gibbs general la $\lambda(u; x)$ depende de x . Así, la intensidad condicional de un proceso de interacción par es:

$$\lambda(u; x) = b_\theta(u) \prod_{\substack{i=1 \\ x_i \neq u}}^{n(x)} h_\theta(u, x_i) \quad (7.4)$$

Nótese que $\lambda(\cdot; x)$ es discontinua en los eventos del proceso y que se ha eliminado la constante normalizadora de (7.2).

Besag (1977) define la pseudoverosimilitud de un proceso con intensidad condicional $\lambda_\theta(u; x)$ en una región $W \subset \mathbb{R}^2$ mediante:

$$PL_\theta(\theta; x) = \left(\prod_{i=1}^{n(x)} \lambda_\theta(u; x_i) \right) \exp\left(-\int_W \lambda_\theta(u; x_i) du\right) \quad (7.5)$$

Para procesos homogéneos la pseudoverosimilitud coincide con la verosimilitud (7.1) salvo por la constante $e^{-|W|}$. Para un proceso de interacción par, la pseudoverosimilitud es:

$$PL_\theta(\theta; x) = \left(\prod_{i=1}^{n(x)} b_\theta(x_i) \prod_{i < j} h_\theta(x_i, x_j) \right) \exp\left(-\int_W b_\theta(u) \prod_{i < j} h_\theta(u, x_i) du\right) \quad (7.6)$$

donde la constante normalizadora de la función de verosimilitud ha sido reemplazada por la integral de la parte derecha de (7.6) como si el proceso fuese Poisson.

Baddeley y Turner (2000) implementaron un método computacional para aproximar el estimador de máxima pseudoverosimilitud, adaptando la técnica para máxima verosimilitud de procesos inhomogéneos propuesto por Berman y Turner (1992). Para un proceso de Gibbs con intensidad condicional $\lambda_\theta(u; x)$, se aproxima la integral de la expresión (7.5) mediante una suma finita usando una regla de cuadratura.

$$\int_A \lambda_\theta(u, x_i) du \approx \sum_{j=1}^m \lambda_\theta(u_j, x) w_j \quad (7.7)$$

donde $\{u_j, j=1, \dots, m\}$ son puntos en W y $w_j > 0$ son los pesos de cuadratura cuya suma es $|W|$. Esto proporciona la siguiente aproximación de la log-pseudoverosimilitud:

$$\log PL(\theta; x) = \log \sum_{i=1}^{n(x)} \lambda_\theta(x_i, x) - \sum_{j=1}^m \lambda_\theta(u_j, x) w_j \quad (7.8)$$

Cuando los puntos de cuadratura $\{u_j, j=1, \dots, m\}$ incluyen a los eventos observados $\{x_1, \dots, x_n\}$ se puede reescribir (7.8) como:

$$\log PL(\theta; x) = \log \sum_{i=1}^{n(x)} (y_j \log \lambda_j - \lambda_j) w_j \quad (7.9)$$

donde $\lambda_j = \lambda_\theta(u_j)$ e $y_j = z_j/w_j$ y

$$z_j = \begin{cases} 1 & \text{si } u_j \in x \\ 0 & \text{si } u_j \notin x \end{cases}$$

La expresión (7.9) se puede maximizar usando software estándar para ajustar modelos lineales generalizados, pues la intensidad condicional $\lambda_\theta(u; x)$ para un diseño fijo x , se relaciona con una variable explicativa mediante

$$g(\lambda_\theta(u; x)) = \theta^T S(u, x) \quad (7.10)$$

donde g es la función link, generalmente se utiliza el link logarítmico, y $S(u, x)$ es el vector de covariables definidas en cada punto $u \in A$.

2. ANÁLISIS DE BONDAD DE AJUSTE

Como hemos visto en la sección anterior, el método de máxima pseudoverosimilitud permite ajustar un amplio rango de modelos para procesos puntuales. El siguiente paso será analizar la bondad de ajuste del modelo aplicado. Uno de los métodos posibles para esto es emplear el test gráfico basado en K_{inhom} , donde se comparaba la K-función observada para los datos con las envolturas superior e inferior de la K-función inhomogénea para simulaciones del proceso bajo el modelo seleccionado, que ya hemos utilizado para los estimadores no paramétricos de la intensidad. Este test, al igual que otros basados en características de segundo orden, es adecuado para detectar interacciones entre eventos pero no para detectar inhomogeneidad. Por tanto, será necesario utilizar métodos de contraste de bondad de ajuste que analicen la intensidad del proceso y no sólo las características de segundo orden.

Baddeley et al (2005) pusieron de manifiesto esta necesidad y definieron residuos para procesos puntuales espaciales como extensión de los residuos para procesos temporales de uso habitual en análisis de supervivencia. La clave de esta extensión está en sustituir la intensidad condicional usual del proceso por la intensidad condicional de Papangelou. Una vez definidos los residuos, el hecho de que la intensidad condicional de Papangelou se pueda entender como la respuesta media de un GLM, permite desarrollar de modo inmediato gráficos de diagnóstico a partir de los residuos para estudiar la bondad de ajuste de los modelos.

2.1. Intensidad condicional de Papangelou.

Dado un proceso puntual espacial, la falta de un orden natural en el espacio bidimensional no permite generalizar de modo natural la intensidad condicional de los procesos temporales o espacio-temporales, donde se condiciona al proceso antes del instante t . La medida equivalente en procesos puntuales espaciales es la intensidad condicional de Papangelou $\lambda(u; x)$, que para un proceso puntual finito X en W con densidad $f(x)$ que verifica la propiedad hereditaria se define como:

$$\lambda(u; x) = \begin{cases} f(x \cup \{u\}) / f(x) & \text{si } u \notin x \text{ y } f(x) > 0 \\ 0 & \text{si } f(x) = 0 \\ \lambda(u; x \setminus \{u\}) & \text{si } u \in x \end{cases} \quad (7.11)$$

y se interpreta como la probabilidad de tener un evento en un entorno infinitesimal del punto u , condicionado a que el resto del proceso coincide con x .

Recordemos que un proceso puntual es un proceso de Gibbs cuando su densidad respecto al proceso homogéneo de Poisson con intensidad unitaria se podía expresar en forma exponencial, esto es:

$$f(x) = \exp\{V_0 + V_1(x) + V_2(x) + \dots + V_k(x)\} \quad (7.12)$$

donde $V_1(\cdot)$ es el término correspondiente a la intensidad del proceso y $V_{\geq 2}(\cdot)$ refleja las interacciones de distinto orden entre los eventos. De (7.12) se deduce la siguiente representación para la intensidad de Papangelou:

$$\log \lambda(u; x) = v_1(x) + \sum_i v_2(\{u; x_i\}) + \sum_{i < j} v_3(\{u; x_i, x_j\}) + \dots \quad (7.13)$$

La intensidad condicionada de Papangelou determina unívocamente la densidad $f(x)$ y viceversa. Para procesos puntuales de Markov es preferible construir el modelo del proceso a partir de λ en lugar de mediante $f(x)$, dado que la intensidad condicional juega el mismo papel que las características locales en los procesos aleatorios de Markov.

Para cualquier función no negativa $h(u, x)$ los procesos puntuales verifican que:

$$E \left[\sum_{x \in X} h(u, x \setminus \{x_i\}) \right] = E \left[\int_W h(u, X) \lambda(u, X) du \right] \quad (7.14)$$

Esta igualdad, denominada condición de Georgii-Nguyen-Zessin (GNZ), va a ser la base para desarrollar los gráficos de diagnóstico.

2.2. Residuos para procesos puntuales espaciales.

Consideremos un modelo paramétrico para un proceso puntual espacial X con densidad f_θ , se define la **Innovación del proceso** mediante:

$$I_\theta(B) = n(X \cap B) - \int_B \lambda_\theta(u; X) du \quad (7.15)$$

para cualquier subconjunto $B \subset W$, $n(X \cap B)$ es el número de eventos del proceso que están en B . La innovación es una medida (aleatoria) con peso 1 en cada evento del proceso y $-\lambda_\theta(u; X)$ en los demás puntos de W . Tomando $h(u, x) = 1$ ($u \in B$) en la expresión (7.14) se tiene que

$$E[I_\theta(B)] = 0$$

Podemos reescribir la condición GNZ mediante:

$$E_\theta \left[\int_B h(u, X \setminus \{u\}) dI_\theta(u) \right] = 0$$

esto se corresponde con la propiedad de martingala para la innovación en procesos temporales y espacio-temporales. La innovación mide el error que cometemos al estimar el proceso con el modelo propuesto, esta medida no se puede calcular en la práctica al desconocer el proceso puntual que da lugar al diseño observado.

En la práctica, cuando se ajusta un modelo paramétrico a un diseño observado x , el parámetro desconocido θ se sustituye por un estimador $\hat{\theta} = \hat{\theta}(x)$, usando este estimador se definen los **residuos directos** mediante:

$$R_{\hat{\theta}}(B) = n(x \cap B) - \int_B \hat{\lambda}(u; x) du \quad (7.16)$$

donde $\hat{\lambda} = \lambda_{\hat{\theta}}$. Los incrementos de $R_{\hat{\theta}}$ se corresponden con los residuos de un modelo lineal. Los residuos $R_{\hat{\theta}}(B)$ son una medida (aleatoria) con peso 1 en cada evento del proceso y $-\hat{\lambda}(u; X)$ en los demás puntos de W . Los residuos permiten medir la bondad de ajuste del modelo propuesto.

Otros autores habían propuesto con anterioridad medidas de diagnóstico que solo se podían calcular para los eventos del proceso, los residuos propuestos por Baddeley y Turner (2005) se pueden calcular en cualquier punto de la región W . Se debe tener en cuenta que la información respecto a un proceso puntual no se reduce a la posición de los eventos, su ausencia en otras localizaciones también aporta información.

En el análisis de modelos estadísticos es habitual reescalar los residuos directos con el fin de obtener residuos estandarizados. El análogo en este contexto es escalar las sumas de residuos $R_{\hat{\theta}}$. Esto se hace sin más que cambiar la función h en la fórmula de GNZ (7.14), así para cualquier función no negativa $h(u, X)$ se define la innovación reponderada de un proceso puntual espacial con intensidad condicional de Papangelou λ como:

$$I_{\theta}(B, h, \lambda) = \sum_{x_i \in X \cap B} h(x_i, X \setminus \{x_i\}) - \int_B h(u, x) \lambda(u; X) du \quad (7.17)$$

Se puede interpretar $\Lambda I(x_i) = h(x_i, X \setminus \{x_i\})$ como el incremento de la innovación (error) asociado a un evento x_i del proceso X y $dI_{\theta}(u) = h(u, x) \lambda(u; X)$ como el incremento asociado a un punto $u \in W$. Por (7.14) la innovación tiene media 0.

Una vez ajustado un modelo paramétrico al diseño observado x mediante el estimador $\hat{\theta} = \hat{\theta}(x)$ y estimada la intensidad $\hat{\lambda}(u; x) = \lambda_{\hat{\theta}}(u; x)$, para cualquier función no negativa que depende también del parámetro $\hat{h}(u, x) = h_{\hat{\theta}}(u, x)$, se definen los residuos reponderados según h mediante:

$$R_{\hat{\theta}}(B, \hat{h}, \hat{\theta}) = I(B, \hat{h}, \hat{\theta}) = \sum_{x_i \in X \cap B} \hat{h}(x_i, X \setminus \{x_i\}) - \int_B \hat{h}(u, x) \hat{\lambda}(u; x) du \quad (7.18)$$

La media de la innovación es 0 y se espera que la media de los residuos también sea 0 cuando el modelo es correcto.

$$E[R_{\hat{\theta}}(B, \hat{h}, \hat{\theta})] = 0 \quad (7.19)$$

La elección $h(u, x) = 1$ da lugar a la innovación y los residuos directos (7.15) y (7.16).

Por analogía con los residuos de Pearson para regresión de Poisson log-lineal, se considera la función de pesos $h(u, x) = 1/\sqrt{\lambda(u, x)}$, dando lugar a la **medida de Innovación de Pearson**:

$$I_{\theta}\left(B, \frac{1}{\sqrt{\lambda}}, \lambda\right) = \sum_{x_i \in X \cap B} \frac{1}{\sqrt{\lambda(x_i, X)}} - \int_B \sqrt{\lambda(x_i, X)} du \quad (7.20)$$

que tiene media 0. Los **residuos de Pearson** correspondientes son:

$$R_{\hat{\theta}}\left(B, \frac{1}{\sqrt{\hat{\lambda}}}, \hat{\lambda}\right) = \sum_{x_i \in X \cap B} \frac{1}{\sqrt{\hat{\lambda}(x_i, x)}} - \int_B \sqrt{\hat{\lambda}(x_i, x)} du \quad (7.21)$$

Para que esta medida esté bien definida debe verificarse que $\hat{\lambda}(x_i, x) > 0, \forall x_i \in x$.

A continuación se presentan algunas propiedades notables de los residuos de un proceso puntual espacial.

- Los residuos suman 0: cuando se ajusta un modelo sin tendencia, con intensidad condicional log-lineal $\lambda_{\theta}(u, x) = \exp\{\beta + \eta T(u, x)\}$ donde $\theta = (\beta, \eta)$ y $T(u, x)$ no es constante. Si se ajusta el modelo por máxima verosimilitud se tiene que $R_{\hat{\theta}}(W) = 0$.
- Media de los residuos: supongamos que se ha ajustado un modelo paramétrico al diseño observado x mediante el estimador $\hat{\theta} = \hat{\theta}(x)$, donde x es una realización de

un proceso puntual X cuya densidad satisface la propiedad hereditaria. Entonces los residuos (7.18) tiene esperanza:

$$E\left[R\left(B, \hat{h}, \hat{\theta}\right)\right] = \int_B E\left[h_{\hat{\theta}(X \cup \{u\})}(u, X)\lambda(u; X) - h_{\hat{\theta}(X)}(u, X)\lambda_{\hat{\theta}(X)}(u; X)\right] du \quad (7.22)$$

En particular, para los residuos directos y de Pearson, se tiene:

$$E\left[R_{\hat{\theta}}\left(B, 1, \hat{\lambda}\right)\right] = E\left[\int_b (\lambda(u; X) - \lambda_{\hat{\theta}}(u; X)) du\right] \quad (7.23)$$

$$E\left[R_{\hat{\theta}}\left(B, \frac{1}{\sqrt{\hat{\lambda}}}, \hat{\lambda}\right)\right] = \int_B E\left[\frac{\lambda(u; X)}{\sqrt{\lambda_{\hat{\theta}(X \cup \{u\})}(u; X)}} - \sqrt{\lambda_{\hat{\theta}(X)}(u; X)}\right] du \quad (7.24)$$

Dado que la intensidad del proceso puntual es $\lambda(u) = E[\lambda(u; x)]$ se puede interpretar (7.23) como una medida del sesgo negativo del estimador de la intensidad del proceso.

- Varianza de los residuos: la expresión de las varianzas de la medida de innovación y de los residuos se obtiene a partir de la intensidad condicional de dos puntos:

$$\lambda(u, v; x) = \lambda(u; x \setminus \{v\})\lambda(v; x \cup \{u\}) = f(x \cup \{u, v\}) / f(x \setminus \{u, v\})$$

La varianza de la innovación será:

$$\text{var}\left[I(B, h, \lambda)\right] = \int_B E\left[h(u, X)^2 \lambda(u; X) du\right] + \int_B \int_B E\left[S(u, v, X)\right] \quad (7.25)$$

Donde

$$S(u, v, X) = \lambda(u; x)\lambda(v; x)h(u, x)h(v, x) + \lambda(u, v; x)h(v, x \cup \{u\})\left[h(u, x \cup \{v\}) - 2h(u, x)\right]$$

En el caso particular de un proceso homogéneo de Poisson, se tiene:

$$\text{var}\left[I(B, 1, \lambda)\right] = \int_B \lambda(u) du \quad (7.26)$$

$$\text{var}\left[I\left(B, \frac{1}{\sqrt{\lambda}}, \lambda\right)\right] = |B| \quad (7.27)$$

El cálculo de la varianza de los residuos es más complejo al entrar en juego características tanto del modelo ajustado como del proceso puntual subyacente. Si se ajusta una realización de un proceso con intensidad $\lambda(u)$ mediante un modelo de Poisson con intensidad $\lambda_{\hat{\theta}}(u)$, la varianza de los residuos directos vendrá dada por:

$$\begin{aligned} \text{var}[R(B)] &= \int_B \lambda(u) du + \int_B \int_B \text{cov}(\lambda_{\hat{\theta}(x)}(u), \lambda_{\hat{\theta}(x)}(v)) dv du \\ &\quad - 2 \int_B \int_B E \left[\lambda_{\hat{\theta}(x \cup \{u\})}(v) - \lambda_{\hat{\theta}(x)}(v) \right] (\lambda(u)) dv du \end{aligned} \quad (7.28)$$

2.3. Gráficos de diagnóstico.

En esta sección se proponen gráficos de diagnóstico para detectar tanto tendencia como efectos de covariables. En el contexto de los modelos lineales generalizados (GLM) para estudiar el efecto de las covariables se representaban los residuos frente a índices, covariables incluidas en el modelo y otras variables explicativas no incluidas.

En el marco de los procesos puntuales espaciales, se proponen tres gráficos de diagnóstico para los residuos: el gráfico de marcas, el gráfico de residuos suavizados y los lurking-variable plots (gráficos de variables al acecho)

- **Gráfico de marcas:**

Para construir este gráfico se asocia a cada evento del proceso una marca que corresponde al valor del residuo en ese punto, $\hat{h}(x_i, x \setminus \{x_i\})$. A continuación se crea una imagen donde se representa la densidad, $\hat{h}(u, x) \hat{\lambda}(u; x)$, junto con círculos centrados en los eventos de radio el residuo en cada uno de ellos. Este gráfico permite identificar datos extremos. Sin embargo, el diagnóstico de los residuos está orientado al análisis de la suma de los residuos en subregiones, por lo que debemos buscar gráficos que tengan en cuenta estas sumas.

- **Residuos suavizados:**

Tomando una función núcleo $k(\cdot)$, el valor del residuo suavizado en cada punto es:

$$\begin{aligned} s(u) &= c(u) \int_W k(u-v) dR(v, \hat{h}, \hat{\theta}) \\ &= c(u) \left[\sum_{x_i \in x} k(u-x_i) \hat{h}(x_i, x \setminus \{x_i\}) - \int_W k(u-v) \hat{h}(u, x) \hat{\lambda}(u; x) du \right] \end{aligned} \quad (7.29)$$

donde $c(u)^{-1} = \int_A k(u-v) dv$ es el corrector del efecto frontera. Los residuos suavizados se representan en un gráfico de contorno. El análogo para la medida de innovación tiene media 0 y se espera que $s(u) \approx 0$ si el modelo es correcto.

- **Lurking-variable plot:**

En modelos lineales, cuando se sospecha que los datos dependen de una variable que no se ha incluido en el modelo, se representan los residuos frente a la variable en cuestión de modo que cualquier patrón en el gráfico indica que el modelo no es correcto y sugiere incluir la covariable.

Por analogía, en el marco de los procesos puntuales podemos representar los residuos frente a determinadas covariable o frente a las coordenadas cartesianas con

el fin de investigar la presencia de tendencia espacial que el modelo no refleja o para verificar si el modelo ajusta correctamente la tendencia del diseño observado.

Para una covariable espacial $Z(u)$ definida en cada punto de la región A , se evalúan los residuos en las siguientes subregiones:

$$W(z) = \{u \in A; Z(u) \leq z\}$$

Dando lugar a la función de residuos acumulada:

$$A(z) = n(x \cap W(z)) - \int_{W(z)} \hat{\lambda}(u, x) du \quad (7.30)$$

Nótese que $H(z) = n(x \cap W(z)) / n(x)$ es la función de distribución (CDF) de los valores de la covariable observados en los eventos del proceso. $A(z)$ es un ajuste de esta CDF y su media bajo el modelo es aproximadamente 0.

La lurking variable plot es el gráfico de $A(z)$ frente a la covariable Z . Para modelos sin interacciones, en este gráfico se añaden las envolturas superior e inferior correspondientes a los límites 2σ para un modelo inhomogéneo de Poisson. Suponiendo que se cumple el teorema central del límite, estos límites tienen la interpretación habitual de significación puntual.

3. TEST DE BONDAD DE AJUSTE PARA PROCESOS INHOMOGÉNEOS DE POISSON.

Guan (2008) propuso un test formal basado en los residuos de Baddeley y Turner (2005), descritos en la sección anterior, para contrastar la bondad de ajuste de un modelo inhomogéneo de Poisson. Pues, como se ha comentado anteriormente, la extensión de los test de Monte Carlo para la K-función bajo CSR al caso inhomogéneo tiene poca potencia para detectar la falta de ajuste, debido a que se está evaluando la K-función, basada en la estructura de segundo orden, para analizar la intensidad de primer orden. Por tanto se espera que un test basado en la intensidad tenga mayor potencia.

Consideremos un proceso puntual X observado en la región acotada W con intensidad de primer orden $\lambda(\cdot)$, y sea $\lambda_c(\cdot; \theta)$ una clase de modelos paramétricos para $\lambda(\cdot)$. El objetivo del test es contrastar la hipótesis nula $H_0: \lambda(\cdot) = \lambda_c(\cdot; \theta_0)$.

Sea $\hat{\theta}$ un estimador de θ . Para una forma sencilla $S \subset \mathbb{R}^2$, por ejemplo un círculo, sea $B(x, t)$ el conjunto de Borel de centro x y radio t con la misma forma que S y $N(x, t)$ el número de eventos observados en $B(x, t) \cap W$, entonces se define:

$$r_c(x, t, \hat{\theta}) = N(x, t) - \int_{B(x,t) \cap W} \lambda_c(u; \hat{\theta}) du \quad (7.31)$$

que es un caso particular de (7.16). Si el estimador de la intensidad es adecuado, el cuadrado de los residuos es un estimador insesgado de la varianza de $N(x, t)$. Por otra parte, al suponer que el proceso es de Poisson, $N(x, t)$ es un estimador insesgado de la misma cantidad. Esta coincidencia sugiere definir la siguiente medida de discrepancia:

$$D_c(t; \hat{\theta}) = \int_w \left(r_c(x, t, \hat{\theta})^2 - N(x, t) \right) dx \quad (7.32)$$

Guan (2008) demuestra que, cuando el ajuste es correcto, esta medida de discrepancia tiene distribución asintótica normal con media 0 y desviación típica $\sigma_c(\hat{\theta})$. Teniendo en cuenta esta propiedad se desarrolla el test de bondad de ajuste para el modelo ajustado $\lambda_c(\cdot; \hat{\theta})$. Fijado un radio t , se calcula el estadístico:

$$T(\hat{\theta}) = \frac{D_c(\hat{\theta}) - \text{sesgo}[D_c(\hat{\theta})]}{\sigma_c(\hat{\theta})} \quad (7.33)$$

que tiene distribución aproximada $N(0,1)$ bajo la hipótesis nula $H_0: \lambda(\cdot) = \lambda_c(\cdot; \theta_0)$. Así para un contraste con nivel de significación α , se rechaza la hipótesis nula, concluyendo falta de ajuste, cuando $T(\hat{\theta}) > Z_\alpha$, donde Z_α es el valor crítico de la derecha correspondiente a un nivel de significación α .

4. AJUSTE PARAMÉTRICO PARA LOS FUEGOS.

Se han ajustado distintos modelos paramétricos a los conjuntos de datos que nos ocupan. Para seleccionar el modelo más adecuado se ha estudiado su bondad de ajuste se ha utilizado el gráfico de diagnóstico implementado en *Spatstat*. Este gráfico presenta en una misma figura la imagen de los residuos marcados (esquina superior izquierda) los residuos suavizados (esquina inferior derecha) y las lurking variable plot respecto a los ejes de coordenadas (X: inferior izquierda, Y: superior derecha). En esta se han utilizado los residuos de Pearson y los límites 2σ se han calculado usando la varianza de las innovaciones que sobreestima la de los residuos. Una vez seleccionado el modelo se ha aplicado el contraste gráfico basado en la K-función inhomogénea, tomando como intensidad variable la propuesta para confirmar si el modelo es correcto o hay indicios de interacciones.

Tras comprobar que otros modelos no ajustaban bien los datos y coincidiendo con los resultados obtenidos por Genton et al (2008) y Yang (2007), hemos visto que los modelos más adecuados para describir la estructura de nuestros diseños son los que tienen intensidad con forma exponencial polinómica, esto es:

$$\lambda(x) = \exp\{\theta^T p_{x,m}\}$$

donde $p_{x,m}$ es un polinomio de orden m en \mathbb{R}^2 y θ es el vector de coeficientes asociado al polinomio. Hemos ajustado polinomios desde grado 1 hasta 6 a cada uno de los diseños analizados con el fin de seleccionar el mejor modelo posible en cada caso.

En primer lugar se presentan los gráficos del ajuste del modelo para los fuegos registrados en el intervalo 1991-2008 mediante polinomios de grado 2 y 4. Los gráficos de residuos respecto a las coordenadas indican que el modelo de orden 4 ajustaría mejor los datos, sin embargo en este caso tendremos que quedarnos con el modelo de menor grado, dado que con un polinomio de orden 4 el algoritmo GLM del método de máxima verosimilitud no converge. Por otra parte, si nos fijamos en el gráfico de los residuos suavizados, vemos que en la zona de mayor actividad el error del modelo es bastante grande, indicando que este modelo no muestra adecuadamente la estructura del diseño.

En el gráficos 2 se muestran los tests basados en la K-función inhomogénea, con estimación no paramétrica de la intensidad y con el estimador polinómico de orden 2. Ambas imágenes muestran indicios de agregación, siendo ligeramente mayor la distancia entre la K-empírica y la envoltura superior para el modelo paramétrico.

En los gráficos 3-5 se presentan los gráficos de diagnóstico correspondientes al mejor ajuste mediante la intensidad exponencial polinómica de grados 1-6. Para conatos e incendios el mejor ajuste se obtiene con un polinomio de grado 4, si bien en el primer caso las lurking plots muestran que el modelo no refleja correctamente el efecto de Y , pues los residuos acumulados a ciertas distancias están fuera de las envolturas. Para quemas, nos quedamos con un polinomio de grado 2, aunque al igual que en conatos hay puntos donde los residuos acumulados caen fuera de las envolturas. Respecto al K-test inhomogéneo correspondiente a los estimadores de la intensidad que acabamos de proponer, se mantienen los indicios de clustering vistos en estimación no paramétrica, siendo los radios de interacción 6 km para conatos, 4 km para incendios y 8 para quemas.

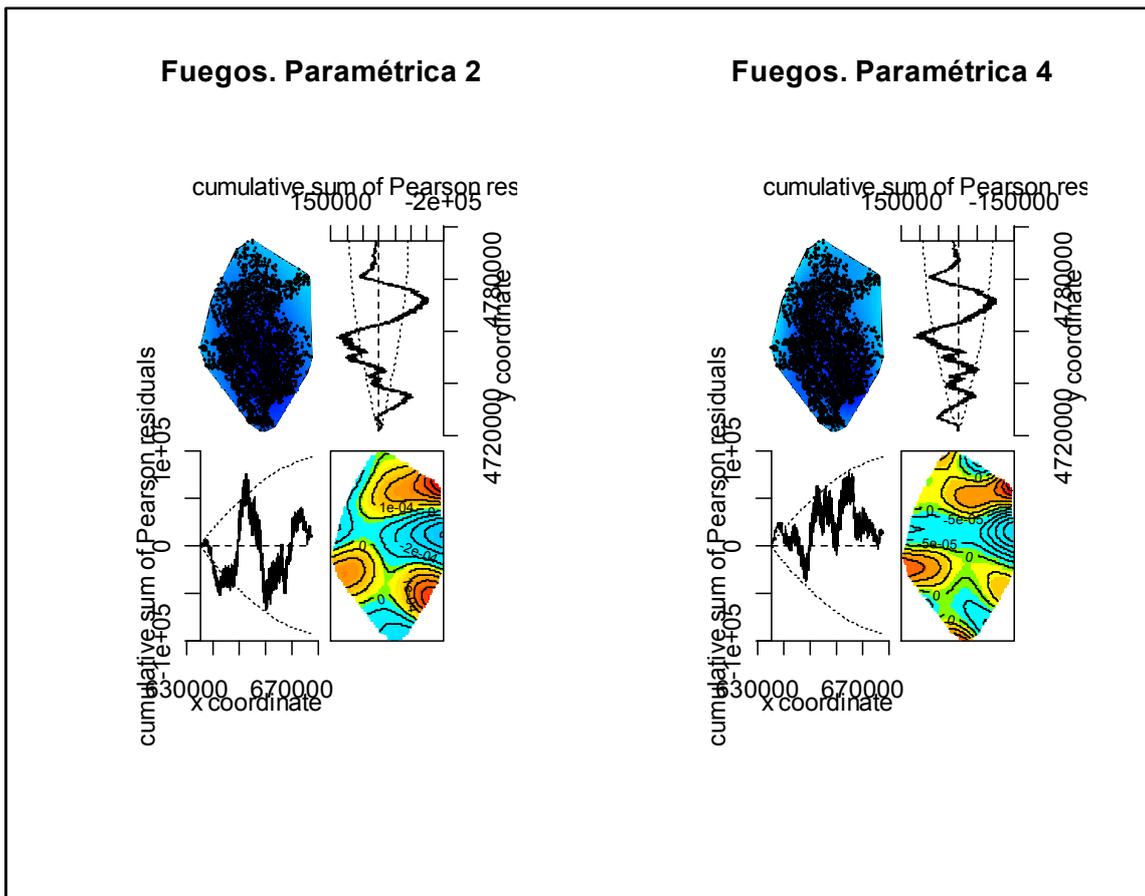


Gráfico 1: Grafico de diagnóstico para un modelo con exponencial polinómica de grados 2 y 4, correspondiente a los fuegos registrados en el período 1991-2008.

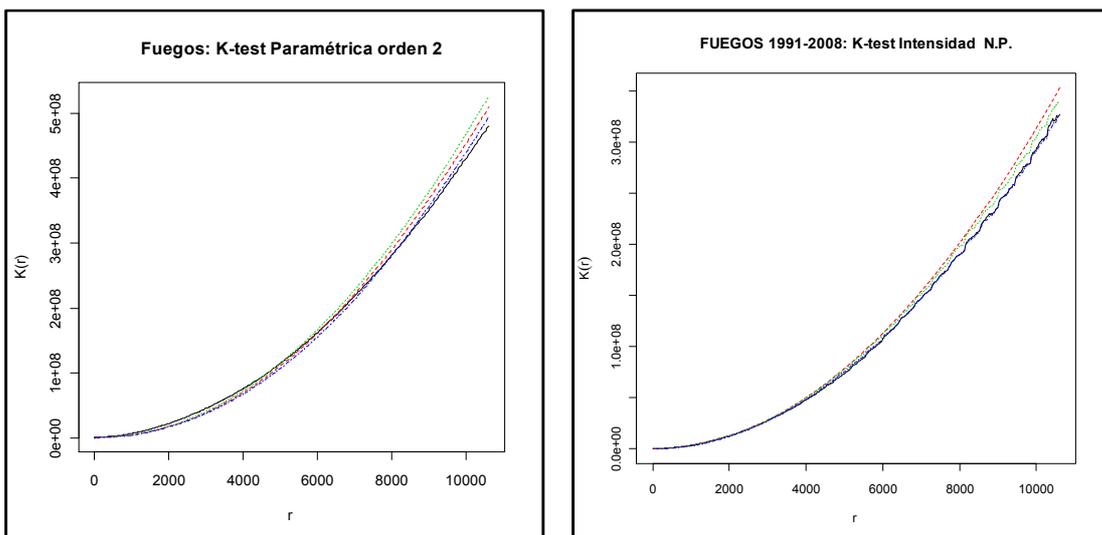


Gráfico 2: K test inhomogéneo, paramétrico y no paramétrico para fuegos. Distribución empírica (negro), L(t) (azul), U(t) (verde) distribución media de las simulaciones (rojo)

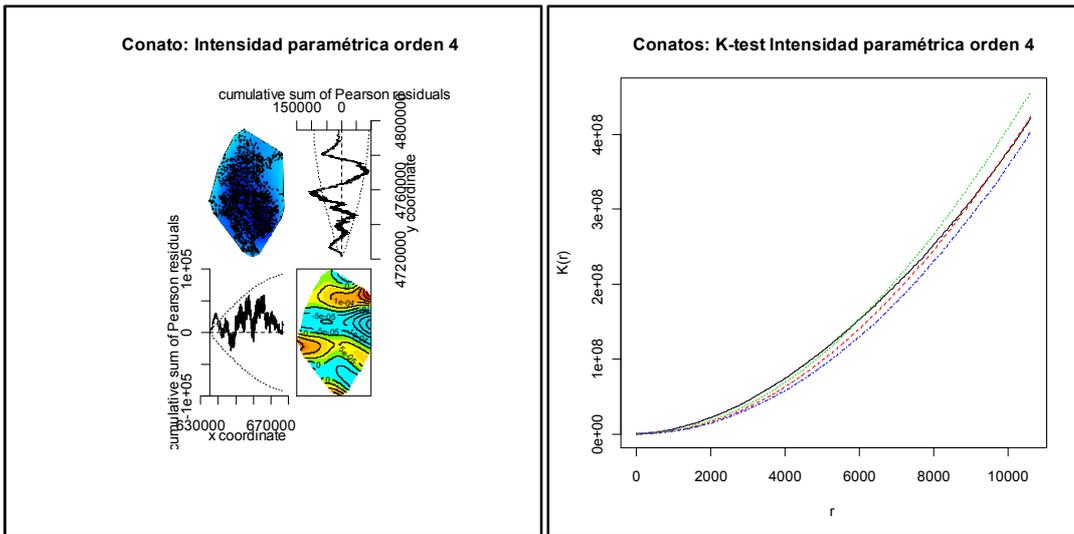


Gráfico 3: Grafico de diagnóstico y K-test inhomogéneo para conatos.

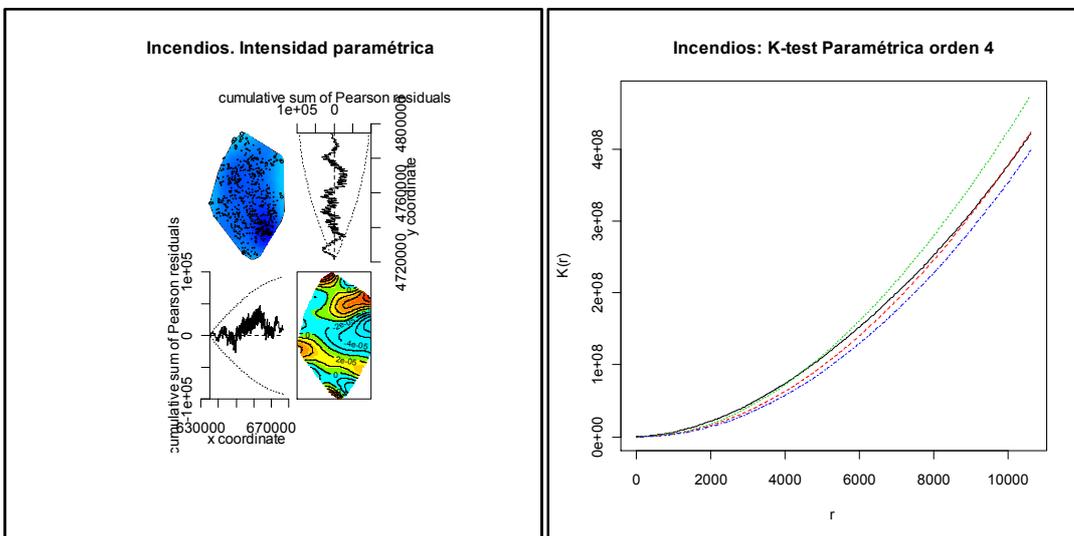


Gráfico 4: Grafico de diagnóstico y K-test inhomogéneo para incendios.

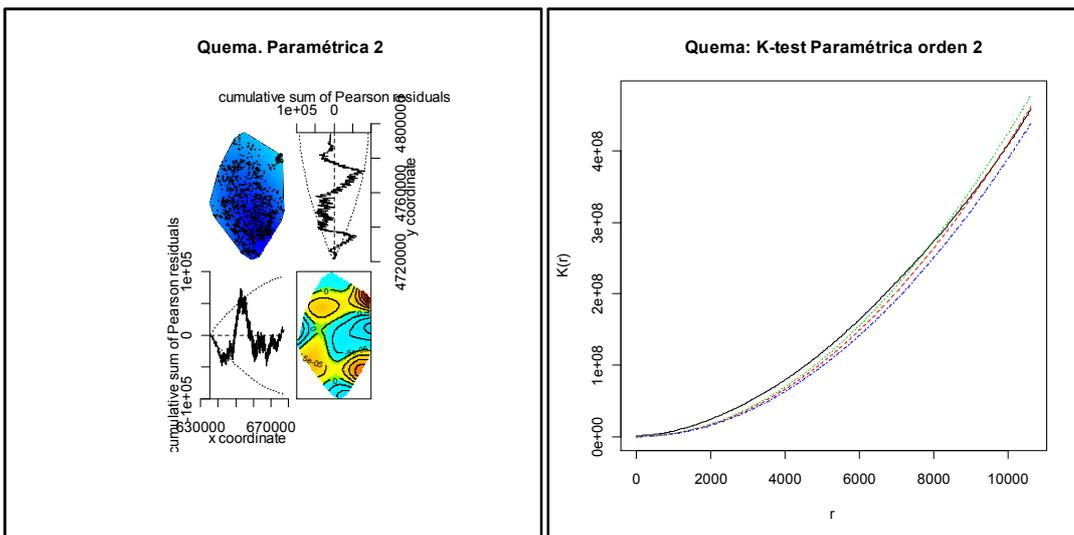


Gráfico 5: Grafico de diagnóstico y K-test inhomogéneo para quemas.

Respecto al análisis de los diseños correspondientes a los fuegos registrados en función de la causa que los produce, en los tres conjuntos de datos se ha visto que el mejor modelo correspondía a la intensidad exponencial con polinomio de grado 2. Aunque para los fuegos intencionados vemos el modelo no refleja completamente el efecto de las coordenadas sobre la intensidad del proceso. Observando la K-función se mantienen las sospechas de análisis agregado manifestadas en el análisis no paramétrico para fuegos intencionados, con radio de interacción de 4 km, y debidos a causas naturales, radio de 6 km, mientras que para los fuegos con causa desconocida, aunque la K-empírica está cerca de la envoltura superior se mantiene por debajo de ella, indicando que podríamos aceptar que no hay interacción entre este tipo de fuegos.

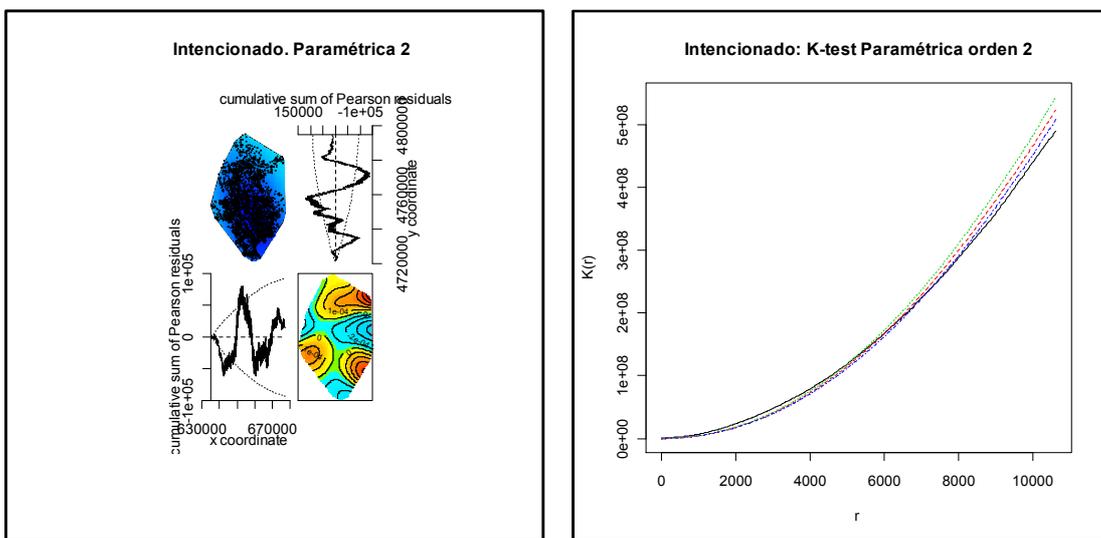


Gráfico 6: Grafico de diagnóstico y K-test inhomogéneo para fuegos intencionados.

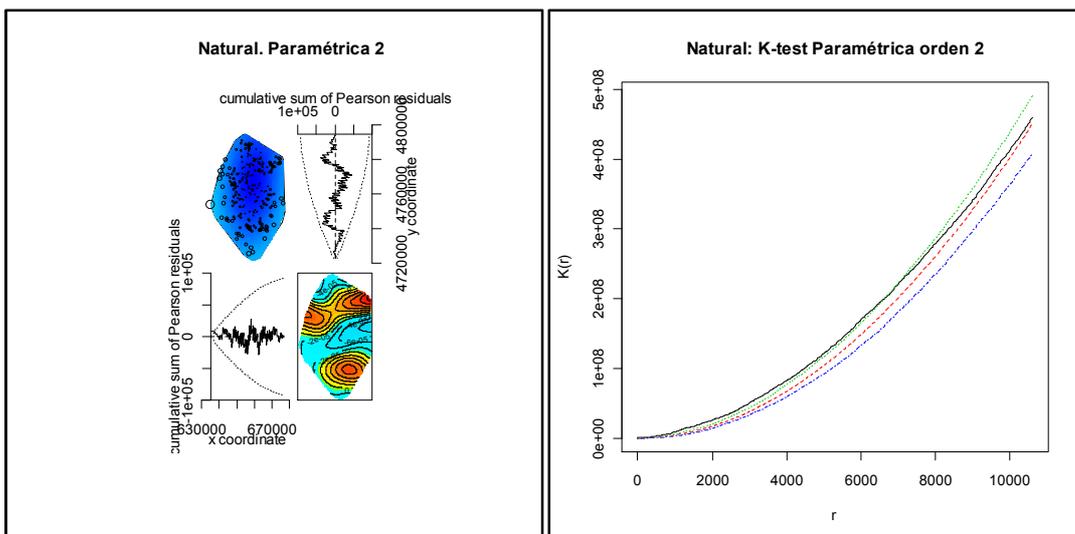


Gráfico 7: Grafico de diagnóstico y K-test inhomogéneo para fuegos debidos a causas naturales.

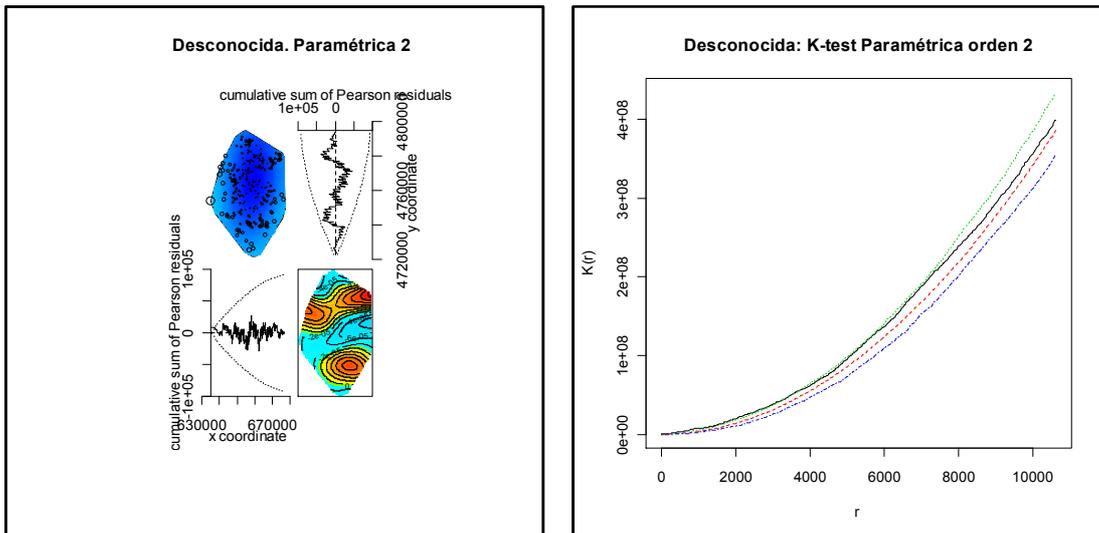


Gráfico 8: Gráfico de diagnóstico y K-test inhomogéneo para los fuegos con causa desconocida.

Por último se ha aplicado estimación paramétrica a los datos registrados en 1992. Lo gráfico 9 muestra que el modelo que mejor ajusta el diseño de todos los fuegos es el que tiene intensidad polinómica de orden 4, el K-test aplicado a este modelo presenta indicios de agregación con un radio de interacción de 3 km. Respecto a los fuegos clasificados por tipos, el modelo exponencial polinómico de orden 2 ajusta de modo razonable los tres conjuntos de datos, por su parte los K-tests correspondientes muestran indicios de diseño agregado con radios de interacción de 5 km para incendios, 6 km para conatos y 3 km para quemas.

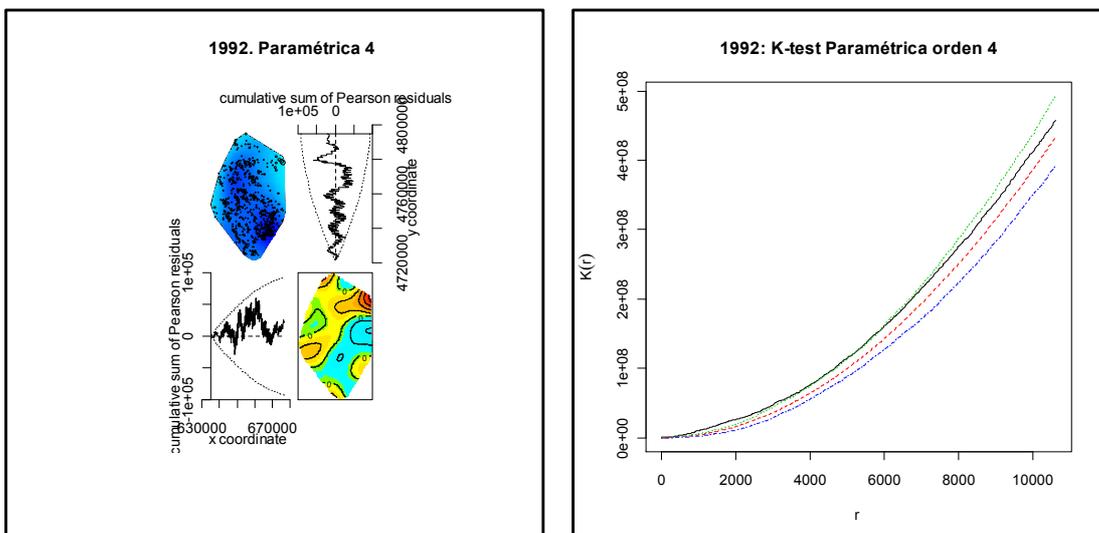


Gráfico 9: Gráfico de diagnóstico y K-test inhomogéneo para los fuegos registrados en 1992.

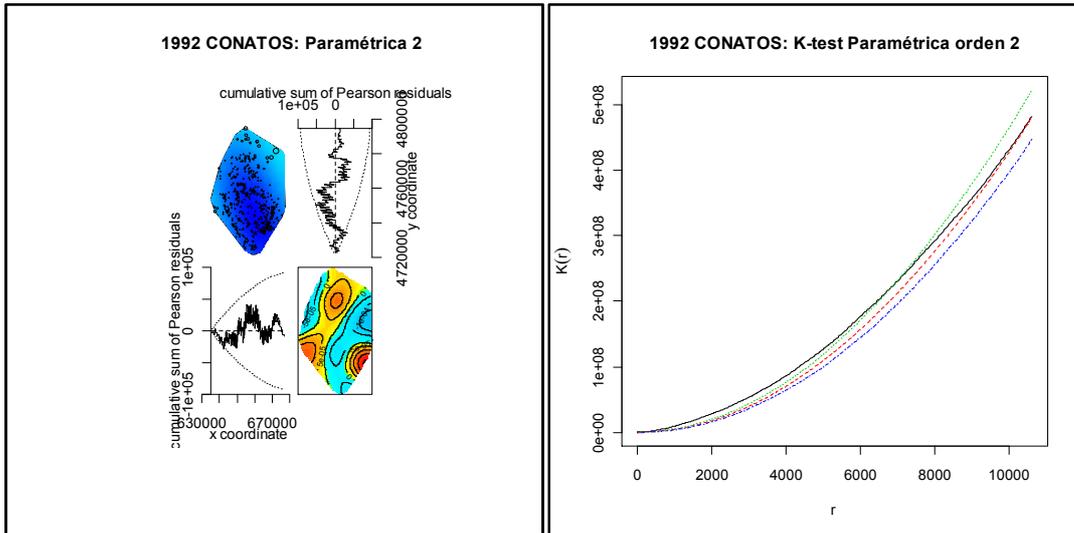


Gráfico 10: Gráfico de diagnóstico y K-test inhomogéneo para los conatos registrados en 1992.

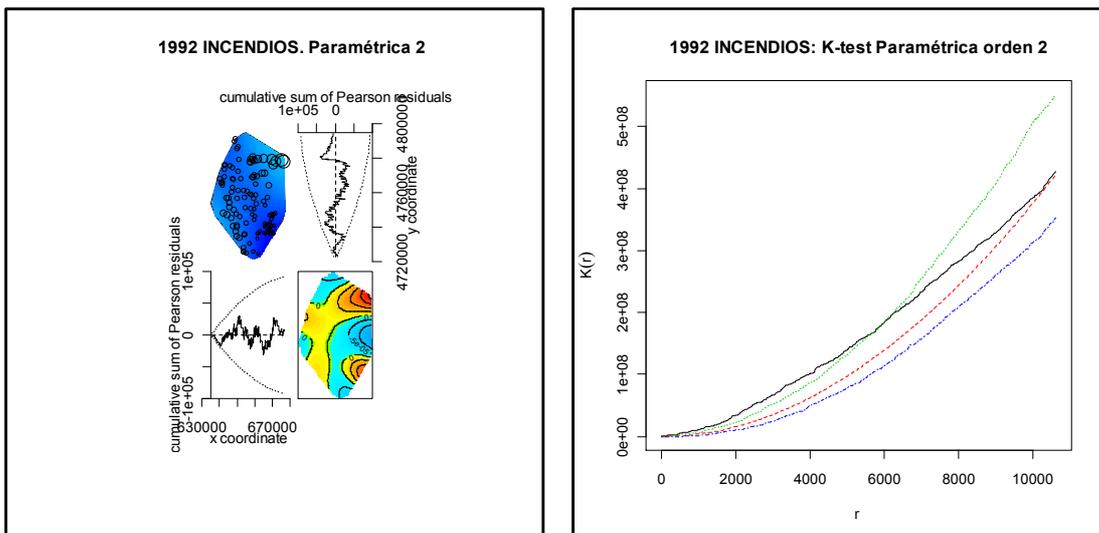


Gráfico 11: Gráfico de diagnóstico y K-test inhomogéneo para los incendios registrados en 1992.

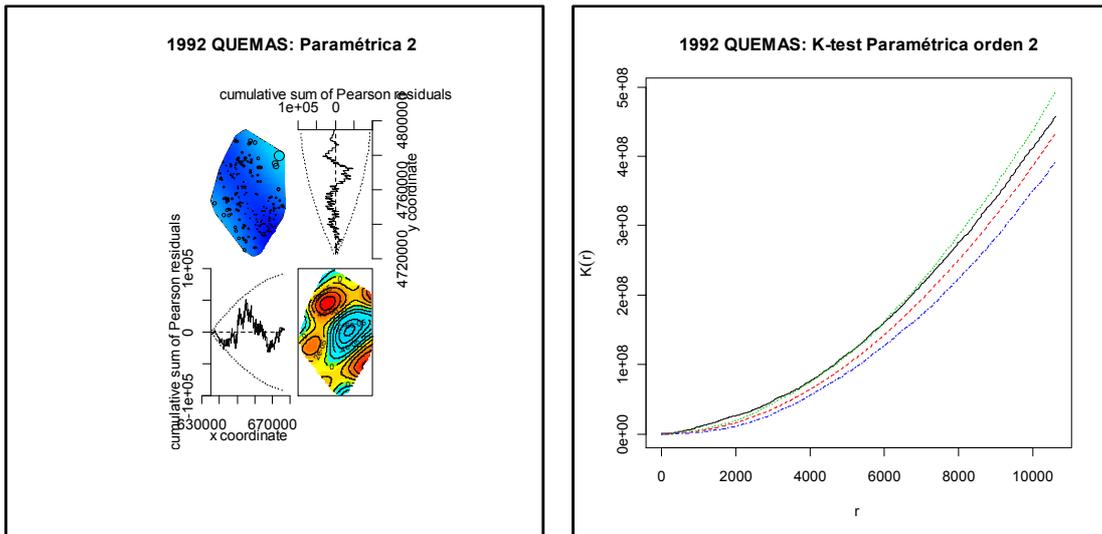


Gráfico 12: Gráfico de diagnóstico y K-test inhomogéneo para las quemaduras registradas en 1992.

En los gráficos 13-14 se presentan las imágenes correspondientes a los estimadores paramétricos de la intensidad, tanto para el total de fuegos registrados en 1992 como para los diseños correspondientes a su clasificación en tipos. En estos gráficos vemos que los estimadores paramétricos de la intensidad, al igual que los no paramétricos, identifican el sureste del distrito, Suarna y Pedrafita do Cebreiro, como la zona más conflictiva.

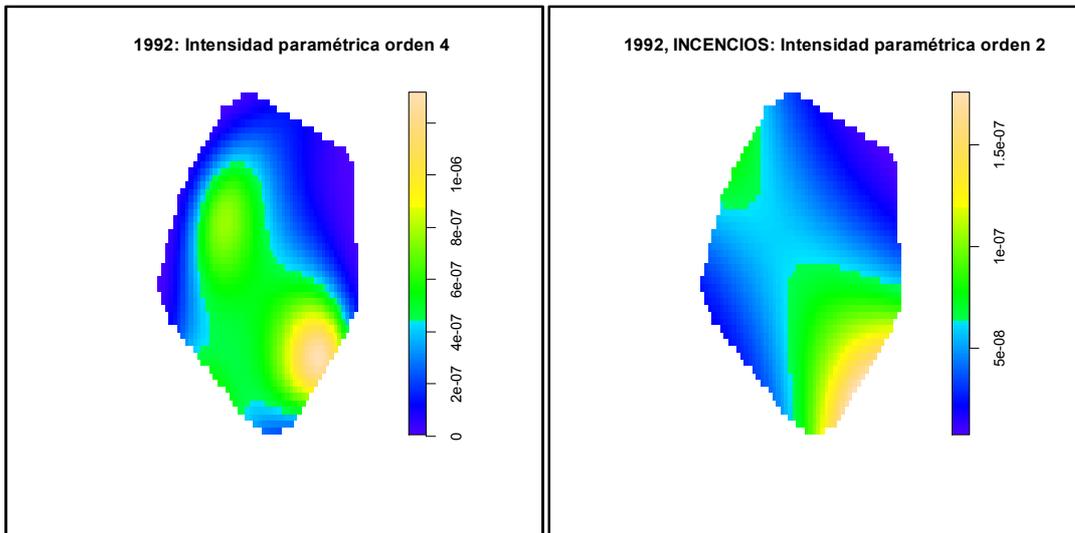


Gráfico 13: Estimación paramétrica de la intensidad para fuegos, incendios, conatos y quemaduras registradas en 1992.

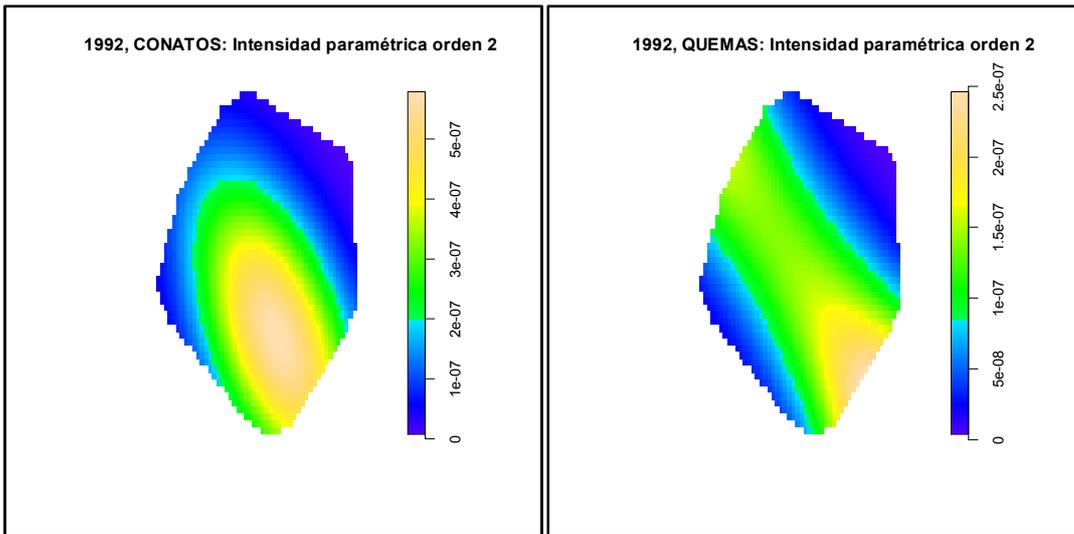


Gráfico 14: Estimación paramétrica de la intensidad para conatos y quemas registrados en 1992.

CONCLUSIÓN

1. CONCLUSIONES DEL ANÁLISIS DE LOS INCENDIOS FORESTALES.

Como resultado de la aplicación de las distintas técnicas de análisis exploratorio e inferencia en procesos puntuales espaciales a los incendios forestales registrados en el distrito Fonsagrada- Ancares durante el periodo 1991-2008 se han extraído las siguientes conclusiones.

- Los contrastes de CSR han descartado la hipótesis de distribución aleatoria de los incendios. Proporcionando además una fuerte evidencia de diseños agregados.
- Si se aproxima la intensidad variable de los procesos mediante un estimador no paramétrico, las imágenes obtenidas identifican correctamente las zonas más conflictivas del área. Si se utilizan estas intensidades no paramétricas para estimar la K-función inhomogénea y se aplica un test de independencia análogo al K-test de CSR, se siguen teniendo indicios de agregación para la mayoría de los diseños. Sin embargo, comparando estos gráficos con los de CSR vemos que bajo homogeneidad se sobreestimaba el clustering al identificar como interacción lo que en realidad es tendencia espacial.
- Al suponer que el proceso es inhomogéneo y ajustar un modelo con intensidad exponencial polinómica, de grado 2 o 4 según el conjunto de datos, se obtienen estimaciones de la intensidad similares a las no paramétricas y el test basado en la K-función inhomogénea vuelve a mostrar indicios de clustering para la mayoría de los diseños analizados.
- Al estudiar la interacción entre tipos de fuegos, el test basado en la K-función bivalente indica atracción entre los pares incendios-conatos, incendios-quemas y conatos-quemas. Este confirma la sospecha de que los pirómanos intentan provocar reiteradamente un incendio en la misma zona hasta que lo consiguen.
- El análisis de interacción entre años indica que no hay evidencias de dependencia entre las localizaciones de los incendios entre un año y los siguientes.

Los comentarios anteriores sugieren que el próximo paso a dar en este estudio sería ajustar un modelo con interacciones en el que se tenga en cuenta la atracción detectada entre los distintos tipos de fuegos.

2. FUTURAS LINEAS DE INVESTIGACIÓN.

En este trabajo se ha visto la utilidad de los procesos puntuales espaciales en una aplicación tan importante como es el análisis de incendios forestales. Sin embargo lo que ha presentado no es más que una primera aproximación. Pues contamos con información y herramientas metodológica suficientes para seguir estudiando el comportamiento de los fuegos y tratar de construir un modelo que nos ayude a predecir y, por tanto, evitar la localización de los incendios forestales.

El SIG permite vincular al diseño espacial variables físicas tales como la pendiente, la orientación, tipo de propiedad, distancia a núcleos urbanos y carreteras, y modelos de combustible existentes en el punto de inicio de cada fuego. La importancia de las variables socioeconómicas como, distancias a núcleos urbanos y carreteras o tipo de propiedad, radica en el hecho de que los pirómanos van a provocar fuegos zonas de fácil acceso y, por otra parte la cercanía a núcleos de población o a carreteras hacen que la intervención pueda ser más rápida evitando que las quemaduras o conatos pasen a convertirse en incendios. Variables como orientación y tipo de combustible afectan a la velocidad de propagación del fuego, por tanto, también pueden ayudar a predecir la probabilidad de que un conato o una quema alcance la condición de incendio.

En este trabajo solo se han analizado los datos correspondientes al distrito Fonsagrada- Ancares. Para estudios futuros sería de gran interés contar con información de toda Galicia, esto permitirá añadir la componente temporal al modelo. Así, al contar con una superficie mayor y un número mayor de datos, se puede introducir la fecha como variable y hacer un análisis espacio-temporal más completo, teniendo en cuenta aspecto como la interacción entre conatos registrados un día e incendios en días posteriores.

Desde el punto de vista metodológico han quedado abiertas las siguientes cuestiones:

- Aplicación del contraste de bondad de ajuste propuesto por Guan (2008) a los modelos ajustados para los fuegos.
- Desarrollo de un método bootstrap suavizado que permita aproximar el MISE de un estimador de la intensidad y obtener la expresión asintótica del MISE*.

BIBLIOGRAFÍA

1. PROCESOS PUNTUALES ESPACIALES

1. Allen, M.P. and Tildesley, D.J. (1987). *Computer Simulation of Liquids* Oxford University Press, Oxford..
2. Baddeley, A.J., Møller, J., and Waagepetersen, R. (2000). Non and semi- parametric estimation of interactions in inhomogeneous point patterns. *Statistica Neerlandica*, **54**, 329-50.
3. Baddeley, A. and Turner, R. (2000). Practical maximum pseudolikelihood for spatial point patterns (with discussion). *Australian and New Zealand Journal of Statistics*, **42**, 283-322.
4. Baddeley,A. Møller,J. and Waagepetersen, R.(2000) Non- and semiparametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, **54(3)**:329-350.
5. Baddeley,A. Turner,R. Møller,R. And Hazelton, M. (2005). Residual analysis for spatial point processes (with discussion). *Journal of the Royal Statistical Society, series B*, 67(5):617-666,.
6. Baddeley,A. and Turner,R. (2005) Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software* 12: 6 1-42.www.jstatsoft.org.
7. Baddeley,A. Analysing Spatial Point Patterns in R. Workshop Notes. Published online by CSIRO, Australia, 10 february 2008.
8. Bartlett, M.S. (1964). Spectral analysis of two-dimensional point processes. *Biometrika*, **51**, 299-311.
9. Berman, M. and Diggle, P. (1989). Estimating weighted integrals of the second order intensity of a spatial point process. *Journal of the Royal Statistical Society, B* **51**, 81-92.
10. Berthelsen , K.K. and Møller, J: (2002). A primer on perfect simulation for spatial point processes. *Bulletin of the Brazilian Mathematicla Society*, **33**, 351-67.
11. Besag, J. (1977). Contribution on the discussion of Dr. Ripley's paper. *Journal of the Royal Statistical Society, B* **39**, 193-5.
12. Besag, J. (1978). Some methods of statistical analysis for spatial data. *Bulletin of the International Statistical Institute*, **47**, 77-92.
13. Cox, D.R. (1977). The role of significance tests. *Scandinavian Journal of Statistics*, **4**, 49-70.
14. Cowling,A., Hall,P. and Phillips, M.J.(1996). Bootstrap confidence regions for the intensity of a Poisson point process, *Journal of the American Statistical Association*, vol. 91, pp. 1516–1524,
15. Cressie, N.A.C. (1993). *Statistics for Spatial Data*, second edn, Wiley, New York.
16. Diggle, P.J (1985b). A kernel method for smoothing point process data. *Applied Statistics*, **34**, 138-47.
17. Diggle, P.J (2003). *Statistical Analysis of Spatial Point Patterns*. Oxford University Press.
18. Diggle, P.J., Fiksel, T., Grabarnik, P. Ogata, Y., Stoyan, D. and Tanemura, M. (1994). On parameter estimation for pairwise interaction point processes. *International Statistitcal Review*, **62**, 99-117.

19. Diggle, P.J, Gates, D.J. and Stibbard, A. (1987). A non parametric estimator for pairwise interaction point processes. *Biometrika*, **74**, 763-70.
20. Diggle, P.J. and Gratton, R.J. (1984). Monte Carlo methods of inference for implicit statistical models (with discussion). *Journal of the Royal Statistical Society*, B **46**, 193-227.
21. Diggle, P.J. and Matèrn (1981). On sampling designs for the estimation of point- event nearest neighbour distributions. *Scandinavian Journal of Statistics*, **7**, 80-4.
22. Diggle, P.J. and Milne (1983a). Negative binomial quadrat counts and point processes. *Scandinavian Journal of Statistics*, **10**, 257-67.
23. Diggle, P.J. and Milne (1983b). Bivariate Cox processes: some models for bivariate spatial point patterns. *Journal of the Royal Statistical Society*, B **45**, 11-21.
24. Efron, B. and Tibshirani, R.J. (1993). *An introduction to the Bootstrap*. London: Chapman & Hall.
25. Fernández, R., Ferrari, P.A. and García, N.L: (2002). Perfect simulation for interacting point processes , loss networks and Ising models. *Stochastic Processes and their Applications*, **102**, 63-88.
26. Fiksel, T. (1984). Simple spatial-temporal models for sequences of geological events. *Elektronische Informationsverarbeitung und kybernetick*, **20**, 480-87.
27. Foss, S.G. and Tweedie, R.L. (1998). Perfect simulation and backward coupling. *Stochastic Models*, **14**, 187-2003.
28. Genton, M. G., Butry, D. T., Gumpertz, M. L., and Prestemon, J. P. (2006), "Spatio-temporal analysis of wildfire ignitions in the St Johns River Water Management District, Florida," *International Journal of Wildland Fire*, **15**, 87-97.
29. Geyer, C.J. and Moller, J. (1994). Simulation procedures and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics*, **21**, 359-73
30. Geyer, C.J. and Thompson, E.A. (1992). Constrained Monte Carlo maximum likelihood for dependet data (with discussion). *Journal of the Royal Statistical Society*, B **54**, 657-99.
31. Greig-Smith, P. (1952). The use of random and contiguous quadrats in the study of the structure of plant communities. *Annals of Botany*, **16**, 293-316.
32. Guan, Y. A Goodness-of-fit Test for Inhomogeneous Spatial Poisson Processes. *Biometrika*, **95**,4, pp: 831-845
33. Häggström, O. van Lieshout, M.N.M. and Moller, J. (1999). Characterization results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes. *Bernoulli*, **5**: 641-659.
34. Heikkinen, J. and Penttinen, A. (1999). Bayesian smoothing in the estimation of the pair potential function of Gibbs point processes. *Bernoulli*, **5**, 1119-36.
35. Hering, A. S., Bell, C. L., and Genton, M. G. (2009), "Modeling spatio-temporal wildfire ignition point patterns," *Environmental and Ecological Statistics*, Special Issue on Statistics for Wildfire Processes, **16**, 225-250.
36. Hope, A.C.A. (1968), A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society*, B **30**, 582-98.
37. Hsuan, F. (1979). Generating uniform polygonal random pairs. *Applied Statistics*, **28**, 170-2.
38. Kelly, F.P. and Ripley, B.D. (1976). A note on Strauss model for clustering. *Biometrika*, **63**, 357-60.

39. Kendall, W.S. (1998). Perfect simulation for the area-interaction point process, in L. Accardi & C. Heyde (eds). *Probability towards 2000*, Springer Lecture Notes in Statistics 126. Springer Verlag, New York, pp. 218-234.
40. Kendall, W.S. and Moller, J. (2000). Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes. *Advances in Applied Probability*, **32**, 844-865.
41. Kingman, J.F.C: (1977). Remarks on the spatial distribution of a reproducing population. *Journal of Applied Probability*, **14**, 577-83.
42. Lewis, P.A.W. and Shedler, G.S. (1979). Simulation of non- homogeneous Poisson processes by thinning. *Naval Research Logistics Quarterly*, **26**, 403-13.
43. Lotwick, H.W. (1981), PhD thesis, University of Bath.
44. Lotwick, H.W. and Silverman, B.W. (1982). Methods for analysing spatial processes of several types of points. *Journal of the Royal Statistical Society, B*, **44**, 406-13.
45. Matèrn, B. (1960). *Spatial Variation*. Modellanden fran statens skogsforsningsinstitut, Vol. 49(5). Stockholm: statens skogsforsningsinstitut.
46. Matèrn, B. (1971). Doubly stochastic Poisson processes in the plane. In *Statistical Ecology*. Vol. 1, ed. G.P. Patil, E.C. Pielou and W.E. Waters, 195-213. University Park: Pennsylvania State University Press.
47. J. MATEU, P. GREGORI, P. JUAN, F. SAURA & A. CALLE. *Modelos espacio-temporales para la evaluación y caracterización de índices de riesgo de incendios forestales en la Comunidad Valenciana* (2006). **Davalos-Fletcher**. ISBN: 84-689-7396-3. Depósito Legal: CS-85-2006.
48. Mead, R. (1974). A test for spatial pattern at several scales using data from a grid of contiguous quadrats. *Biometrics*, **30**, 296-307.
49. Meyn, S.P. and Tweedie, R.L. (1993). *Markov Chains and Stochastic Stability*. Springer-Verlag, London.
50. Moller, J. and Sorensen, M. (1994). Parametric models of birh-and-death processes with a view to modelling linear dune fields. *Scandinavian Journal of Statistics*, **21**, 1-19.
51. Moller, J. and Waagepetersen, R.P. (2003). Statistical inference and simulation for spatial point processes *Monographs on Statistics and Applied Probability*, 100. Chapman & Hall.
52. Nelder J.A. and Mead, R. (1965). A simplex method for for function minimization. *Computer Journal*, **7**, 308-13.
53. Penttinen, A. (1984). *Modelling Interactions in Spatial Point Patterns: Parameter Estimation by the Maximum Likelihood Method*. Jyväskylä Studies in Computer Science. Economics and Statistics, 7. University of Jyväskylä.
54. Preston, C.J: (1977). Spatial birth-and-death processes. *Bulletin of the International Statistical Institute*, **46**, 371-91.
55. Propp, J.G. and Wilson, D.B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, **9**, 223-52.
56. Ripley, B.D. (1976). The second order analysis of stationary point processes. *Journal of Applied Probability*, **13**, 255-66.
57. Ripley, B.D. (1977). Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society, B*, **39**, 172-212.
58. Ripley, B.D. (1979a). Tests of "randomness " for spatial point patterns. *Journal of the Royal Statistical Society, B*, **41**, 368-74.

59. Ripley, B.D. (1979b). Simulating Spatial patterns: dependent samples from a multivariate density. *Applied Statistics*, **28**, 109-12.
60. Ripley, B.D. (1988). *Statistical Inference for Spatial Processes*. Cambridge: Cambridge University Press.
61. Ripley, B.D. and Kelly, F.P. (1977). Markov point processes. *Journal of the London Mathematical Society*, **15**, 188-92.
62. Stoyan, D., Kendall, W.S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications*, second edn, Wiley, Chichester.
63. Yang et al., 2007 J. Yang, H.S. He, S.R. Shifley and E.J. Gustafson, Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands, *Forest Science* 53 (2007), pp. 1–15.

2. INCENDIOS FORESTALES.

64. Andrews P.L., 1986. BEHAVE: Fire behavior prediction and fuel modeling system-BURN subsystem, Part 1, Gen. Tech. Rep. INT-194. Ogden, UT: US Department of Agriculture, Forest Service, Intermountain Research Station, 130 p.
65. Bachmann, A. and Allgower, B., 2000. The need for a consistent wildfire risk terminology. In: Neuenschwander, L., Ryan, K., Golberg, G. (Eds.), *Crossing the Millennium: Integrating Spatial Technologies and Ecological Principles for a New Age in Fire Management*. The University of Idaho and the International Association of Wildland Fire, Moscow, ID, pp: 67–77.
66. Breslow NE and Clayton DG. 1993. Approximate inference in generalized linear models. *Journal of American Statistical Association* 88: 9–25. Brillinger, D.R., 2003. Three environmental probabilistic risk problems. *Statistical Science*, 18: 412-421.
67. Brillinger, David R., Haiganoush K. Preisler and John W. Benoit, 2003. Risk assessment: a forest fire example. *Science and Statistics. Lecture Notes in Statistics* 40: 177-196.
68. Brillinger, D.R., Preisler, H.K. and Naderi, H.M., 2004. Wildfire chances and probabilistic risk assessment. *Proceedings of Accuracy 2004 and Ties 2004*, Portland, Maine.
69. Brillinger, D. R., Preisler, H. K. and Benoit, J. W., 2006. Probabilistic risk assessment for wildfires. *Environmetrics*, 17: 623-
70. Chas, M. 2007. Forest fires in Galicia (Spain): Threats and challenges for the future. *Journal of Forest Economics* 13: 1–5.
71. Chou Y.H., Minnich R.A., Salazar L.A., Power J.D. and Dezzani, R.J. 1990. Spatial autocorrelation of wildfire distribution in the Idyllwild Quadrangle, San Jacinto mountains, California, USA. *Photogrametric Engineering and Remote Sensing* 56: 1507-1513.
72. Chou Y.H., Minnich R.A. and Chase R.A. 1993. Mapping probability of fire occurrence in San Jacinto mountains, California, USA. *Environmental Management* 17: 129-140.
73. Chuvieco E. and Salas, J. 1996. Mapping the spatial distribution of forest fire danger using GIS. *International Journal of Geographical information Systems* 10:333-345.
74. Cox D.R. and Snell E.J. 1989. *Analysis of Binary Data*, 2nd ed., Chapman & Hall, London, 236.

75. Cumming, S.G., 2001. A parametric model of the fire-size distribution. *Canadian Journal of Forest Research*, 31, 1297-1303
76. Cunningham A.A. and Martell, D.L. 1972. a stochastic models for forest fires. *Ecological Modelling* 3:309-313.
77. Deeming, J., Lancaster, J., Fosberg, M., Furman, R. and Shroeder, M.,1972. National fire-danger rating system. Research Paper RM- 84, USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, 165.
78. DGCN, Los incendios forestales en España. Decenio 1996–2005, Dirección General de Conservación de la Naturaleza, MMA, Madrid, 2006.
79. Diaz-Delgado,R., Lloret,F. and Pons,X., 2004. Spatial patterns of fire occurrence in Catalonia, NE, Spain. *Landscape Ecology*, 19: 731- 745.
80. Dickson,B.G, John W. Prather, J.W. Yaguang X, Haydee M. Hampton, H.M, Aumack, E.N., and Sisk, T.D. 2006. Mapping the probability of large fire occurrence in northern Arizona, USA. *Landscape Ecology* 21:747–761.
81. Finney M.A. 1994 Modeling the spread and behavior of prescribed natural fires, Proc. 12th Conf. Fire and Forest Meteorology, 138–143.
82. Floyd, M.L., Hanna, D.D. and Romme, W.H., 2004. Historical and recent fire regimes in Piñon-Juniper woodlands on Mesa Verde, Colorado, USA. *Forest Ecology and Management*, 198: 269-289.
83. González,J.R. and Pukkala,T., 2007. Characterization of forest fires in Catalonia (north-east Spain). *European Journal of Forest Research*, 126: 421-429.
84. González, J.R., Kolehmainen, O. and Pukkala, T. 2007 Using expert Knowledge to model forest stands vulnerability to fire. *Computers and Electronics in Agriculture* 55: 1007-114.
85. Green, P.J. and Silverman, B.W. 1994. *Nonparametric Regression and Generalized Linear Models*. London: Chapman and Hall.
86. Hardy, C.C., Menakis, J., Long, D., Brown, J. and Bunnell, D. 1998. Mapping historic fire regimes for the western United States: integrating remote sensing and biophysical data. In: *Proceedings of the Seventh Biennial Forest Service Remote Sensing Applications Conference*, Nassau Bay, TX, 6–9 April 1998, American Society for Photogrammetric Engineering and Remote Sensing, Bethesda, MD. 288– 300.
87. Hardy, C.C., Schmidt, K., Menakis, J. and Sampson, R. 2001. Spatial data for national fire planning and fuel management. *International Journal of Wildland Fire* 10: 353–372.
88. Hardy, C.C. 2005. Wildland fire hazard and risk: Problems, definitions and context. *Forest Ecology and Management* 211: 73-82.
89. Heinselman, M. L. 1981. Fire and succession in the conifer forests of North America. Pages 374 -406 in D. C. West, H. H. Shugart, and D. B. Botkin, editors. *Forest succession: concepts and application*. Springer-Verlag, New York, New York, USA.
90. Hessburg, P.F., Smith,B.G., Salter, R.B., Ottmar, R.D. and Alvarado, E., 2000. Recent changes (1930s-1990s) in spatial patterns of interior northwest forests, USA. *Forest Ecology and Management*, 136:53-83.
91. Hosmer D.W. and Lemeshow S. 2000. *Applied logistic regression*, 2nd ed., Wiley, Series in Probability and Statistics, New York.
92. Hoxie, G.L., 1910. How fire helps forestry. *Sunset* 34: 145–151.
93. Ketch, J.J. and Byram G.m. 1968. A drought index for forest fire control. USDA Forest Service Research Paper SE-38.

94. Kushla, J.D. and Ripple, W.J. 1997. The role of terrain in a fire mosaic of a temperate coniferous forest. *Forest Ecology and Management*, 136: 53-83.
95. Johnson E.A. and Van Wagner C.E. 1985. The Theory and use of two fire history models. *Canadian Journal of Forest Research* 15: 214-220.
96. Johnson, N.L., Kotz, S., and Balakrishnan, N. 1994. Continuous univariate distributions. 2nd ed. Vol. 1. JohnWiley & Sons, Inc., New York.
97. Lloret, F., Calvo, E., Pons, X. and Díaz-Delgado, R. 2002 Wildfires and landscape patterns in the Eastern Iberian Peninsula. *Landscape Ecology* 17: 745-759.
98. Loehle, C. 2004. Applying landscape principles to FIRE hazard reduction. *Forest Ecology and Management*, 198: 261-267.
99. Malamud, B.D., Morein, G. and Turcotte, D.L. 1998. Forest fires: an example of self-organized criticality. *Science*, 281: 1840–1842.
100. Mandallaz, D. and Ye, R. 1997. Prediction of forest fires with Poisson models. *Canadian Journal of Forest Research*, 27: 1685-1694.
101. Marey Pérez, M.F., Rodríguez Vicente, V. and Crecente Maseda, R. (2006). 'Using GIS to measure changes in the temporal and spatial dynamics of forestland: experiences from north-west Spain'. *Forestry* Vol. 79: 409-423.
102. Marey Pérez, M.F. and Rodríguez Vicente, V. (2008). 'Forest transition in Northern Spain: Local responses on large-scale programmes of field-afforestation'. *Land Use Policy*, Article in Press.
103. Martell D.L., Otukol S. and Stocks B.J. 1987. A logistic model for predicting daily people-caused forest fire occurrence in Ontario. *Canadian Journal of Forest Research* 17: 394-401.
104. Martell D.L., Bevilacqua E. and Stocks B.J. 1989. Modelling seasonal variation in daily people-caused forest fire occurrence. *Canadian Journal of Forest Research* 19: 1555-1563.
105. Mermoz, M., Kitzberger, T. and Veblen, T.T. 2005. Landscape influences on occurrence and spread of wildfires in Patagonian forests and shrublands. *Ecology* 86: 2705-2715.
106. Merrill D.F. and Alexander M.E. 1987. Glossary of forest fire management terms. 4^a Ed. Natural Resources Council Canadian Community Forest Fires Management. Ottawa. Canada.
107. Minnich, R. A. 1983. Fire mosaics in southern California and northern Baja California. *Science* 219:1287–1294.
108. Minnich, R.A. and Bahre, C.J. 1995. Wildland fire and chaparral succession along the California-Baja California Boundary. *International Journal of Wildland Fire* 5:13-24.
109. Moreira, F. Rego, F.C. and Ferreira, P.G. 2001 Temporal (1958-1995) pattern of change in a cultural landscape of northwestern Portugal: implications for fire occurrence. *Landscape Ecology* 16: 557-567
110. Morgan, P., Hardy, C., Swetnum, T., Rollins, M. and Long, D., 2001. Mapping fire regimes across time and space: understanding coarse and fine-scale fire patterns. *International Journal of Wildland Fire* 10:329-342
111. Moritz, M.A. 2003. Spatiotemporal analysis of controls on shrubland and fire regimes: age dependency and fire hazard. *Ecology* 84: 351-361.
112. Ministry of Forests (MOF), 1997. Glossary of Forest Terms. Ministry of Forests, Province of British Columbia, Canada. Available on the web: <http://www.for.gov.bc.ca/pab/publctns/glossary/glossary.htm>.

113. National Academy of Sciences (NAS), 1983. Risk Assessment in the Federal Government: Managing the Process. National Academy Press, Washington, DC, 192.
114. National Research Council (NRC), 1989. Improving Risk Communication. National Academy Press, Washington, DC.
115. National Wildfire Coordinating Group (NWCG), 2003. Glossary of Wildland Fire Terminology. Available on the web: <http://www.nwcg.gov/teams/pmo/products/wfglossary/content.htm>.
116. Nunes M.C.S. Vasconcelos, M.J. Pereira, J.M.C., Dasgupta, N., Alldredge, R.J. and Rego, F.C. 2005 Land cover type and fire in Portugal: do fires burn land cover selectively?. *Landscape Ecology* 20: 661-673.
117. Pereira, J.M.C., João M.B. Carreiras and Maria J. P. de Vasconcelos (1998) Exploratory data analysis of the spatial distribution of wildfires in Portugal, 1980-1989. *Geographical Systems*, 5: 355-390.
118. Pereira, M.G., Trigo, R.M., Camara, C.C., Pereira, J.M.C. and Leite, S.M. 2005. Synoptic patterns associated with large summer forest fires in Portugal. *Agricultural and Forest Meteorology*, 129: 11-25.
119. Pollet J. and Omi P.N. 2002. Effect of thinning and prescribed burning on crown fire severity in ponderosa pine forests, *International Journal of Wildland Fire* 11: 1–10.
120. Preisler, H.K. Brillinger, D.R. Burgan, R.E. and Benoit, J.W., 2004. Probability based models for estimation of wildfire risk. *International Journal of Wildland Fire*, 13: 133-142.
121. Preisler, H. K. and Westerling, A. L. 2005. "Estimating Risk Probabilities for Wildland Fires." *Proceedings of the 2005 Joint Statistical Meetings*. Minneapolis, Minnesota.
122. Preisler, H.K. and Westerling, A.L. 2007 . Statistical Model for Forecasting Monthly Large Wildfire Events in Western United States. *Journal of Applied Meteorology and Climatology*, 46: 1020-1030.
123. Quintanilha, J. A. and Ho, L. L., 2006. Analyzing wildfire threat counts using a negative binomial regression model. *Environmetrics*, 17: 529-538.
124. Reed W., 1994. Estimating the Historic Probability of Stand-Replacement Fire Using the Age-Class Distribution of Undisturbed Forest, *Forest Science*. 40: 104–119.
125. Reed, W.J. and McKelvey, K.S., 2002. Power-law behaviour and parametric models for the size-distribution of forest fires. *Ecological Modelling* 150: 239-254.
126. Riera, P. and Mogas, J. 2004. Evaluation of a risk reduction in forest fires in Mediterranean region. *Forest Policy and Economics* 6: 521-528.
127. Schoenberg, F.P., Peng, R. and Woods, J. (2003). On the distribution of wildfire sizes. *Environmetrics* , 14: 583—592.
128. Tábara, D., Saurí, D. and Cerdan, R. 2003. Forest Fire Risk Management and Public Participation in Changing Socioenvironmental Conditions: a Case Study in a Mediterranean Region. *Risk Analysis* 23: 249-260.
129. Teensma, P.D.A. 1987. Fire History and Fire Regimes of the Central Western Cascades of Oregon. Ph.D. Dissertation. University of Oregon, Eugene, OR, 188.
130. Thompson W.A., Vertinsky I., Schreier H. and Blackwell B.A. 2000. Using forest fire hazard modelling in multiple use forest management planning, *Forest Ecology and Management* 134: 163–176.
131. Van Wagner C.E. 1977. Conditions for the start and spread of crown fire, *Canadian Journal of Forest Research* 7: 23–34.

- Van Wagner C.E. 1987. Development and structure of the Canadian Forest Fire Weather Index System. Canadian Forest Service, Ottawa, Ontario Forestry Technical Report nº 35.
132. Vázquez, A. and Moreno, J.M., 2001. Spatial distribution of forest fires in Sierra de Gredos (Central Spain). *Forest Ecology and Management*, 147: 55-65.
133. Velez R. 2002. Causes of fires in the Mediterranean Basin, *EFI proceedings* 45: 35–42.
134. Viedma, O. Moreno, J.M. and Rieiro, I. 2006 Interactions between land use/land cover change forest fires and landscape structure in Sierra de Gredos (Central Spain). *Environmental Conservation* 33: 212-225.