

DISTRIBUIÇÃO ESPACIAL DOS INCÊNDIOS FLORESTAIS NO DISTRITO DE SANTARÉM

RESUMO

Pretende-se realizar uma análise da “Distribuição espacial dos incêndios florestais no distrito de Santarém” ocorridos entre 2009 e 2013. Na primeira parte do estudo, pretende-se avaliar a distribuição do padrão pontual, em que se apresentaram diversos parâmetros estatísticos básicos, que fornecem uma visualização da localização espacial dos eventos, designados por parâmetros descritores;

1. Enquadramento da área de estudo

O distrito de Santarém pertence às Unidades Territoriais da Lezíria do Tejo e do Médio Tejo. É limitado a norte pelos distritos de Leiria e de Castelo Branco, a leste pelo distrito de Portalegre, a sul pelos distritos de Évora e Setúbal e a Oeste pelos de Lisboa e Leiria. É o terceiro maior distrito de Portugal e é composto por 21 concelhos (Figura 1).

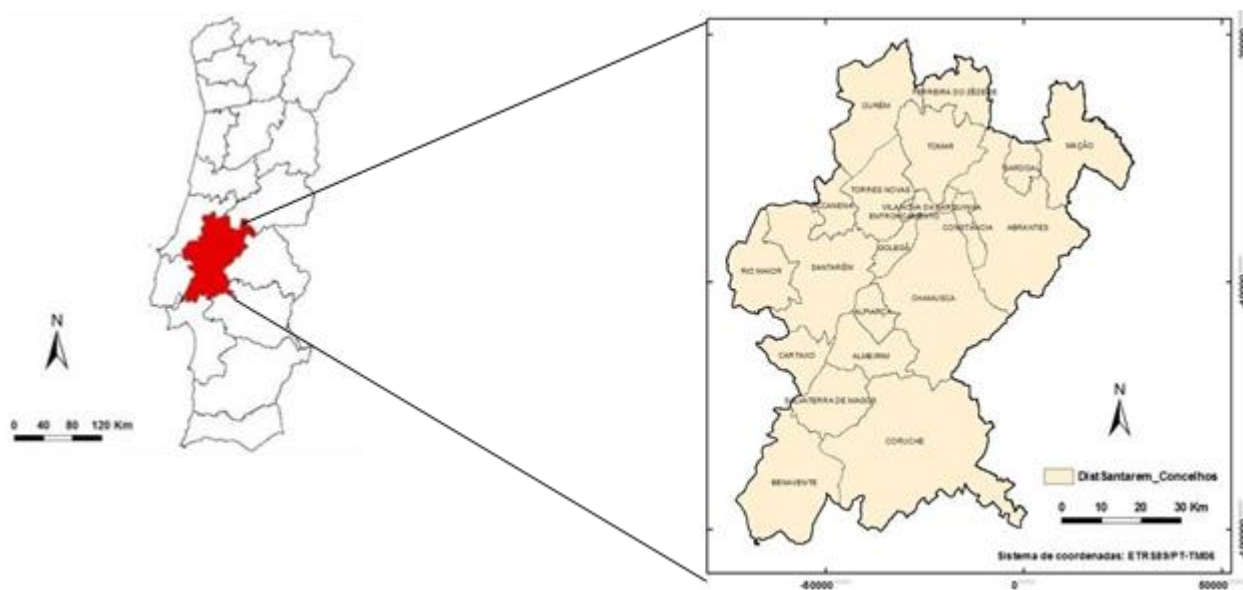


Figura 1 – Enquadramento do Distrito de Santarém

2. Dados

2.1– Fontes de Informação

Para a concretização deste trabalho, terá que adquirir os dados em diversas fontes, por forma a poder explorar e analisar os focos de incêndios florestais.

Tabela 1 – Dados recolhidos e informação associada.

Dados	Tipo	Fonte
CAOP Continente: Carta Administrativa Oficial de Portugal 2018.	Vetorial (polígono)	DGT – Direcção Geral do Território. http://www.dgterritorio.pt/cartografia_e_geodesia/cartografia/carta_administrativa_oficial_de_portugal_caop/caop__download/carta_administrativa_oficial_de_portugal__versao_2019__em_vigor/
Focos de Incêndio Florestal	Tabela Excel	ICNF – Instituto de Conservação da Natureza e das Florestas http://www2.icnf.pt/portal/florestas/dfci/inc/estat-sgif#list

2.2– Tratamento dos Dados

1. Verifique se a informação geográfica, referente aos focos de incêndio florestal, se encontra no sistema de coordenadas ETRS89/PT-TM06, se não se encontrar deverá proceder à sua transformação.
2. No que se refere à *shapefile* da CAOP, deve criar uma nova layer apenas com os limites do distrito de Santarém.
3. Em seguida, deve gerar uma layer de pontos apenas com os focos de incendio florestal no distrito de Santarém. Elimine os incêndios agrícolas, falsos alarmes e queimadas.

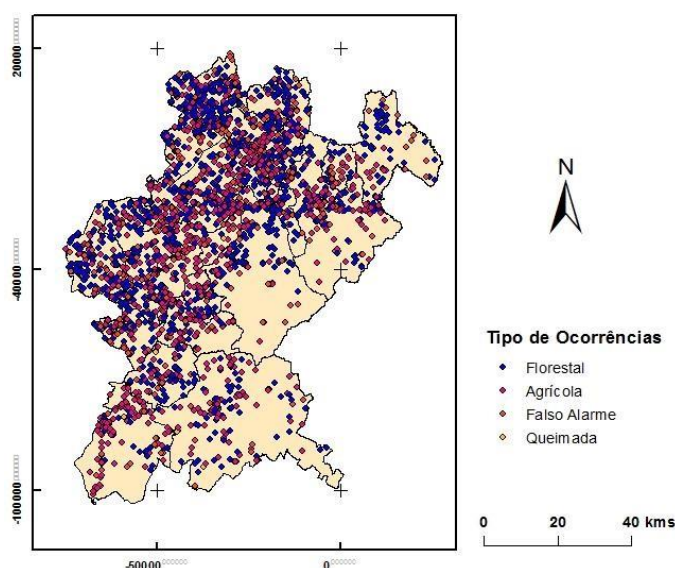


Figura 2 – Distribuição espacial dos incêndios florestais por classificação de ocorrências.

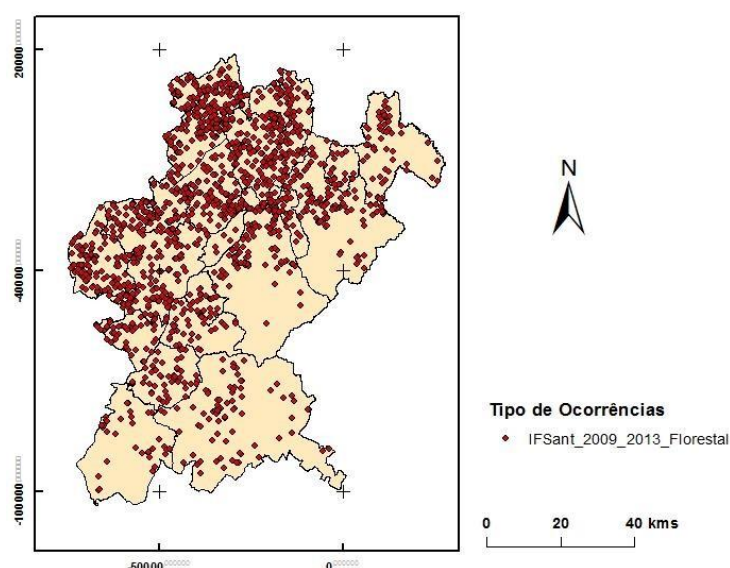


Figura 3 – Distribuição espacial dos incêndios florestais por classificação de ocorrências.

Antes de iniciar a visualização dos dados espaciais, os gráficos 1 e 2 dão uma noção, através da estatística clássica, do número de incêndios florestais ocorridos no período de 2009 a 2013.

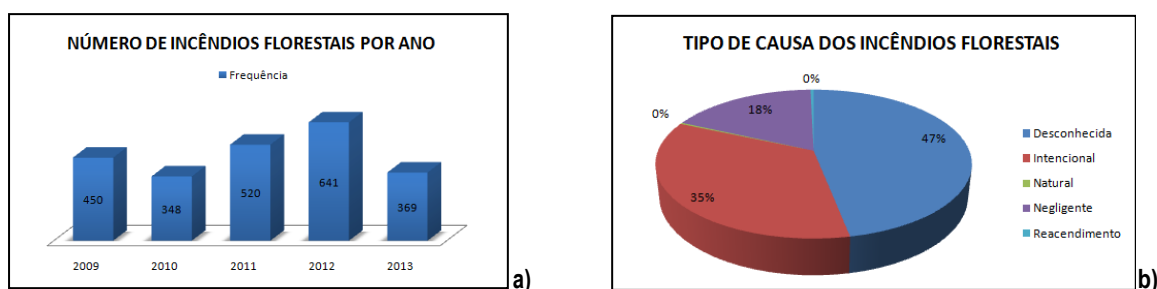


Gráfico 1 – Incêndios florestais entre os anos de 2009 e 2013: a) Frequência; b) Tipo de causa.

3. Análise da Distribuição de Padrões Pontuais

Na análise da distribuição de padrões pontuais, existem diversos parâmetros estatísticos básicos que fornecem uma visualização da localização espacial dos eventos. Através destes parâmetros, é possível avaliar o padrão de distribuição dos focos de incêndio identificando a existência de tendências espaciais, ou concluir que o padrão é distribuído aleatoriamente.

Ao gerar uma layer dos dados para o SIG verifique se existem *outliers* nos seus dados, por exemplo, focos que não pertencem a este distrito.

3.1– Visualização dos Dados Espaciais

Na fase da visualização dos dados espaciais, usam-se os descritores de padrões de pontos que são parâmetros de estatística básicos, os quais dão uma perceção da distribuição dos pontos no espaço em termos da sua variação e orientação, tais como:

Frequência (n) – o número de eventos que ocorrem numa área de estudo. A tabela 1 e o gráfico 1 mostram que, analisando os dados por ano, os valores mínimo e máximo variaram entre 348 (no ano de

2013) e 641 (no ano de 2012), respetivamente. No conjunto global, entre 2009 e 2013, ocorreram um valor máximo de 2328 eventos.

Densidade (λ) – a razão entre o número de eventos que ocorrem numa região e a área da mesma (Frequência / Área).

Tabela 2 – Frequência e densidade dos incêndios florestais referente ao período entre 2009 e 2013.

	Distrito de Santarém					
Descritores	2009	2010	2011	2012	2013	2009-2013
Frequência	450	348	520	641	369	2328
λ (eventos/km ²)	0,0670	0,0518	0,0774	0,0954	0,0549	0,3465

Central Feature – parâmetro que identifica a localização do evento mais central (Figuras 4a e 4b). Esta *feature* tem associada a menor distância total às restantes *features*.

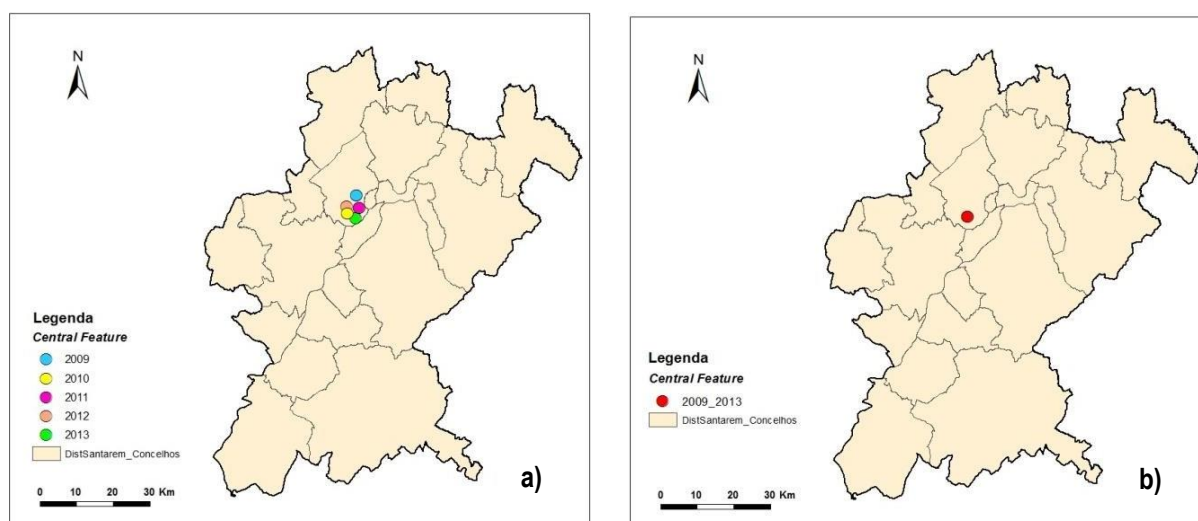


Figura 4 – *CentralFeature*: a) Conjunto individual (por anos); b) Conjunto global (período de 2009 a 2013).

Pela análise da figura 4, verifica-se que os eventos mais centrais se localizam todos, por anos e no conjunto global, na zona norte do distrito e no mesmo concelho, o de Torres Novas.

Centro geométrico – indicador da tendência central da distribuição (Figura 5).

Este parâmetro identifica a localização média de todos os eventos e é determinado pela média das suas coordenadas X e Y. Para cada ano, verifica-se que os centros geométricos estão localizados na zona norte do distrito, tal com a *Central Feature*. Na análise individual por anos, as referidas posições localizam-se no concelho de Torres Novas em 2011 e 2013 e, para os anos de 2009, 2010 e 2012, no concelho da Golegã.

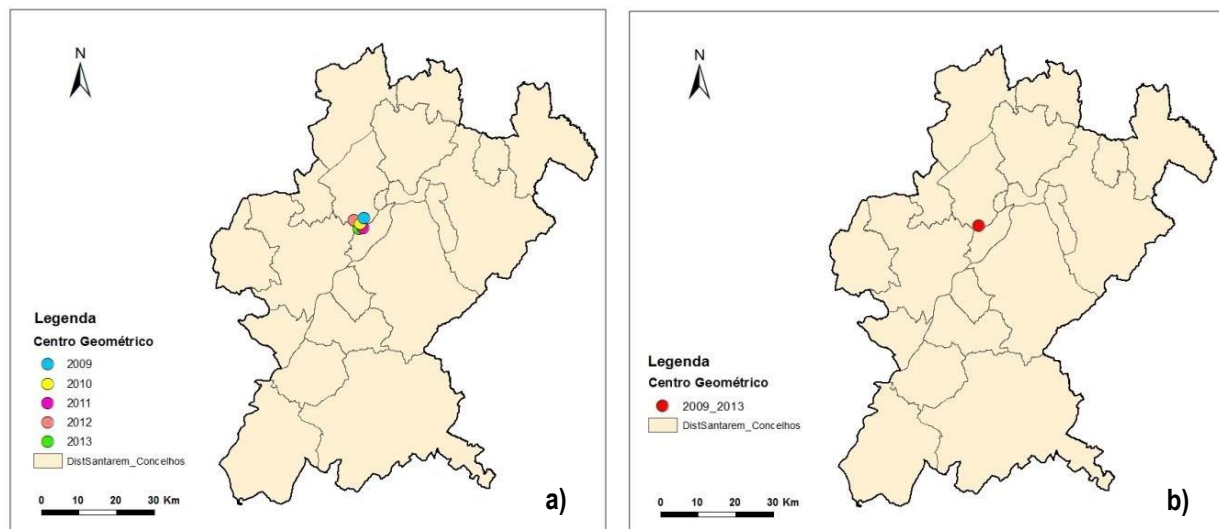


Figura 5 – Centro Geométrico: a) Conjunto individual (por anos); b) Conjunto global (período de 2009 a 2013).

Distância Padrão (ou dispersão espacial) – uma medida do grau de dispersão, ou de concentração, da distribuição espacial em torno do seu centro geométrico. A figura 6 mostra a concentração dos pontos a 68% e 95% de confiança.

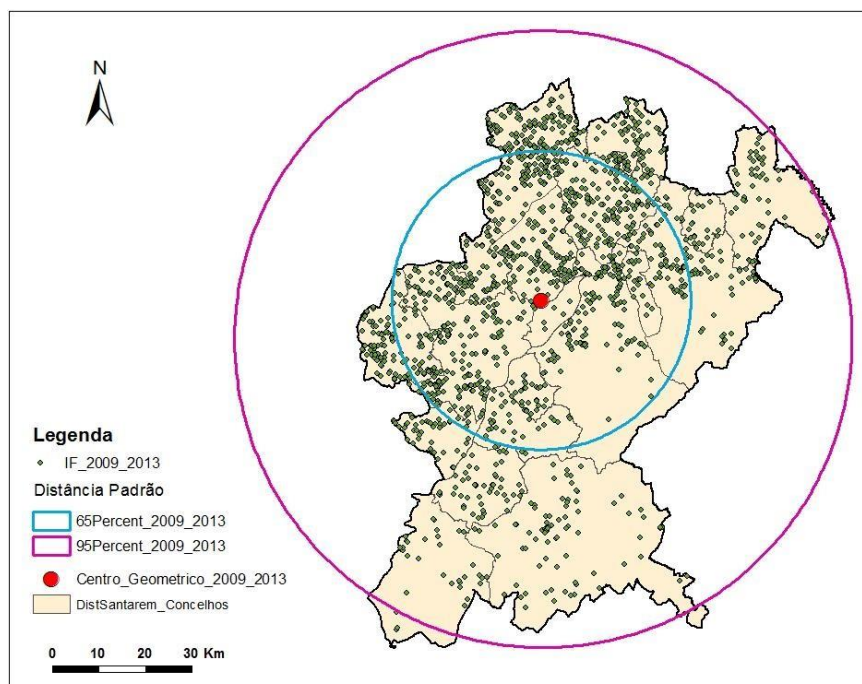


Figura 6 – Distância Padrão com 68% e 95% de confiança.

Na figura 6 estão representadas os círculos centrados no centro geométrico dos dados com raios distintos: 68% de confiança, a azul e a 95% de confiança, a roxo.

Elipse padrão – indica a orientação da distribuição espacial dos eventos através dos seus semi-eixos (Figura 7). O semi-eixo maior define a direção de máxima dispersão da distribuição e o semi-eixo menor define a mínima dispersão.

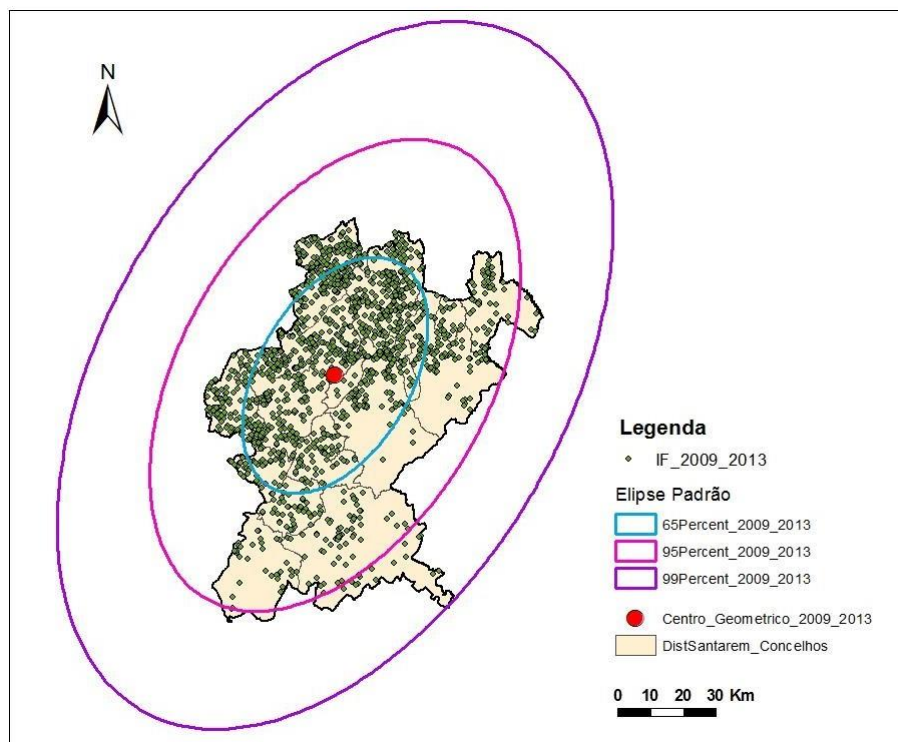


Figura 7– Elipse Padrão com 68% (a azul), 95% e 99% de confiança.

Analisando a figura 7, constata-se que, no conjunto global dos eventos entre 2009 e 2013, existe um efeito direcional da elipse padrão mais acentuado na orientação SW-NE.

A que se deve este facto?

Quais os concelhos que apresentam uma maior dispersão ?

3.2– Análise Exploratória dos Dados

Na etapa da análise exploratória dos dados, usam-se os detetores de padrões de pontos, que são técnicas da análise exploratória dos dados que estão divididas em propriedades de 1ª e de 2ª ordem. Nas de 1ª ordem, estão os métodos da contagem por quadrantes e do estimador de intensidades (globais ou de grande escala) e nas de 2ª, os métodos baseados na medição de distâncias, método do vizinho mais próximo e a função K (locais ou de pequena escala). Estes métodos de análise, de um padrão espacial de incidência pontual, vão ser explorados e analisados nas etapas que se seguem.

3.2.1– Método da Contagem por Quadrantes

O método da contagem de quadrantes consiste em criar uma grelha regular na qual são contabilizados os eventos pontuais. É essencial ter em atenção que quadrantes grandes não permitem uma boa descrição do padrão pontual e quadrantes demasiado pequenos podem não conter eventos.

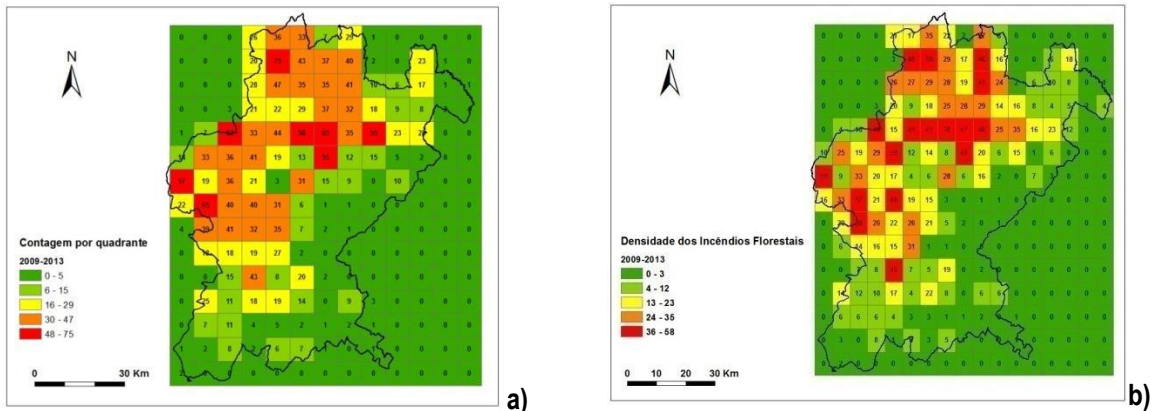


Figura 8 – Mapa resultante do método de contagem por quadrantes no conjunto global: a) grelha 8000x8000 e b) 6000x8000.

Tabela 3–Método da Contagem por Quadrantes

Variáveis	Grelha 8x8	Grelha 6x8
Nº Quadrantes	155	255
Pontos (Nº de Incêndios Florestais)	2328	2328
Variância	280,6811	174,1889
Média	11,9333	9,1255
VMR	23,5208	19,0882
Teste do Qui-Quadrado ($X^2 = (n-1) \cdot \text{VMR}$)	4563,028	4848,395
X_{teste}^2	155,7215	209,8829

Com base na tabela 3, em ambos os casos, rejeita-se a hipótese nula de que o padrão é aleatório, verificando-se também que o VMR é superior a 1, que comprova a existência de padrão de aglomerados (*clusters*) e, portanto, uma variabilidade espacial grande.

É de destacar, ainda, que a concentração dos *clusters* se encontra a noroeste e no centro do distrito.

3.2.2– Estimador de densidade (Kernel Density)

O método *Kernel density* estima em cada célula de uma grelha a densidade de eventos na região de estudo, permitindo identificar as zonas onde existe uma maior e menor concentração de eventos. Este método usa uma função densidade que minimiza os efeitos de transição entre zonas com grande e baixa densidade de eventos.

Para este estudo, geraram-se vários mapas resultantes do método kernel: um com os dados de todos os anos e outros com os dados de cada ano em análise. A largura de banda foi sempre a mesma e o raio foi de 8500 metros (Figuras 9 e 10).

O mapa Kernel representa o resultado da interpolação dos eventos considerados na análise, onde se observa a intensidade pontual de ocorrência destes na área em estudo. As cores vermelhas e laranja mostram as manchas com densidade muito alta e alta, respetivamente, localizadas na zona norte,

noroeste e central. As manchas referentes às densidades média e baixa, cor amarela e verde claro, estão mais distribuídas no distrito, concentrando-se, à semelhança das outras cores, na parte norte e oeste da área de estudo. A cor verde escura, que corresponde à mancha com densidade muito baixa, concentra-se nos concelhos localizados na zona este e sul do distrito de Santarém.

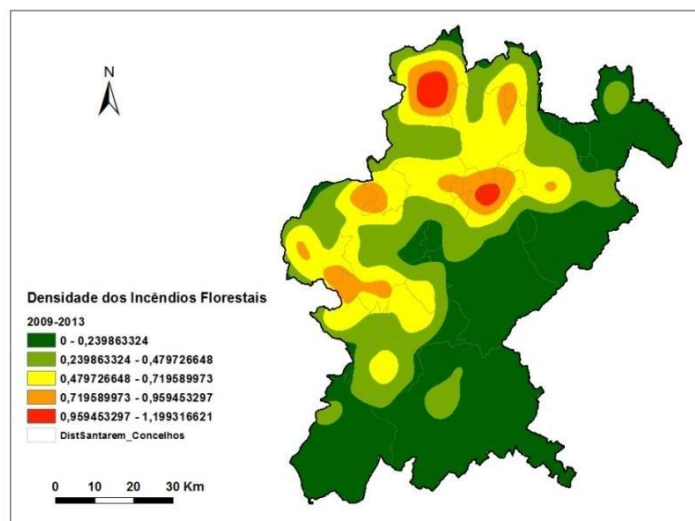


Figura 9 – Mapa de Intensidade dos Incêndios Florestais desde o ano 2009 ao 2013.

Nota: Necessidade de formatação dos valores numéricos da legenda para 2 casas decimais no máximo.

De seguida, apresentam-se os mapas Kernel para o conjunto individual, que permitem visualizar a distribuição espacial dos incêndios florestais desde 2009 a 2013 (Figura 10).

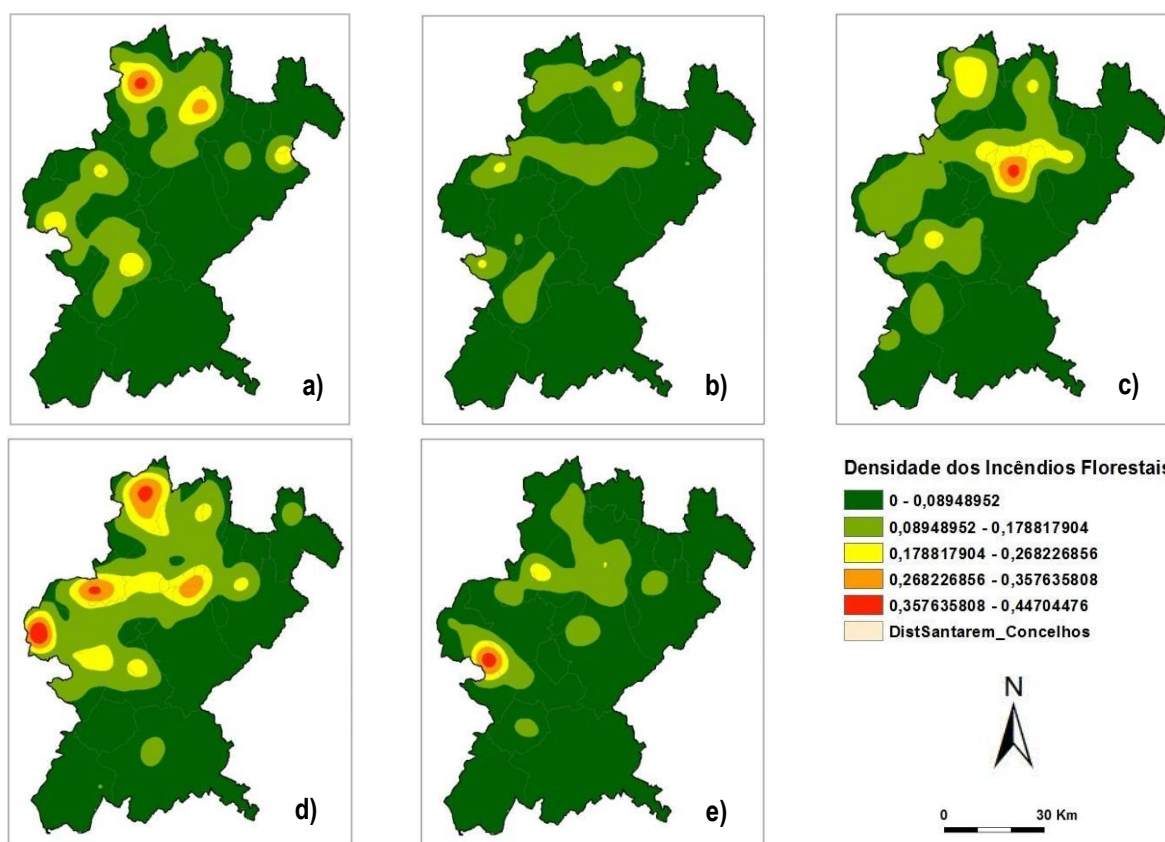


Figura 10 – Mapa de Densidade dos Incêndios Florestais: a)2009; b) 2010; c) 2011; d) 2012 e e) 2013.

Numa análise global, observa-se que o ano de 2012 foi o que teve mais incêndios florestais, dado que é o que assinala mais manchas com a cor vermelha (*hot-spots*), isto é, em que existe uma forte tendência de agregação. A nível individual, verifica-se que:

- ✓ Em 2009, os concelhos mais afetados foram os de Ourém e Tomar, localizados a noroeste do distrito;
- ✓ 2010 foi um ano mais calmo, tendo ocorrido incêndios florestais com manchas de baixa (verde claro) e muito baixa densidade (verde escuro), sendo esta última mancha a predominante;
- ✓ Em 2011, voltaram a aparecer as manchas com muito alta e alta densidade (vermelho e laranja), lesando a zona a oeste do concelho da Chamusca, localizado a noroeste da área em estudo;
- ✓ 2012, tal como já referido, foi o ano em que ocorreram mais incêndios florestais;
- ✓ Em 2013, os *hot-spots* incidiram nos concelhos de Santarém e Rio Maior, a sul.

Pode-se concluir que, no quinquénio em estudo, aproximadamente 50% do distrito foi afetado por incêndios florestais, tendo as zonas norte e centro sido as mais atingidas, evidenciando que os *hot-spots* se assinalaram nos concelhos mais a noroeste e a oeste do distrito de Santarém.

3.2.3– Método do Vizinho mais próximo

Os métodos de análise de efeitos de segunda ordem baseiam-se na distância entre os diferentes pontos procurando inferir a natureza da sua distribuição local.

O método do vizinho mais próximo baseia-se na estimação da distribuição dos eventos através da distância entre os pontos na região de interesse. Os métodos normalmente utilizam a distância euclideana entre os pontos e a determinação é feita de ponto a ponto.

O Método do vizinho mais próximo baseia-se na análise de função G determinada de forma cumulativa. O gráfico da função G dá informação sobre a forma como os eventos estão distribuídos na região de interesse.

Um gráfico que aumenta rapidamente em pequena distância indica um padrão de aglomerados. Um gráfico que aumenta lentamente até aproximadamente à distância entre eventos e depois aumenta rapidamente indica que os eventos apresentam um padrão de dispersão.

A função G no ArcGIS está implementada através da *Spatial Statistics Toolbox > Average Nearest Neighbor*. Após o algoritmo correr sobre os dados é elaborado um relatório que apresenta dois descritores estatísticos que devem ser analisados: o *z-score* e o *p-value*. O *z-score* e o *p-value* apresentam a medida de significância estatística que indica se a hipótese nula deve ou não ser rejeitada. A hipótese nula neste algoritmo é de estarmos na presença de uma amostra com uma distribuição aleatória.

O *p-value* indica a probabilidade de que o padrão pontual seja aleatório. Um valor elevado indica um padrão aleatório e um valor pequeno indica um padrão não aleatório. O *z-score* é o desvio padrão da amostra. Os dois descritores indicados estão associados à distribuição gaussiana. O manual do ArcGIS apresenta uma tabela que ajuda a interpretação dos resultados do teste estatístico de hipóteses com a hipótese zero sendo da *Complete Spatial Randomness (CSR)*:

z-score (Standard Deviations)	p-value (Probability)	Confidence level
< -1.65 or > +1.65	< 0.10	90%
< -1.96 or > +1.96	< 0.05	95%
< -2.58 or > +2.58	< 0.01	99%

Informação sobre o algoritmo de Average Nearest Neighbor pode ser encontrada em linha no sítio:
<http://resources.arcgis.com/en/help/main/10.1/index.html#//005p00000008000000>

Para mais informação sobre o z-score e p-value ver:
<http://resources.arcgis.com/en/help/main/10.1/index.html#//005p00000006000000>

Para o caso particular dos eventos dos incêndios florestais o relatório gerado relativo à aplicação do método do vizinho mais próximo encontra-se na Figura 11.

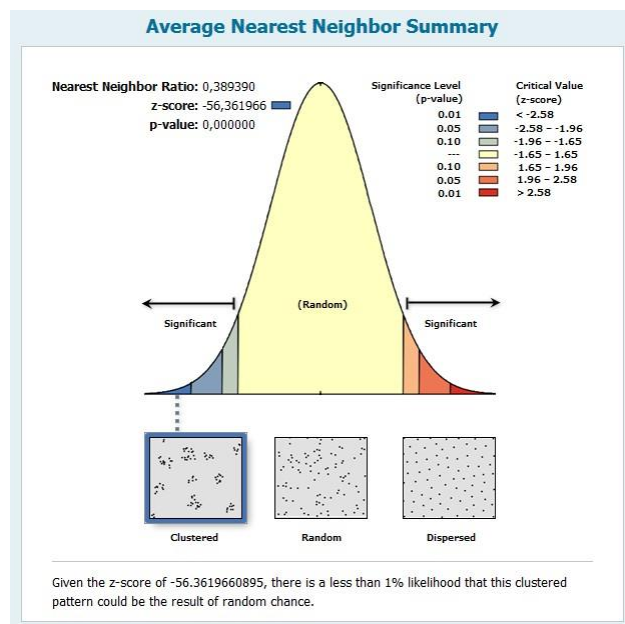


Figura 11 – Resultado do método do vizinho mais próximo entre os anos de 2009 e 2013, conjunto global.

Os resultados obtidos (z-score de -56.36 e p-value de 0.00) indicam que a hipótese nula (CSR) pode ser rejeitada com um nível de confiança de 99%. Ou seja, existe menos de um 1% de probabilidade deste padrão ser aleatório, confirmando-se que estamos na presença de um padrão de cluster.

Aplicou-se a mesma análise para cada ano individualmente, tal como se mostra de seguida. Verifica-se a tendência para o aglomerado em cada ano, tal como para a análise conjunta de todos os anos.

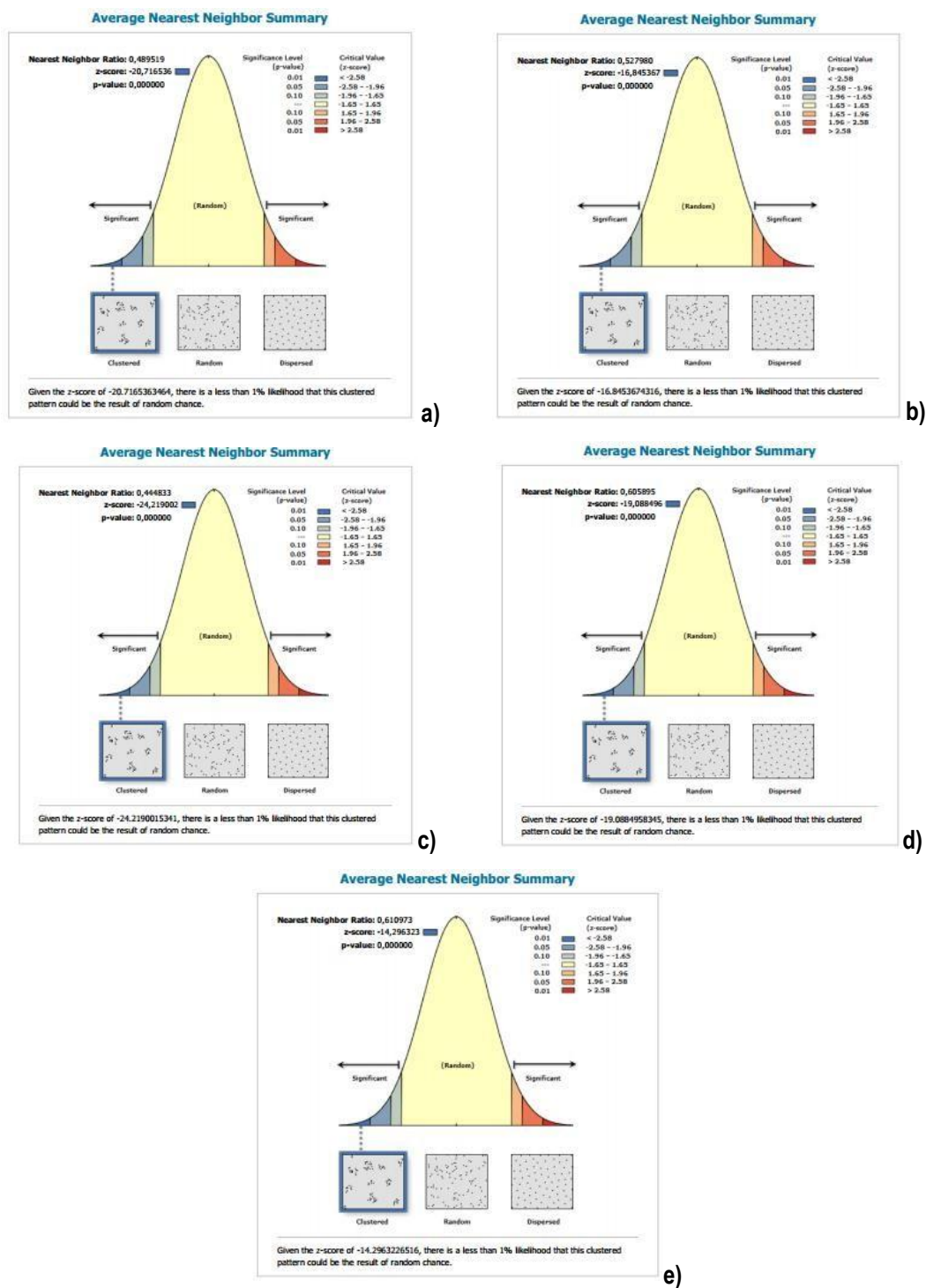


Figura 12 – Resultado do método do vizinho mais próximo para os anos: a) 2009; b) 2010; c) 2011; d) 2012 e e) 2013.

3.2.4– Função K-Ripley

O método do vizinho mais próximo proporciona uma primeira visão da distribuição espacial dos pontos, no entanto, tem como desvantagem considerar apenas a distância mínima entre dois pontos.

A função *K-Ripley* baseia-se nas distâncias entre todos os eventos da área em estudo, sendo mais sensível a distâncias maiores, ou seja, ajuda a determinar o agrupamento em diferentes distâncias. Esta função varia com o diâmetro dos clusters pelo que, fazendo variar o raio d , é possível detetar o padrão espacial em diferentes escalas.

Neste método é possível comparar o gráfico da função K do padrão espacial observado com o gráfico de uma função K no pressuposto de um padrão completamente aleatório (hipótese CSR). Para tal, é necessário obter valores para esta função com envelopes de simulação. Estes permitem validar, para um dado nível de confiança, se existe ou não a aleatoriedade espacial completa. Os envelopes de confiança são obtidos pela realização de permutações.

Para a realização deste método, foram criados envelopes de confiança com 9 permutações (Figura 13).

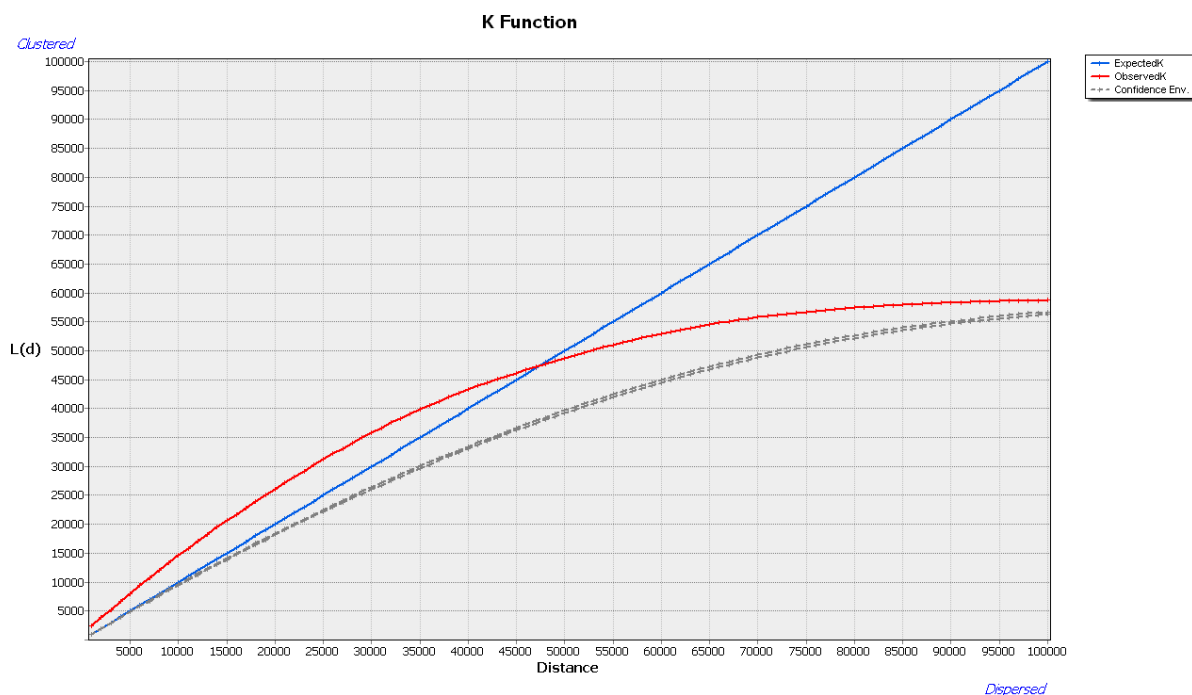


Figura 13– Resultado da aplicação da função K Ripley.

Na figura 13, observa-se que o padrão de distribuição espacial dos incêndios florestais variou em diferentes escalas. Constata-se que, até aproximadamente aos 45 km, o gráfico da função dos eventos observados está acima do gráfico da função na hipótese CSR, logo até esta distância o padrão observado apresenta tendência para agrupamento e a partir dos 45 km o padrão observado tende a dispersar.