

LABORATÓRIO 1: ANÁLISE DE PADRÕES DE DISTRIBUIÇÃO DE PONTOS

O objetivo deste laboratório é ilustrar as várias formas de analisar padrão de pontos, a partir de alguns conjuntos de dados. As ferramentas de análise disponíveis no SPRING são: o Interpolador Kernel, o Método do Vizinho Mais Próximo e a Função K.

Aplicação do Estimador de densidade por Kernel

Primeiramente aplicou-se o estimador de densidade por Kernel, o qual refere-se a uma medida de um suavizador de pontos para os dados de violência no Município de Porto Alegre/RS. Para isso, utilizou-se o banco de dados Porto Alegre no Spring.

Inicialmente ativou-se o banco de dados e o projeto com os respectivos dados. Visualizou-se os dados pontuais correspondentes aos eventos de homicídios, suicídios e acidentes de trânsito e após isso procedeu-se a fase da aplicação do método de Kernel. No menu “Análise” acessou-se “estatística espacial” e posteriormente “estimador de densidade por Kernel”. Na interface apresentada, seleciona-se o tipo de dado (neste caso, ponto – associado ao PI “eventos_mortalidade”), define-se a largura da banda e seleciona-se a categoria e o PI de saída (o qual será MNT).

Para a análise dos resultados, foram utilizadas duas larguras de banda diferentes, apresentadas nas Figuras 1 a 6, as quais foram 1500 e 5000 metros. Nas figuras 1 e 2 é apresentada a grade regular gerada a partir do estimador de densidade por Kernel.

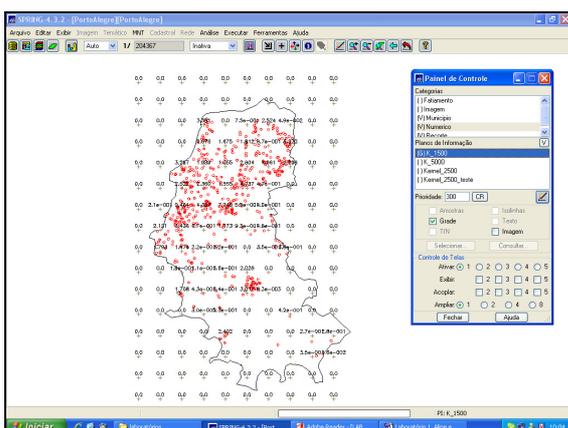


Figura 1: Grade regular – banda 1500m.

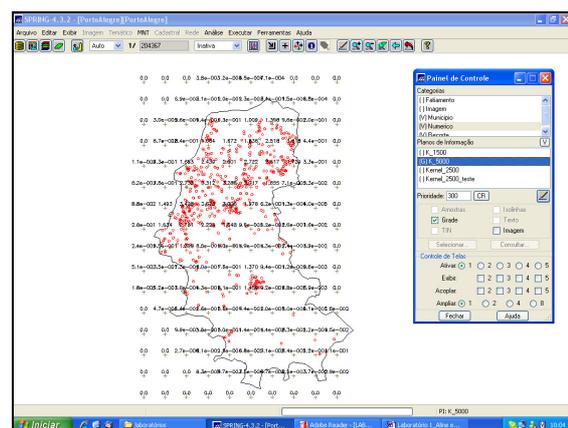


Figura 2: Grade regular – banda 5000m.

A fim de refinar e melhorar a interpretação do resultado obtido a partir de uma grade numérica fez-se a transformação do tipo Grade \rightarrow Imagem, resultando nas Figuras 3 e 4 apresentadas a seguir.

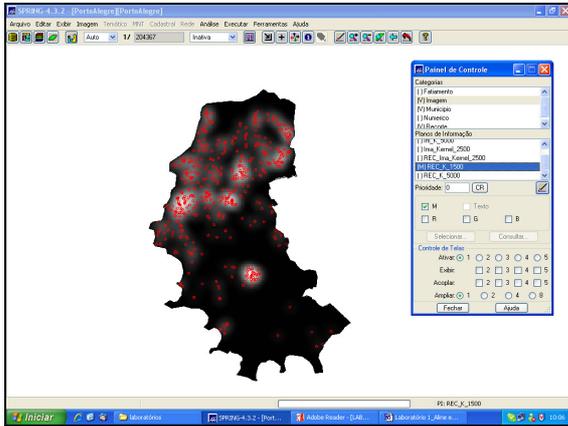


Figura 3: Imagem – banda 1500m.

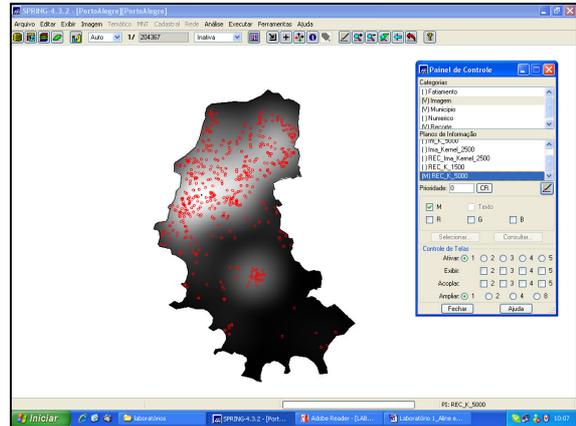


Figura 4: Imagem – banda 5000m.

Na etapa posterior realizou-se o fatiamento na grade gerada a partir do método de Kernel. Essa etapa foi cumprida no programa em LEGAL, no menu “análise” do Spring. As classes de densidade pelo fatiamento foram classificadas em: baixa, baixa-média, média, média-alta e alta, considerando um intervalo fixo determinado a partir da diferença entre a cota mínima e a cota máxima, em ambos os valores da banda do Kernel. As figuras 5 e 6 demonstram as diferenças após o fatiamento da grade.

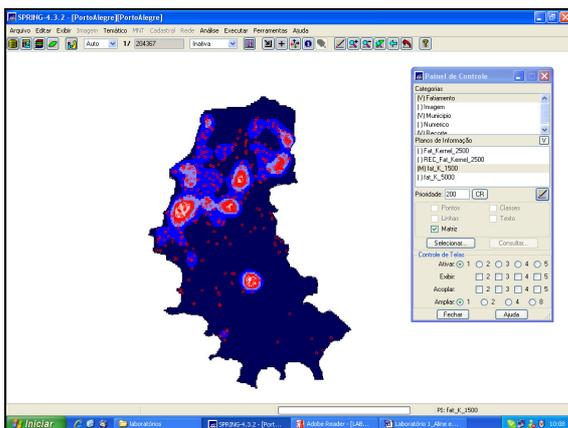


Figura 5: Fatiamento da grade – banda 1500m.

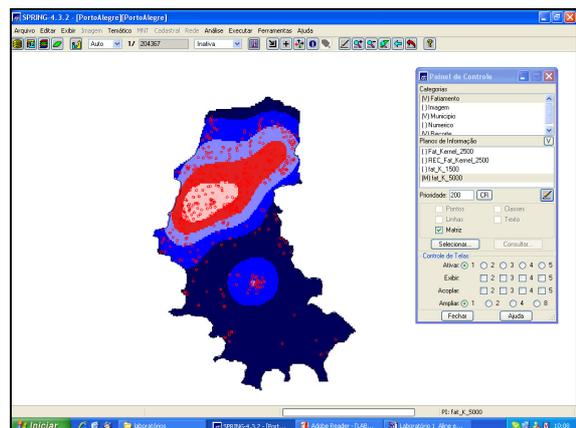


Figura 6: Fatiamento da grade – banda 5000m.

Os resultados apresentados segundo a análise do estimador de densidade por Kernel, fornecem a idéia do que compreende a suavização das características pontuais de acordo com o valor de banda utilizado. No primeiro caso, com o valor da banda de menor número (1000m), a imagem gerada apresenta menor generalização dos dados se comparada ao resultado encontrado no mapa com uma banda de número maior. A suavização apresentada pelo segundo resultado (banda com valor de 5000m), torna os dados menos confiáveis se considerar como objetivo uma análise mais detalhada.

Aplicação do Kernel considerando o valor do atributo

Nesta aplicação, o Kernel refere-se a uma medida de quantidade total do atributo por unidade de área. Para isso, utilizou-se o banco de dados com bairros de São Paulo, com atributos sobre o percentual de idosos (mais de 70 anos).

Inicialmente ativou-se o banco de dados e o projeto com os respectivos dados. Visualizou-se os dados de áreas correspondentes aos bairros da parte central e leste da cidade de São Paulo e após isso procedeu-se a fase da aplicação do método de Kernel. No menu “Análise” acessou-se “estatística espacial” e posteriormente “estimador de densidade por Kernel”. Na interface apresentada, seleciona-se o tipo de dado (neste caso, área – associado ao PI “mapa_bairros”), define-se a largura da banda e seleciona-se a categoria e o PI de saída (o qual será MNT).

Para fins de comparação, foram utilizadas duas larguras de banda diferentes, apresentadas nas Figuras 7 a 12, as quais foram 2000 e 5000 metros. Nas figuras 7 e 8 é apresentada a grade regular gerada a partir do Kernel.

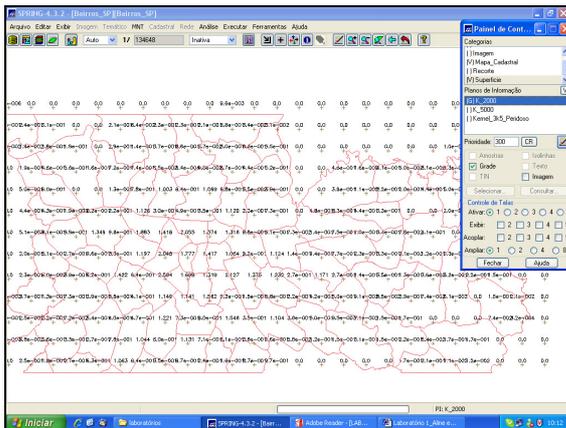


Figura 7: Grade regular – banda 2000m.

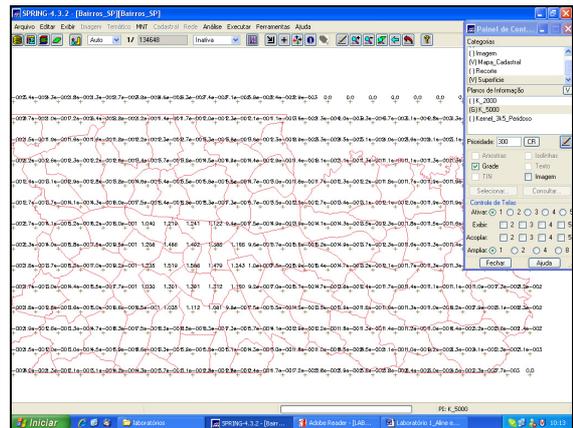


Figura 8: Grade regular – banda 5000m.

Da mesma forma ao exemplo anterior, para refinar e melhorar a interpretação do resultado obtido a partir de uma grade numérica transformou-se o tipo Grade em Imagem, resultando nas Figuras 9 e 10.

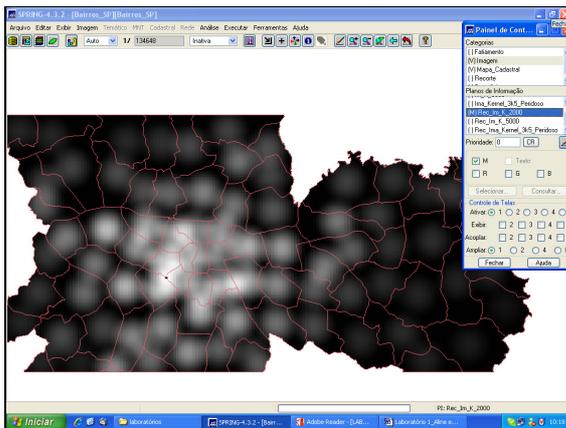


Figura 9: Imagem – banda 2000m.

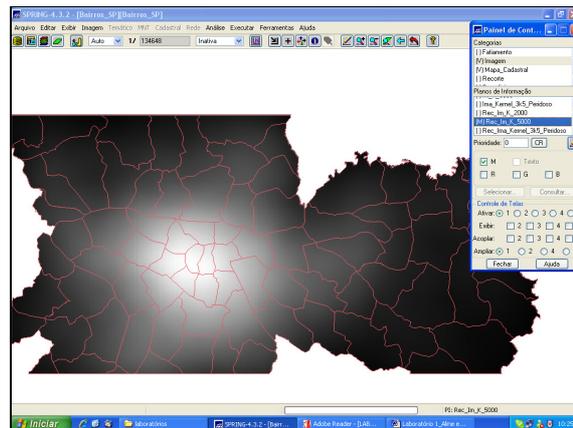


Figura 10: Imagem – banda 5000m.

Na etapa posterior realizou-se o fatiamento na grade gerada a partir do método de Kernel. Essa etapa foi cumprida no programa em LEGAL, no menu “análise” do Spring. As classes do fatiamento foram classificadas em: baixa, baixa-média, média, média-alta e alta, considerando um intervalo fixo determinado a partir da diferença entre a cota mínima e a cota máxima, em ambos os valores da banda do Kernel. As figuras 11 e 12 demonstram as diferenças após o fatiamento da grade.

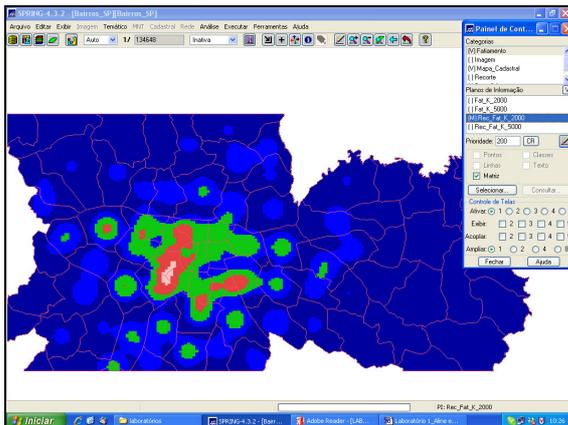


Figura 11: Fatiamento da grade – banda 2000m.

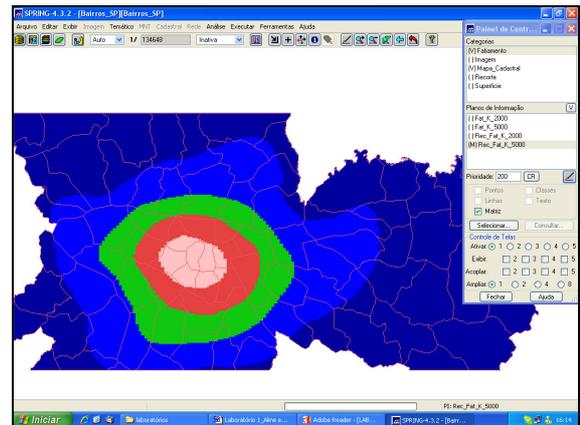


Figura 12: Fatiamento da grade – banda 5000m.

A análise do método de Kernel de acordo com o valor do atributo considerando amostras de áreas, apresenta resultados semelhantes ao exercício anterior. Os valores das bandas que foram utilizadas correspondem a 2000 e 5000 metros, respectivamente. Da mesma forma, quanto menor o número da banda utilizada mais detalhado é o resultado final. No entanto, a banda com maior número apresenta uma suavização maior dos dados, generalizando as superfícies de áreas estudadas.

Análise do vizinho mais próximo – banco de dados de Porto Alegre

A técnica de análise pelo método do vizinho mais próximo consiste em um gráfico de frequência acumulada da distância de cada ponto (evento) em relação ao vizinho mais próximo.

Ao iniciar o Spring, ativou-se o banco de dados e o projeto referente aos dados de eventos da cidade de Porto Alegre. Para esta análise, no menu “análise” seleciona-se “estatística espacial” e logo, “análise univariada de pontos”. Na interface apresentada, a análise a ser escolhida corresponde ao “vizinho mais próximo”, a distância mínima (0) e o número de intervalos (10) foram mantidos, enquanto a distância máxima foi modificada três vezes com valores diferentes em cada uma delas: 1000m, 1700m e 3400m. Nas figuras 13, 14 e 15 podem ser observados os gráficos resultantes de cada uma das análises.

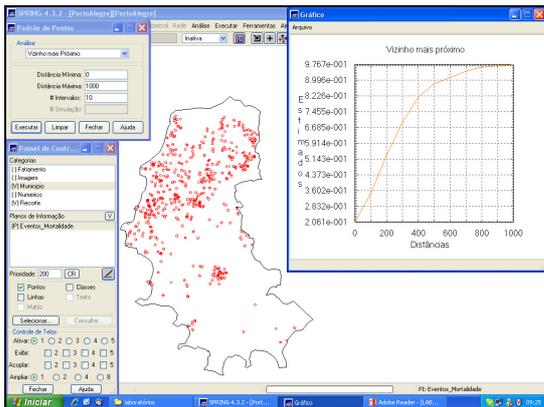


Figura 13: Método vizinho mais próximo – dist.máx. 1000m.

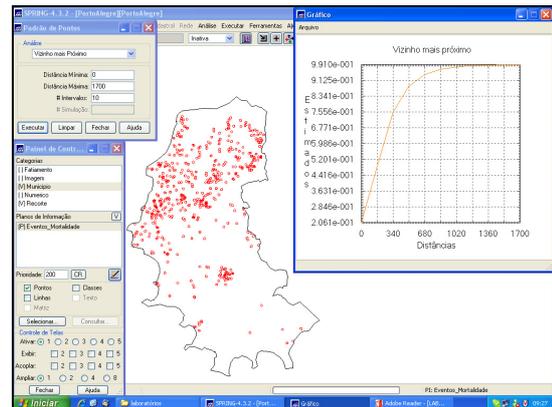


Figura 14: Método vizinho mais próximo – dist.máx. 1700m.

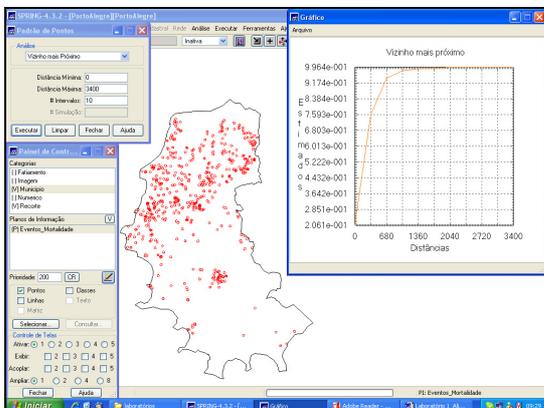


Figura 15: Método vizinho mais próximo – dist.máx. 3400m.

Na análise pelo método do vizinho mais próximo, são apresentados três resultados diferentes. No primeiro caso utilizou-se uma distância máxima de 1000m, configurando um gráfico mais suave, com uma elevação menos brusca em relação aos demais. Esse caso define, segundo a interpretação do gráfico, uma elevação próxima a ser constante (45°), indicando aleatoriedade dos dados. O segundo resultado é um intermediário entre o primeiro e o terceiro gráfico. O gráfico resultante da distância máxima de 3400m mostra uma elevação brusca, sugerindo um processo de agrupamento. Desta forma, quanto maior a distância considerada entre os eventos, pode-se sugerir que ocorre um agrupamento dos dados na escala que foi considerada.

Análise do vizinho mais próximo com simulação – banco de dados de Porto Alegre

O método do vizinho mais próximo com simulação permite a comparação da função acumulada das amostras com as funções de dados gerados aleatoriamente (simulação superior e inferior).

Para este caso, utilizou-se o mesmo procedimento do método anterior do vizinho mais próximo, porém acrescentou-se o número de simulações. Desta forma, foram consideradas as mesmas distâncias do método anterior, com o número de intervalos e número de simulações igual a dez (10) nos três casos.

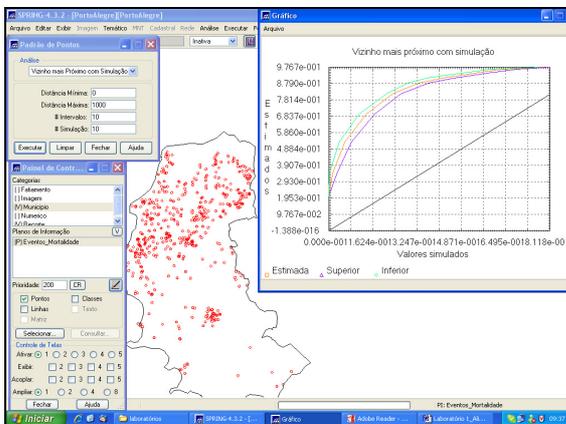


Figura 16: Método vizinho mais próximo com simulação – dist.máx. 1000m.

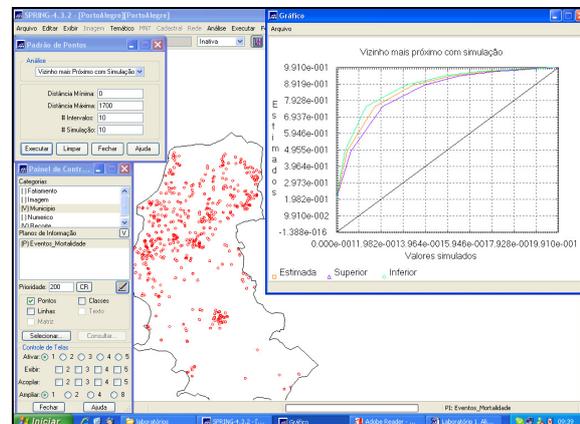


Figura 17: Método vizinho mais próximo com simulação – dist.máx. 1700m.

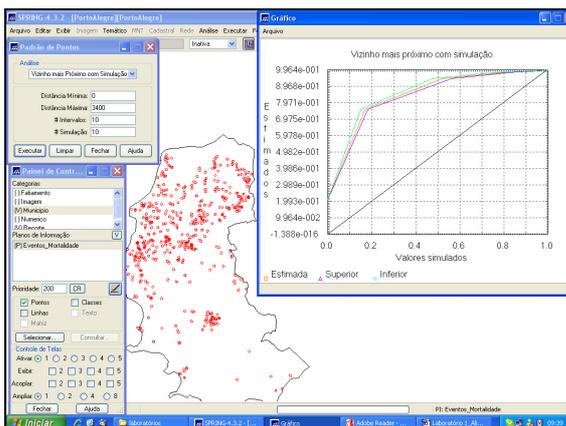


Figura 18: Método vizinho mais próximo com simulação – dist.máx. 3400m.

Na análise do vizinho mais próximo com simulação para o banco de dados de Porto Alegre foram utilizadas as mesmas distâncias calculadas no exercício anterior. Para todas as distâncias, em todos os gráficos apresentados, a função se destacou acima de uma reta de 45°, demonstrando o agrupamento dos dados de acordo com a escala.

Análise da Função L – Derivada da Função K

A função K considera escalas maiores para análise de padrões pontuais. Para uma análise que facilite a interpretação dos resultados, utiliza-se a função auxiliar L.

Para a elaboração dos gráficos a seguir, foi utilizada como análise a Função L, a distância mínima – zero (0), o número de intervalos – dez (10), e a distância máxima no primeiro gráfico é igual a 1000m e no segundo, 10000m.

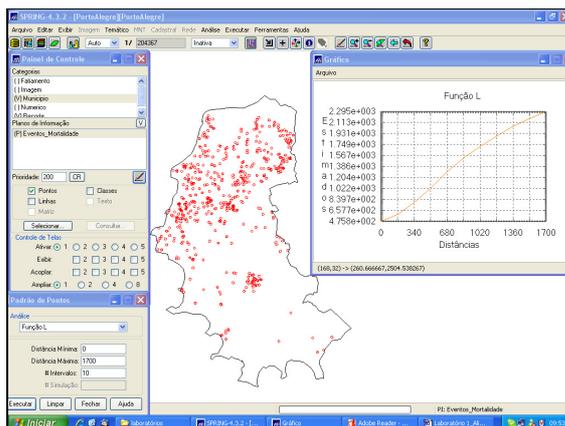


Figura 19: Gráfico da Função L – Dist. Máx. 1000m.

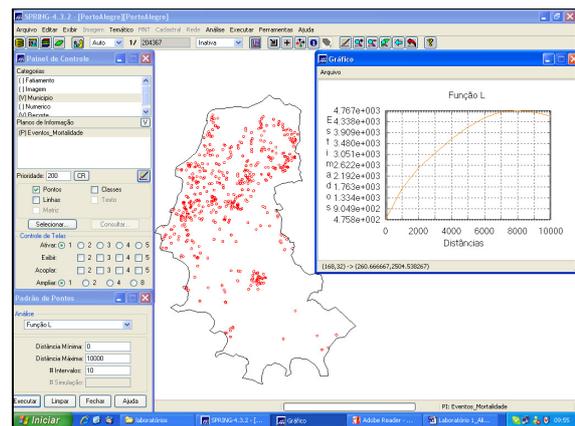


Figura 20: Gráfico da Função L – Dist. Máx. 10000m.

A interpretação dos gráficos da Função L busca explicar se há evidências de agregação a partir da distribuição espacial dos dados (eventos), aleatoriedade ou ordenação regular. Neste caso, de acordo com a interpretação sugerida, os gráficos apresentam resultados de agregação dos dados espacialmente, uma vez que a maior parte das distâncias observadas correspondem a extremos positivos.

Análise da Função L com simulação

A análise da Função L com simulação é semelhante a análise do vizinho mais próximo, a fim de estimar a significância dos desvios da distribuição L (h) em relação a aleatoriedade. O principal objetivo é realizar as simulações sobre a região R e computar os valores, superior e inferior.

Para esta análise, no menu “análise” selecionou-se “estatística espacial” e “análise univariada de pontos”. Após isso definiu-se a análise a ser utilizada, neste caso, Função L com simulação, determinou-se a distância mínima (0), a distância máxima (10000m e 1700m). Para o número de intervalos e de simulação escolheu-se 10 (dez). As figuras 21 e 22 apresentam os gráficos obtidos.

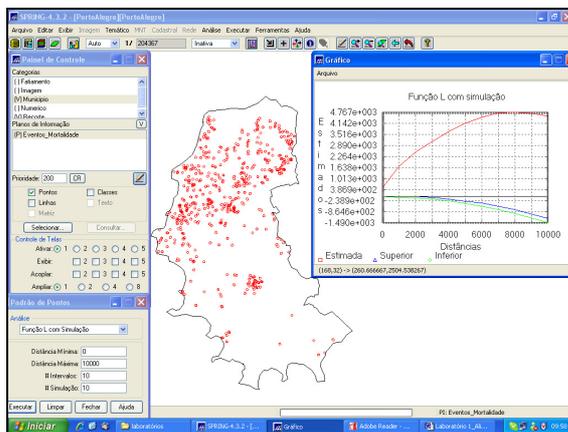


Figura 21: Gráfico da Função L com simulação – Dist. Máx. 10000m.

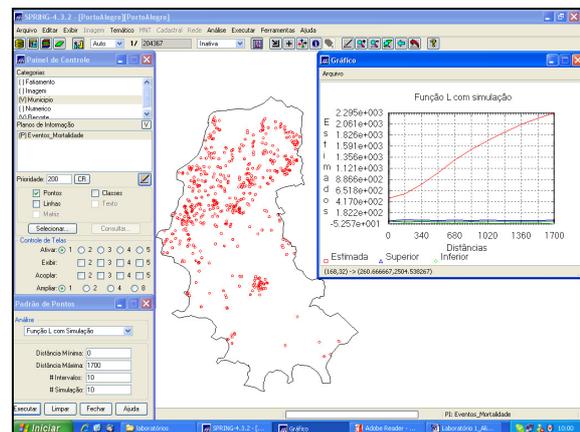


Figura 22: Gráfico da Função L com simulação – Dist. Máx. 1700m.

Para a interpretação dos gráficos gerados a partir da análise da Função L com simulação, parte-se do pressuposto de que se valores de $L^{\wedge}(h)$ são positivos e se estes estiverem acima dos envelopes (superior e inferior), caracteriza-se nesta escala de distância como agrupamento, sendo estes mais fortes nas distâncias que possuem os extremos da curva. Observando os gráficos elaborados com distâncias máximas de 10000m e 1700m, respectivamente, ambos possuem agrupamento dos eventos, sendo que estes, aumentam quanto maior for a distância.