



**UNIVERSIDADE ESTADUAL PAULISTA**  
**“JÚLIO DE MESQUITA FILHO”**  
Câmpus de Presidente Prudente

CARLOS AUGUSTO SILVA DE PROENÇA

**ANÁLISE ESPACIAL E FATORES DE RISCO AMBIENTAL DA  
DENGUE NA ÁREA URBANA DE PRESIDENTE PRUDENTE – SP**

PRESIDENTE PRUDENTE

2023

**CARLOS AUGUSTO SILVA DE PROENÇA**

**ANÁLISE ESPACIAL E FATORES DE RISCO AMBIENTAL DA  
DENGUE NA ÁREA URBANA DE PRESIDENTE PRUDENTE – SP**

Relatório Final de Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Estatística da FCT/UNESP para aproveitamento na disciplina Trabalho de Conclusão de Curso.

Orientador: Prof. Dr. Edilson Ferreira Flores

PRESIDENTE PRUDENTE

2023

P964a Proença, Carlos Augusto Silva  
Análise espacial e fatores de risco ambiental da dengue na  
área urbana de Presidente Prudente - SP / Carlos Augusto  
Silva Proença. -- Presidente Prudente, 2023  
42 p. : il.

Trabalho de conclusão de curso (Bacharelado - Estatística) -  
Universidade Estadual Paulista (Unesp), Faculdade de  
Ciências e Tecnologia, Presidente Prudente  
Orientador: Edilson Ferreira Flores

1. Análise espacial (Estatística). 2. Dengue. 3. Funções de  
Kernel. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da  
Faculdade de Ciências e Tecnologia, Presidente Prudente. Dados fornecidos pelo  
autor(a).

Essa ficha não pode ser modificada.

## TERMO DE APROVAÇÃO

Carlos Augusto Silva de Proença

### ANÁLISE ESPACIAL E FATORES DE RISCO AMBIENTAL DA DENGUE NA ÁREA URBANA DE PRESIDENTE PRUDENTE – SP


Relatório de Final de Trabalho de Conclusão de Curso aprovado como requisito para obtenção de créditos na disciplina Trabalho de Conclusão do curso de graduação em Estatística da Faculdade de Ciências e Tecnologia da Unesp, pela seguinte banca examinadora:

Orientador:



---

Prof. Dr. Edilson Ferreira Flores  
Departamento de Estatística



---

Prof. Dr. Fernando Antonio Moala  
Departamento de Estatística



---

Prof. Dr. Silvely Nogueira De Almeida Salomão Néia  
Departamento de Estatística

Presidente Prudente, 26 de Janeiro de 2023.

## AGRADECIMENTOS

Queria agradecer primeiramente ao Professor Edilson pelo aprendizado, orientação e por todo apoio que ele me proporcionou no desenvolvimento desse trabalho. Assim como a banca composta pelo Professor Fernando Moala e pela professora Silvely Salomão por todas ótimas sugestões e dedicação para que o trabalho pudesse evoluir.

A minha família por ter dado todo o suporte que precisei em toda minha vida, e por sempre apoiar meus sonhos.

Aos meus amigos do curso por toda ajuda nessa parte da minha vida e por toda troca de conhecimento e apoio.

Aos amigos pessoais que tive o imenso contato, e que me auxiliaram nesses 4 anos com todas as dificuldades presentes.

E por fim a minha namorada Stefany que sempre me apoiou em todos os momentos e sempre me incentivou a buscar evoluir cada vez mais.

## RESUMO

A dengue é um dos principais problemas de saúde pública do mundo, é uma infecção viral sistêmica transmitida entre humanos pelo mosquito *Aedes* que está presente em diversos continentes. O município de Presidente Prudente não é diferente e tem grande receptividade para a disseminação do arbovírus, a cidade é uma região de grande ocupação populacional de baixo poder aquisitivo, portanto, estudar como os casos da doença se distribuem através do município, possibilita a criação de projetos de ação para que atinjam melhoras na saúde pública. O principal vetor da dengue é o mosquito *Aedes aegypti*, que se procria em focos de água parada. Visto isso, o presente trabalho estudou a distribuição dos casos da dengue e correlacionou-os com vetores encontrados em bocas de lobo pela cidade. Os dados utilizados na análise foram cedidos pelo Serviço de Zoonose e Vigilância Epidemiológica da Secretária Municipal de Saúde e os dados sobre as bocas de lobo serão cedidos pela Mestranda em Geografia Elaine Bertaco. Assim, a análise espacial utilizou os estimadores de Kernel para encontrar as áreas com maior incidência da doença e de bocas de lobo com água parada e com presença de larvas, e ao correlacionar as camadas foi possível encontrar uma correlação de 57% entre os casos de dengue e bocas de lobo com presença de água. Contudo, foi observado que não é possível afirmar que as bocas de lobo com presença de água são os principais focos da doença, mas foi possível elucidar essa relação de modo que essa possa ser uma possível causa no aumento de casos da doença e, portanto, precisa de atenção os órgãos responsáveis.

**Palavras-chave:** 1. Análise espacial (Estatística). 2. Dengue. 3. Funções de Kernel.

## ABSTRACT

Dengue is one of the main public health problems in the world, it is a systemic viral infection transmitted between humans by the Aedes mosquito that is present on several continents. The municipality of Presidente Prudente is no different and is highly receptive to the spread of the arbovirus, the city is a region of great population occupation with low purchasing power, therefore, studying how the cases of the disease are distributed throughout the municipality, makes it possible to create action projects so that improvements in public health are taken. The main vector of dengue is the Aedes aegypti mosquito, which breeds in stagnant water. Given this, the present work studied the distribution of dengue cases and correlated them with vectors found in storm drains around the city. The data used in the analysis were provided by the Zoonosis and Epidemiological Surveillance Service of the Municipal Secretary of Health and data on culverts will be provided by Master's student in Geography Elaine Bertaco. Thus, the spatial analysis used Kernel estimators to find the areas with the highest incidence of the disease and culverts with standing water and the presence of larvae, and by correlating the layers it was possible to find a correlation of 57% between the cases of dengue fever and culverts with the presence of water. However, it was observed that it is not possible to state that the culverts with the presence of water are the main outbreaks of the disease, but it was possible to elucidate this relationship so that this could be a possible cause in the increase in cases of the disease and, therefore, needs attention from the responsible agency.

**Keywords:** Spatial analysis. Kernel. Dengue. Culverts.

## LISTA DE FIGURAS

Figura 1 - Mapa de Londres com óbitos por cólera identificados em vermelho e poços de água em azul.....	11
Figura 2– Distribuição espacial dos casos de infecção por ZIKA Vírus em Maceió -AL 2016. ....	12
Figura 3- À esquerda a superfície do teor de argila e à direita a variância de krigagem. ....	13
Figura 4 - Representação densidade demográfica, separada por cada estado nos anos de 2000. ....	14
Figura 5 – Demonstração do cálculo do estimador Kernel.....	16
Figura 6 – Classificação do estimador do vizinho mais próximo. ....	18
Figura 7 – Saída gráfica do software ArcGis para função K.....	21
Figura 8 - Mapa de Presidente Prudente com casos positivos de Dengue. ....	24
Figura 9 - Mapa de Presidente Prudente com casos graves de Dengue identificados. ....	25
Figura 10 - Mapa de Presidente Prudente com bocas de lobo.....	26
Figura 11 - Mapa de Presidente Prudente com bocas de lobo com larvas identificados em laranja.....	27
Figura 12 - Mapa de Presidente Prudente com bocas de lobo com água.....	28
Figura 13 - Relatório de análise do estimador do vizinho mais próximo para os casos positivos de dengue. ....	29
Figura 14 - Relatório de análise do estimador do vizinho mais próximo para as bocas de lobo com larvas. ....	30
Figura 15 - Relatório de análise do estimador do vizinho mais próximo para as bocas de lobo com água.....	31
Figura 16 - Relatório de análise do estimador K para os casos positivos de dengue. ....	32
Figura 17 - Relatório de análise do estimador K para as bocas de lobo com larvas. ....	33
Figura 18 - Relatório de análise do estimador K para as bocas de lobo com água... ..	34
Figura 19 - Mapa de Kernel para os casos positivos de dengue na cidade de Presidente Prudente.....	35
Figura 20 - Mapa de Kernel para as bocas de lobo com presença de larvas na cidade de Presidente Prudente.....	36



Figura 21 - Mapa de Kernel para as bocas de lobo com água na cidade de Presidente Prudente.....	37
Figura 22 - Mapas de Kernel para as bocas de lobo com presença de larvas e casos positivos de dengue sobrepostos na cidade de Presidente Prudente. ....	38
Figura 23 - Mapas de Kernel para as bocas de lobo com água e casos positivos de dengue sobrepostos na cidade de Presidente Prudente. ....	39

### **LISTA DE QUADROS**

Quadro 1 – Classificação dos casos de dengue. ....	22
Quadro 2 - Matriz de covariância entre dengue e bocas com larvas.....	38
Quadro 3 - Matriz de covariância entre dengue e bocas com água. ....	39

# SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>8</b>
1.1 IDENTIFICAÇÃO SOBRE A DENGUE .....	8
OBJETIVO GERAL .....	9
OBJETIVOS ESPECÍFICOS .....	9
JUSTIFICATIVA.....	9
<b>2 REFERENCIAL TEÓRICO.....</b>	<b>10</b>
2.1 ANÁLISE ESPACIAL.....	10
2.1 EVENTOS OU PADRÕES PONTUAIS.....	11
2.3 ÁREAS COM CONTAGEM OU TAXAS AGREGADAS .....	13
2.4 ESTIMADOR DE KERNEL.....	15
2.5 ESTIMADOR DO VIZINHO MAIS PRÓXIMO .....	18
2.6 FUNÇÃO K .....	19
2.7 SISTEMAS DE INFORMAÇÃO GEOGRÁFICA .....	21
<b>3 METODOLOGIA.....</b>	<b>22</b>
<b>4. RESULTADOS .....</b>	<b>24</b>
<b>6. CONSIDERAÇÕES FINAIS .....</b>	<b>40</b>
<b>REFERÊNCIAS .....</b>	<b>41</b>

# 1 Introdução

## 1.1 Identificação sobre a dengue

De acordo com a Organização das Nações Unidas (ONU), a dengue é um dos principais problemas de saúde pública no mundo que de acordo com a Secretaria de Estado de Saúde, é uma doença febril aguda com amplo espectro de manifestações clínicas (BHATT, *et al.*, 2013).

A dengue é uma infecção viral sistêmica transmitida entre humanos pelo mosquito *Aedes* e a transmissão está presente em todos os trópicos, sendo as zonas de maior risco as Américas e Ásia. O principal vetor da dengue é o mosquito *Aedes aegypti*, que tem um papel fundamental na saúde pública desde o início do século XX. (ZARA, *et al.*, 2016). E hoje é o arbovírus mais importante do mundo em termos de morbidade, mortalidade e danos econômicos (Guzmán *et al.*, 2004; Kindhauser, 2003).

O município de Presidente Prudente/SP tem grande receptividade para a disseminação dos arbovírus, manifestada pela infestação contínua e sustentada dos vetores, principalmente o *Aedes aegypti*, como também grande vulnerabilidade associada a fatores facilitadores de grande extensão territorial, intenso fluxo de pessoas, alta densidade populacional e hábitos de vida. A transmissão da dengue é observada no município desde 1996, ano em que foram notificados os primeiros casos autóctones em um bairro da cidade. Desde então houve ampliação das áreas de transmissão e o aumento de casos notificados sendo que nos anos de 2016 e 2019 vivenciaram duas importantes epidemias de dengue, registrando 13.065 e 7.749 casos respectivamente.

Período no qual houve um crescimento significativo da cidade e de sua população, observa-se que tratando de uma cidade com potencial financeiro expressivo da região oeste do estado de São Paulo, pode-se constatar a entrada e saída diária de centenas de pessoas que moram em cidades no entorno, mas que tem seus empregos em Presidente Prudente/SP. A doença em questão causa não somente prejuízo à saúde como também financeira, tratando-se de uma massa economicamente ativa.

Assim, o presente trabalho buscou estudar os casos de dengue na cidade de Presidente Prudente e correlacioná-los com os dados encontrados nas bocas de lobo da cidade.

### **Objetivo Geral**

O principal objetivo deste trabalho é gerar a distribuição espacial da dengue, utilizando técnicas de estatística espacial, e correlacionar com os vetores encontrados nas bocas de lobo, no município de Presidente Prudente - SP.

### **Objetivos Específicos**

Os objetivos específicos a serem alcançados são:

- Adquirir melhor conhecimento na área de Análise Espacial;
- Mapear os casos positivos de Dengue na cidade de Presidente Prudente – SP, nos anos de 2021;
- Gerar os mapas de Kernel para os casos positivos de Dengue na cidade de Presidente Prudente - SP.
- Gerar os mapas das Bocas de lobo utilizando-se de dados dos vetores encontrados no ano de 2019 e buscar correlações com os casos de Dengue.

### **Justificativa**

A região de Presidente Prudente é uma região de grande ocupação populacional e habitacional de baixo poder aquisitivo, onde encontram-se presente muitas construções precárias e carentes. Com o passar de anos de trabalho desenvolvidos, em conjunto com a secretaria de saúde, assistencial social e limpeza pública, os índices permaneceram elevados, de modo que é necessária uma atenção especializada e estudo na área descrita sob seus problemas e dificuldades de resolução dos casos ativos e constantes conforme dados oficiais.

Desse modo, estudar como se distribuem os casos e encontrar possíveis correlações, pode ajudar no combate à doença e possibilitará conhecer quais áreas

possuem focos recorrentes, quais fatores externos influenciam nas formações desses focos e possibilitará criar previsões de como a doença estará distribuída nos próximos anos, tornando possível criar modelos preditivos que auxiliariam na tomada de decisões e na elaboração de ações preventivas para que seja possível diminuir o número de casos e a fatalidade da doença.

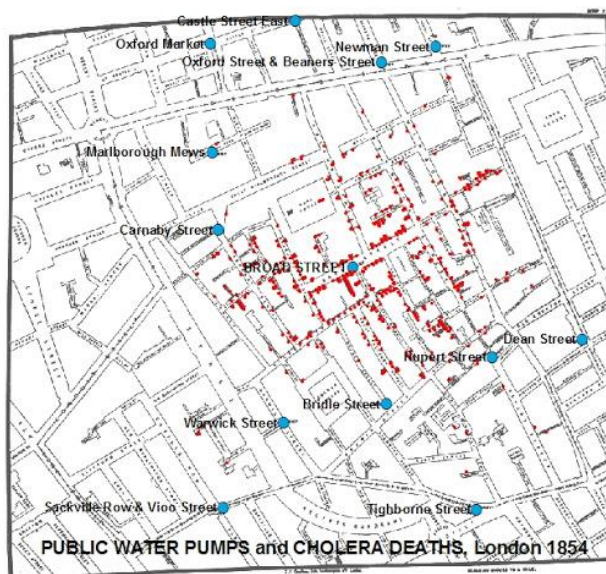
## **2 Referencial Teórico**

### **2.1 Análise Espacial**

A análise espacial pode ser descrita como a ciência que lida com a descrição, identificação e classificação sistemática de um fenômeno estudado onde seu principal objetivo é compreender a distribuição espacial dos dados oriundos desses fenômenos. E embora a análise espacial tenha surgido por volta do século XIX com a análise de John Snow, que foi a precursora da análise espacial que conhecemos hoje, com os avanços nos Sistemas de Informações Geográficas (SIG's) que além de possuírem um baixo custo, contam com interfaces amigáveis que facilitam a execução das análises espaciais cada vez mais.

Conhecido como pioneiro das análises espaciais, o médico John Snow no século XIX realizou um estudo a respeito da epidemia de cólera que estava ocorrendo em Londres em 1854, a associando ao consumo de água insalubre. O médico, com o auxílio de um mapa representado na Figura 1, criou pontos na residência de indivíduos que vieram a óbito por conta da doença e relacionou com as bombas de água que abasteciam a cidade, com esta análise foi possível identificar que ao redor da bomba localizada em Broad Street havia um maior número de casos de cólera, sendo o epicentro da epidemia. Ao saberem dessas informações medidas foram tomadas para o fechamento dos poços e assim evitar a contaminação presente. (DRUCK, CARVALHO, CÂMARA, MONTEIRO, 2004).

Figura 1 - Mapa de Londres com óbitos por cólera identificados em vermelho e poços de água em azul.



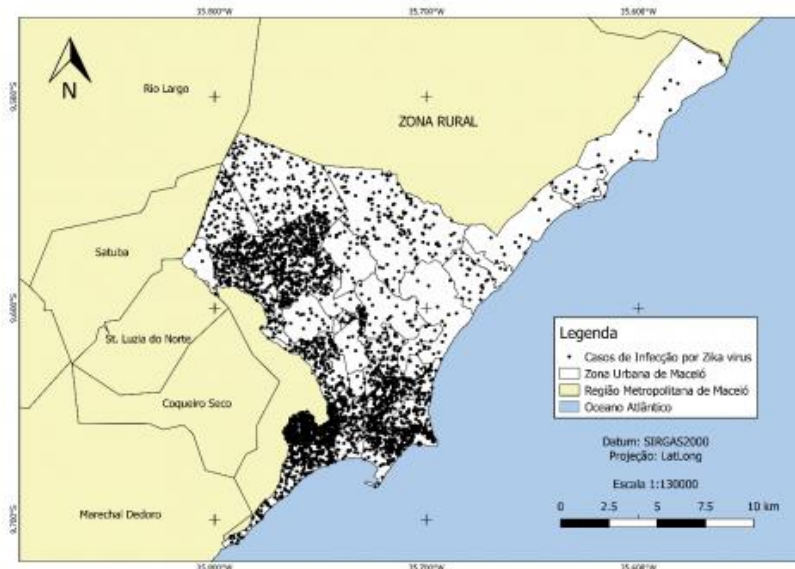
Fonte: Disponível em: <http://www.ufrj.br/institutos/it/de/acidentes/sr10.htm>

Assim, a partir desse pioneirismo de John que a análise espacial pode se desenvolver dando espaço para o surgimento de vários métodos de análises onde após estudos acabou se definindo que na análise espacial a taxonomia mais utilizada para caracterizar os problemas de análise espacial considera três tipos de dados: Eventos ou Padrões Pontuais, superfícies contínuas e áreas com contagens e taxas agregadas (DRUCK et. al., 2004). Cada uma utiliza diferentes métodos para análise dos dados, mas as três têm em comum a necessidade de que os dados sejam georreferenciados.

## 2.1 Eventos ou Padrões Pontuais

Os eventos pontuais são representados por pontos. Estes pontos são associados a ocorrência de fenômenos de interesse, onde as análises relacionadas a este tipo de fenômeno procuram identificar se a distribuição de ocorrências ocorre de forma aleatória ou se existe algum padrão. Um exemplo visível é representado na Figura 2 onde os pontos pretos indicam a localização de casos de Infecção por Zika vírus no município de Maceió – AL.

Figura 2– Distribuição espacial dos casos de infecção por ZIKA Vírus em Maceió -AL 2016.

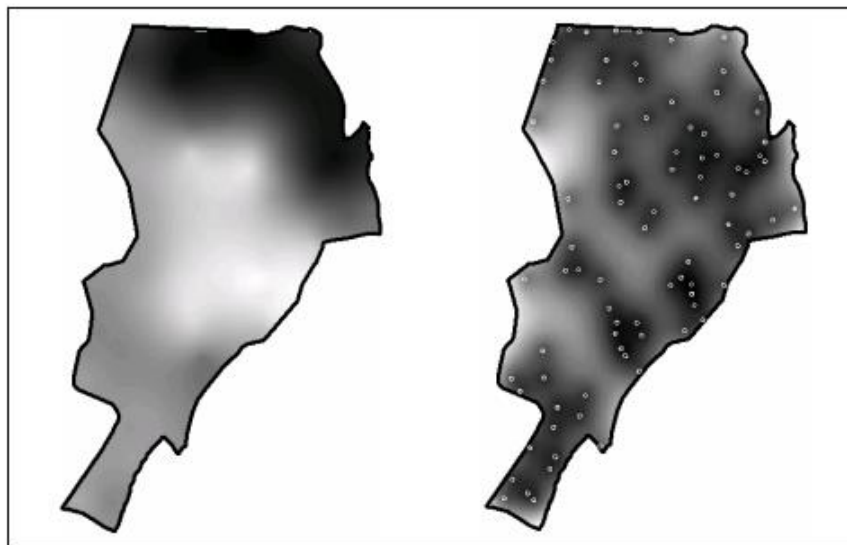


Fonte: SES, 2017. Elaborada por Almeida et al, 2019  
 Disponível em: <https://journals.openedition.org/confins/21976>

## 2.2 Superfícies contínuas

As superfícies contínuas são geradas a partir da interpolação de amostras de campo, podendo estar regular ou irregularmente distribuídas (DRUCK, CARVALHO, CÂMARA, MONTEIRO, 2004). Como geralmente pensamos nesses tipos de dados como resultado de investigações sobre recursos naturais, como temperatura, pressão atmosférica, características do solo etc. Procedimentos de interpolação usando procedimentos geoestatística, como krigagem, são frequentemente necessários. Na Figura 3, temos um exemplo de krigagem numa superfície contínua:

Figura 3- À esquerda a superfície do teor de argila e à direita a variância de krigagem.



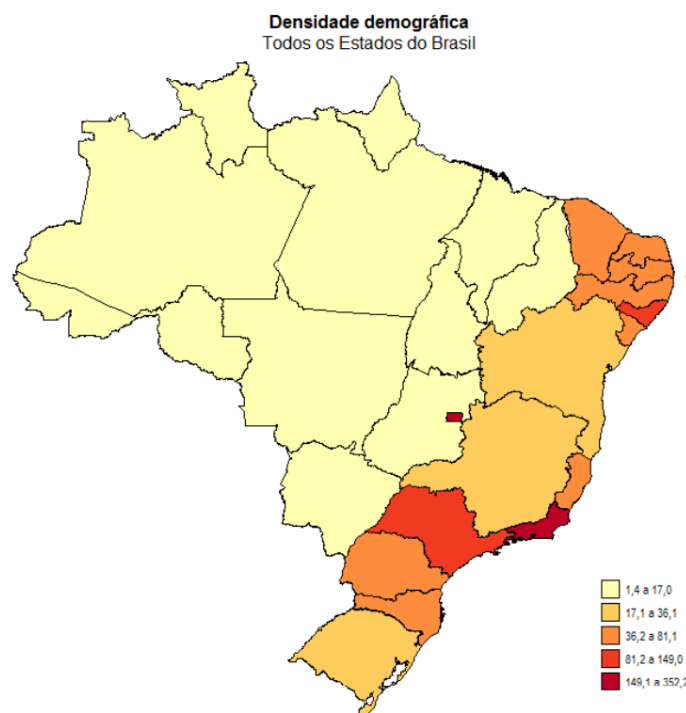
Fonte: Disponível em: <http://www.dpi.inpe.br/gilberto/tutoriais/analise/cap3.pdf> Cap.3 pag. 24.

### 2.3 Áreas com contagem ou taxas agregadas

As áreas com contagem ou taxas agregadas são resultantes de coletas de dados relacionadas a população. Sendo que os dados são agregados e apresentam algumas razões de confidencialidade, tornando assim a formação de áreas delimitadas por polígonos fechados como no mapa coroplético representado na Figura 4.



Figura 4 - Representação densidade demográfica, separada por cada estado nos anos de 2000.



Disponível em: <https://docplayer.com.br/54908600-Capitulo-3-mapas-coropleticos.html>

O presente trabalho tem foco principal na área de análise de processos pontuais, que consiste em fenômenos expressos através de ocorrências identificadas como pontos localizados uma área de interesse, podendo estar associados a atributos de identificação. Esses fenômenos são denominados como processos pontuais. Onde visamos determinar o seu comportamento e relação com outros fenômenos espaciais.

Numa visão, processos pontuais são definidos como um conjunto de pontos irregularmente distribuídos em um terreno, cuja localização foi gerada por um mecanismo estocástico. Para sua caracterização, este processo estocástico pode ser descrito em termos de efeitos de primeira ordem e de segunda ordem (CÂMARA, 2004, p. 57).

Os efeitos de primeira ordem correspondem a variações no valor médio do processo no espaço e são considerados globais ou de larga escala. Assim, o foco é a *intensidade* do processo, geralmente utiliza-se o estimador de *Kernel* para analisar o comportamento de padrões. Os efeitos de segunda ordem, que representam a *dependência espacial* no processo, são geralmente estimados por duas técnicas de

estimação, a do *vizinho mais próximo* para pequenas escalas e a da *função K* para escalas maiores.

## 2.4 Estimador de Kernel

Este estimador forma uma superfície onde o valor será proporcional à intensidade da amostra por unidade de área, e pondera os pontos pela distância até o local de interesse. Esta função conta todos os pontos dentro da área de influência e os pondera pela distância de cada ponto até o local de interesse. Isso é útil para gerar uma visão geral de como os dados são distribuídos. Os estimadores de kernel podem ser descritos estimando a densidade estatística espacial dos dados observados, primeiro analisando os efeitos de primeira ordem correspondentes a mudanças na média do processo no espaço.

Pode ser assimilado com as propriedades de densidade uni variada e alterará o conteúdo de pesquisa da área de estudo. Isso permite medir a distribuição de eventos e calcular a interação de um processo pontual não paramétrico por meio de uma função kernel, onde a entrada necessária serão as coordenadas do ponto.

Mede-se, assim, a distribuição dos eventos calculando a intensidade do processo,  $\lambda(s)$ .

A técnica de kernel consiste num estimador probabilístico de intensidade do processo pontual não-paramétrico através da função de kernel. Usando como entrada para a aplicação as ocorrências da variável (através de um sistema de coordenadas).

Assim, se  $s$  representa uma localização qualquer em  $R$  e  $s_1, s_2, \dots, s_n$  são as localizações dos  $n$  eventos observados, então o estimador de kernel é dado por:

$$\hat{\lambda}_\tau(s) = \sum_{i=1}^n \frac{1}{\tau^2} k\left(\frac{s-s_i}{\tau}\right), s - s_i \leq \tau \quad (1)$$

Onde tem-se que  $\tau$  é o raio de influência em torno de  $s$  e determina o grau de suavização e  $k(\ )$  é uma função kernel bivariada e simétrica chamada de função de interpolação. Assim, a técnica suaviza a superfície, calculando a densidade para cada região da área de estudo, permitindo a construção de uma superfície contínua de ocorrências das variáveis, inferindo para toda a área de estudo a variação espacial da variável.

Já para casos em que são agregados atributos para caracterizar o evento, vemos a utilização da seguinte fórmula:

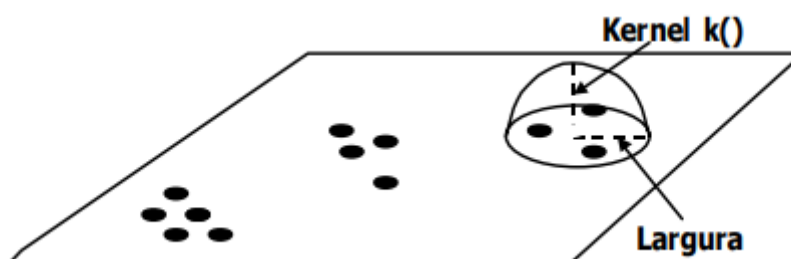
$$\hat{\lambda}_\tau(s) = \sum_{i=1}^n \frac{1}{\tau^2} k\left(\frac{d(u_i, u)}{\tau}\right) y_i, d(u_i, u) \leq \tau \quad (2)$$

Ao observar  $y_i$  dizemos que representa o  $i$ -ésimo atributo referente a  $u_i$ . Mas se o objetivo for calcular a estimativa da intensidade média, vê-se que é necessário a divisão da intensidade dos pontos pela intensidade sem o atributo, sendo assim podemos demonstrar a fórmula como:

$$\hat{\mu}_\tau(s) = \frac{\sum_{i=1}^n k\left(\frac{(s-s_i)}{\tau}\right) y_i}{\sum_{i=1}^n k\left(\frac{(s-s_i)}{\tau}\right)} \quad (3)$$

A partir da fórmula, pode-se dizer que  $(s - s_i)$  é a distância entre o ponto com a intensidade estimada e o ponto  $i$  dentro do raio de influência. Portanto, a definição da área depende da influência do raio centrado no ponto estimado, suavizando a superfície e calculando a densidade da área de estudo por sua vez, e finalmente analisando as possíveis tendências dos dados em escala global para suavizar o fenômeno, vários tipos de funções de interpolação, que podem ser descritas como:

Figura 5 – Demonstração do cálculo do estimador Kernel.



Disponível em: <http://www.dpi.inpe.br/gilberto/livro/analise/cap2-eventos.pdf>

A escolha da função de Kernel  $k(\cdot)$  a ser utilizada fica a critério do pesquisador, não existindo um ponto crítico para a escolha, já o raio de influência  $\tau$  deve ser definido com cautela, pois pode alterar as estimativas final (MINISTÉRIO DA SAÚDE, 2007).

- Kernel Gaussiano ou normal, pondera os pontos dentro do círculo de

forma que os pontos mais próximos possuem um peso maior comparado com os pontos mais afastados:

$$k(h) = \frac{1}{2\pi\sigma} e^{-\frac{h^2}{2\sigma^2}}, \text{ onde } \sigma = \tau \quad (4)$$

- Kernel triangular, dá um maior peso aos pontos próximos do que os pontos distantes dentro do círculo, mas o decréscimo é mais rápido:

$$k(h) = (1 - |h|) \quad (5)$$

- Kernel exponencial negativo, pondera os pontos próximos com peso muito mais intenso do que os pontos distantes:

$$k(h) = \frac{1}{2\pi} e^{-h} \quad (6)$$

- Kernel uniforme, pondera todos os pontos dentro do círculo igualmente:

$$k(h) = \frac{1}{2} \quad (7)$$

- Kernel quártico, pondera com maior peso os pontos mais próximos do que os distantes com um decréscimo gradual:

$$k\left(\frac{s-s_i}{\tau}\right) = \frac{3}{\pi} \left(1 - \left(\frac{s-s_i}{\tau}\right)^2\right)^2, \text{ para } \left(\frac{s-s_i}{\tau}\right)^2 \leq 1 \quad (8)$$

Assim, podemos escrever a intensidade  $\lambda$  desta forma,

$$\hat{\lambda}_\tau = \sum_{i=1}^n \frac{3}{\pi\tau^2} \left(1 - \frac{h_i^2}{\tau^2}\right)^2 \quad (9)$$

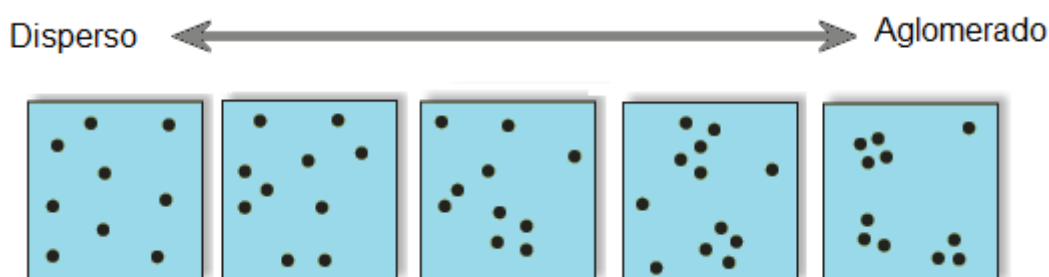
Onde vemos que  $h_i$  é a distância entre a localização em que se deseja calcular a função e o evento observado, e  $\frac{3}{\pi\tau^2}$  é a suavização da função sendo que a distância tende a zero.

A escolha do raio  $\tau$  é delicada e há alguns estudos que sugerem aproximações. Como por exemplo  $\tau = 0.68n^{-0.2}$  para estimar a intensidade quando a região  $R$  for um quadrado unitário e  $n$  for número de eventos observados em  $R$ . Esta sugestão pode ser estendida a regiões de estudo de outros formatos (Bailey & Gatrell, 1995).

## 2.5 Estimador do vizinho mais próximo

O estimador do vizinho mais próximo mede a distância entre cada centroide de recurso e a localização do centroide do vizinho mais próximo. Em seguida, calcula a média de todas as distâncias para os vizinhos mais próximos. Se a distância média for menor que a média da distribuição aleatória assumida, a distribuição das feições analisadas é considerada agrupada. Se a distância média for maior que a distribuição aleatória assumida, o recurso é considerado disperso.

Figura 6 – Classificação do estimador do vizinho mais próximo.



Disponível em: <https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-statistics/average-nearest-neighbor.htm#:~:text=The%20Nearest%20Neighbor%20Index%20is,in%20a%20hypothetical%20random%20distribution.>

A razão média do vizinho mais próximo é calculada como a distância média observada dividida pela distância média esperada (a distância média esperada é baseada em uma distribuição aleatória assumida, com o mesmo número de feições cobrindo a mesma área total). Esta função pode ser usada para verificar a iteração entre eventos.

Dessa forma, a razão média do vizinho mais próximo é calculada por:

$$ANN = \frac{\bar{D}_o}{\bar{D}_e} \quad (10)$$

onde  $\bar{D}_o$  é a distância média observada de cada característica e seus vizinhos mais próximos.

$$\bar{D}_o = \frac{\sum_{i=1}^n d_i}{n} \quad (11)$$

De modo que  $d_i$  seria a distância de cada ponto com para seu vizinho mais próximo.  $n$  seria o número de elementos da característica apresentada, e ainda  $\bar{D}_e$  é a distância média esperada para as características dadas no padrão aleatório.

$$\bar{D}_e = \frac{0,5}{\sqrt{n/A}} \quad (12)$$

tal que  $A$  seria a área mínima onde contenha nossos vizinhos. E a estatística  $Z$  do vizinho mais próximo é calculado por:

$$Z = \frac{\bar{D}_o - \bar{D}_e}{S_e} \quad (13)$$

e  $S_e$  é dado por

$$S_e = \frac{0.26136}{\sqrt{n^2/A}} \quad (14)$$

## 2.6 Função K

A Função K de Ripley (caso uni variado) utiliza a contagem e a distância para considerar a variância de todas as distâncias entre os eventos, diferente do vizinho mais próximo que considera somente a distância entre o evento mais próximo.

Assim, a função K estima o número esperado de eventos dentro de um círculo de raio  $h$  de  $m$  evento arbitrário dividido por  $\lambda$  que representa a intensidade do evento na área.

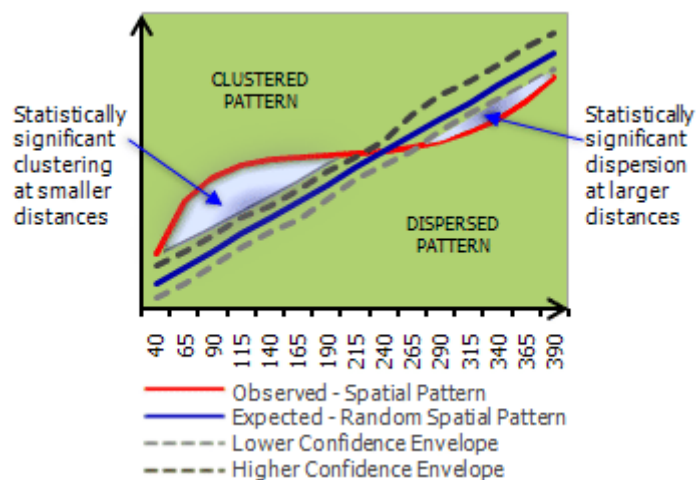
Há variações da função K de Ripley original. Neste trabalho utilizaremos uma transformação comum muitas vezes referida como  $L(d)$  onde o valor esperado de K será igual a distância dada como:

$$\hat{L}(d) = \sqrt{\frac{A \sum_i^n \sum_{j, j \neq i}^n k_{i,j}}{\pi n(n-1)}} \quad (15)$$

Tal que  $d$  é a distância,  $n$  é o número total de eventos observados,  $A$  representa a área da região e  $k_{i,j}$  é o peso. Se não houver uma correção devido ao efeito de borda o peso será igual a um quando a distância entre  $i$  e  $j$  for menor que  $d$  e zero caso contrário.

Já o gráfico de  $\hat{L}(h)$  em relação a distância  $h$ , temos que quando o valor observado de K é maior que o valor esperado de K, podemos afirmar que a distribuição dos dados é mais agrupada que uma distribuição aleatória utilizando aquela distância. Quando o valor observado de K é menor que o valor esperado de K, a distribuição é mais dispersa que uma distribuição aleatória. Temos ainda que quando o valor K observado é maior que intervalo de confiança superior o agrupamento espacial para a distância analisada é estatisticamente significativa. E da mesma forma quando o valor observado é menor que o intervalo de confiança inferior a dispersão espacial é estatisticamente significativa. Como é apresentado na Figura 7:

Figura 7 – Saída gráfica do software ArcGis para função K.



Fonte: Software ArcGis 10.3 Disponível em: <https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/GUID-110520A9-402D-4C17-8486-A7EC0F827D83-web.png>

## 2.7 Sistemas de Informação Geográfica

A ferramenta utilizada na análise espacial são os Sistemas de Informação Geográfica (SIG's), que segundo (Almeida Teixeira, A.L., 1992) são constituídos por uma série de programas e processos de análise, e cuja característica principal é focalizar o relacionamento de determinado fenômeno da realidade com sua localização espacial. Além disso, possuem uma ampla gama de aplicações dentro de diversas áreas como por exemplo:

- Ferramenta para produção de mapas;
- Suporte para análise espacial de fenômenos;
- Banco de dados geográficos, com funções de armazenamento e recuperação de informação espacial.

Porém, para que o fenômeno estudado possa ser representado computacionalmente é necessário fazer algumas transformações nas informações obtidas para que seu uso seja viável nos SIG's.

Assim, a análise espacial é um ramo da estatística que estuda padrões de distribuição de eventos no espaço, planejamento, coleta, tratamento, visualização e



análise dos dados geográficos, com a finalidade escolher um modelo que considere relacionamentos espaciais existentes dos fenômenos.

Nesse estudo serão analisados dados a respeito dos casos de dengue, assim como dados dos vetores encontrados nas bocas de lobo de Presidente Prudente. Onde a partir da análise estatística espacial podem ser avaliadas a variação espacial de taxas de incidência, o que possibilita identificar focos recorrentes da doença, realizar modelos preditivos e orientar ações para evitar a disseminação da doença (LAWSON,2006).

### 3 Metodologia

O objeto de estudo nesse projeto é analisar como se distribuem os casos de Dengue no município de Presidente Prudente - SP e correlacionar estes casos com os vetores obtidos com os dados das bocas de lobo da cidade. Os dados a respeito dos casos de dengue foram cedidos pelo Serviço de Zoonose e Vigilância Epidemiológica da Secretária Municipal de Saúde e os dados sobre as bocas de lobo serão cedidos pela Mestranda em Geografia Elaine Bertaco. A experimentação baseou-se na classificação de todos os pacientes que constatavam a doença enumerado por 10, já para casos que eram suspeitas e não fossem casos de dengue foram denotados por 5, e para casos de dengue com complicações eram separados por 11, e 12 eram casos de dengue grave (Quadro 1).

*Quadro 1 – Classificação dos casos de dengue.*

Para Classificação Final usar:
Negativos: 5
Positivos: 10
Dengue c/ complica- ções: 11
Dengue grave: 12

*Fonte: SUCEN de Presidente Prudente*

A análise espacial será feita a partir de *softwares* de Sistemas de Informação Geográfica (SIG's) que são uma importante ferramenta computacional utilizada na

análise espacial e serão utilizados neste trabalho. Eles irão auxiliar no armazenamento e no tratamento dos dados geográficos, foram utilizados os softwares *TerraView 3.0*, *ArcGIS 10.3*, *ArcView GIS 3.2* e algum outro *software* caso se torne necessário ao longo do desenvolvimento do trabalho.

Os mapas que serão utilizados serão obtidos através da Base do IBGE e da SUCEN, nesses mapas serão atribuídos os dados que foram coletados previamente, para que seja possível uma análise visual. Uma vez que os dados estiverem georreferenciados, serão utilizadas técnicas de análise exploratória, estatística espacial para que seja possível através da utilização das técnicas dos estimadores de Kernel, do vizinho mais próximo e da função k, a fim de evidenciar o comportamento de padrões dos pontos e sua provável causa e assim avaliar se as condições sociais, demográficas e ambientais influenciam nos casos e se sim, como funciona essa influência.

A análise terá foco nas áreas de análise pontual, onde inicialmente foi feito o mapeamento da localização dos casos de Dengue e das bocas de lobo, buscando caracterizar também seus atributos. Posteriormente será aplicado as técnicas de estatística espacial para fins de localizar e analisar os possíveis pontos críticos.

E por fim a será utilizada a ferramenta Band Collection Statistics do ArcGIS 10.3 que fornece estatísticas para análise multivariada de um conjunto de camadas *raster*. Que retorna uma matriz de correlação entre as camadas.

A equação para calcular a correlação é a seguinte:

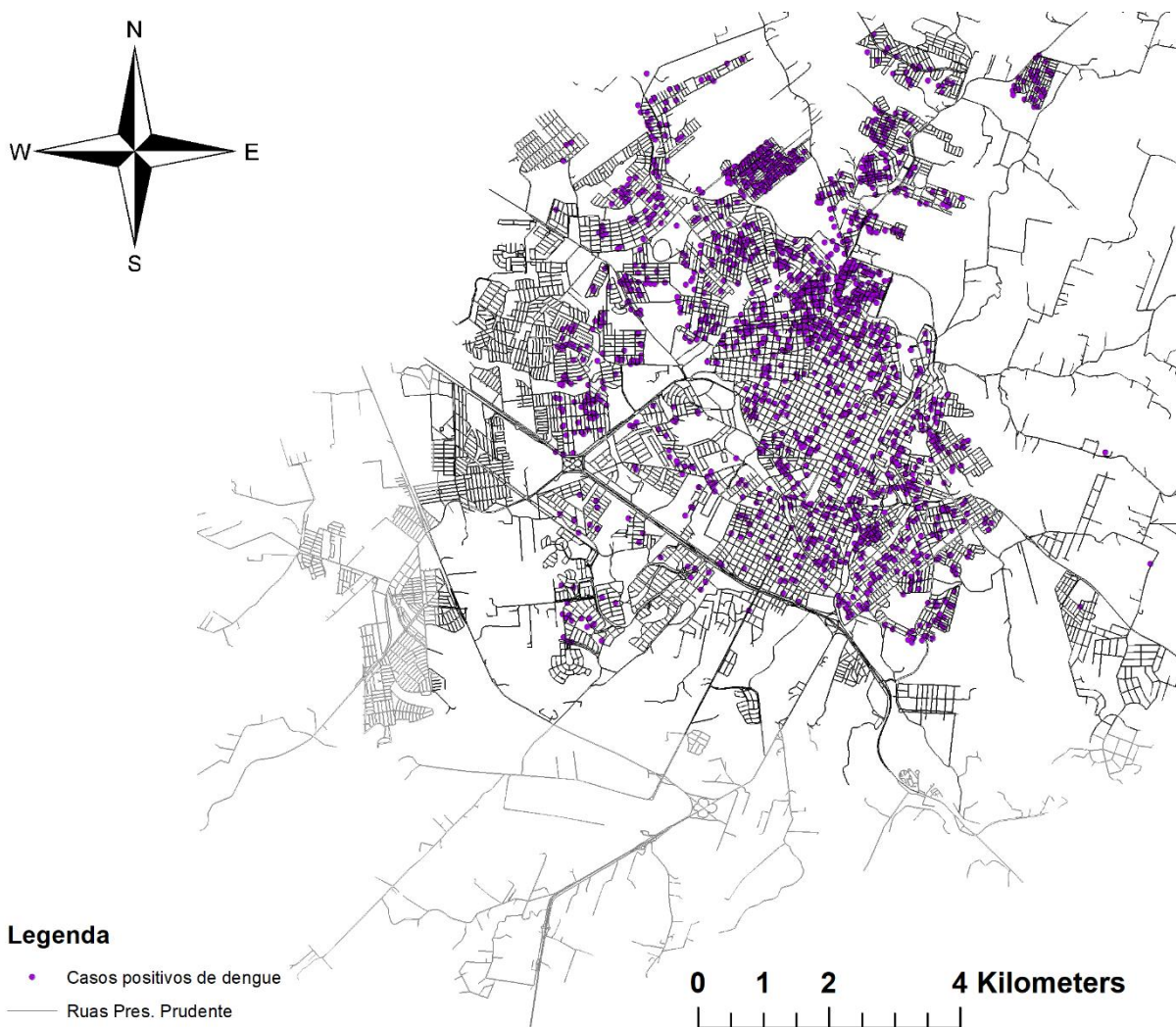
$$Corr_{ij} = \frac{Cov_{ij}}{\sigma_i \sigma_j} \quad (15)$$

A correlação varia de -1 a 1. Tal que uma correlação positiva e próxima de 1, indica uma relação direta entre duas camadas, ou seja, quando os valores celulares de uma camada aumentam, os valores celulares da outra camada também aumentam. Uma correlação próxima de zero significa que as duas camadas são independentes uma à outra, e quando o valor para a correlação está próximo de -1, temos uma correlação negativo, onde conforme os valores de uma camada aumenta, na outra camada esses valores tendem a baixar.

## 4. Resultados

No software, *ArcGIS 10.3*, foi feita a análise espacial dos casos de dengue positivos e das bocas de lobo. Os resultados são apresentados nas figuras 8 a 23.

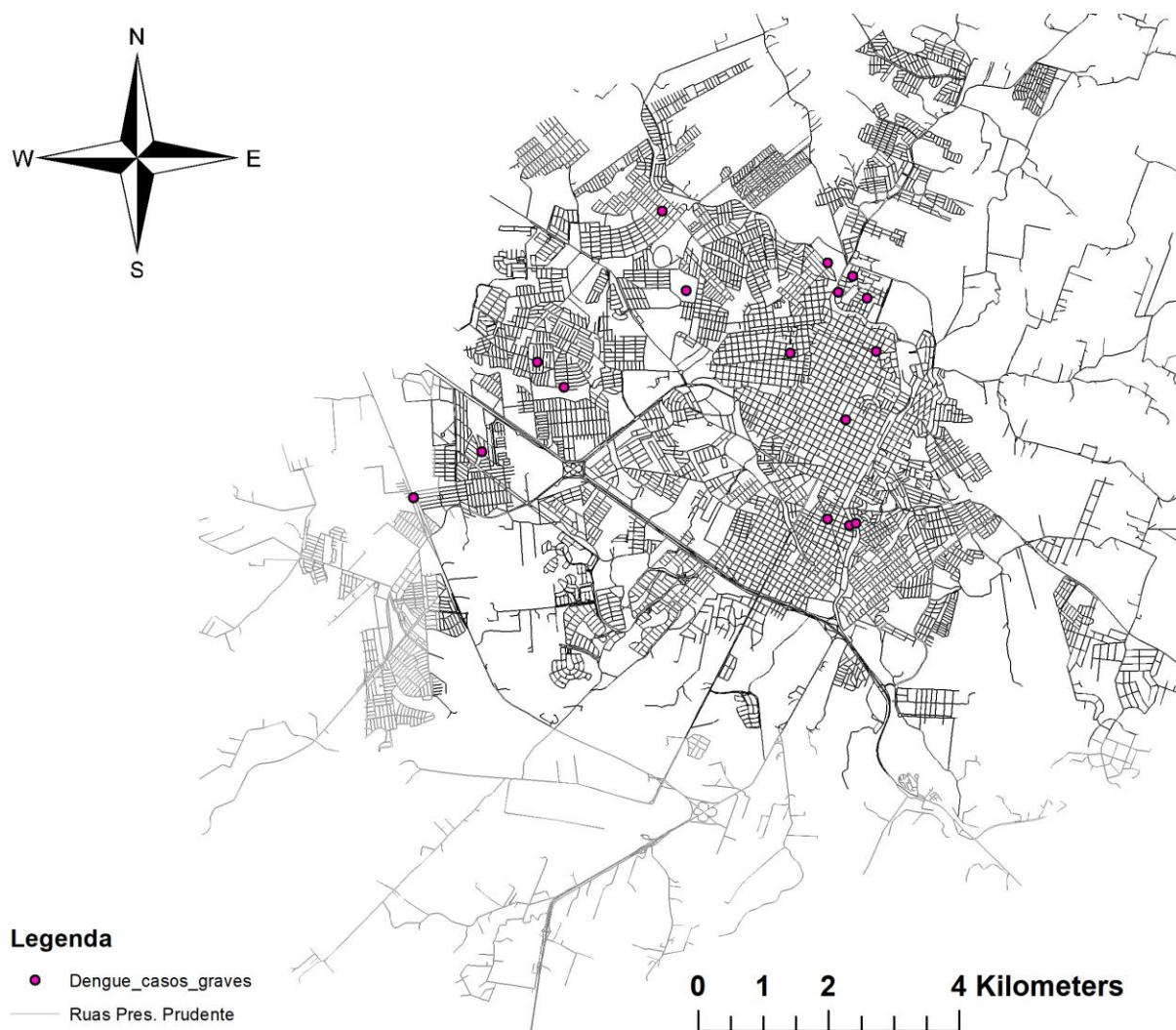
*Figura 8 - Mapa de Presidente Prudente com casos positivos de Dengue.*



Fonte: Autor

Analisando a Figura 8 é válido lembrar que os casos de dengue estudados estão limitados na área em que as bocas de lobo foram estudadas, dito isso, é possível observar que considerando a área estudada pode-se encontrar casos de dengue por toda a cidade, com alguns locais com maiores concentrações que outros.

Figura 9 - Mapa de Presidente Prudente com casos graves de Dengue identificados.

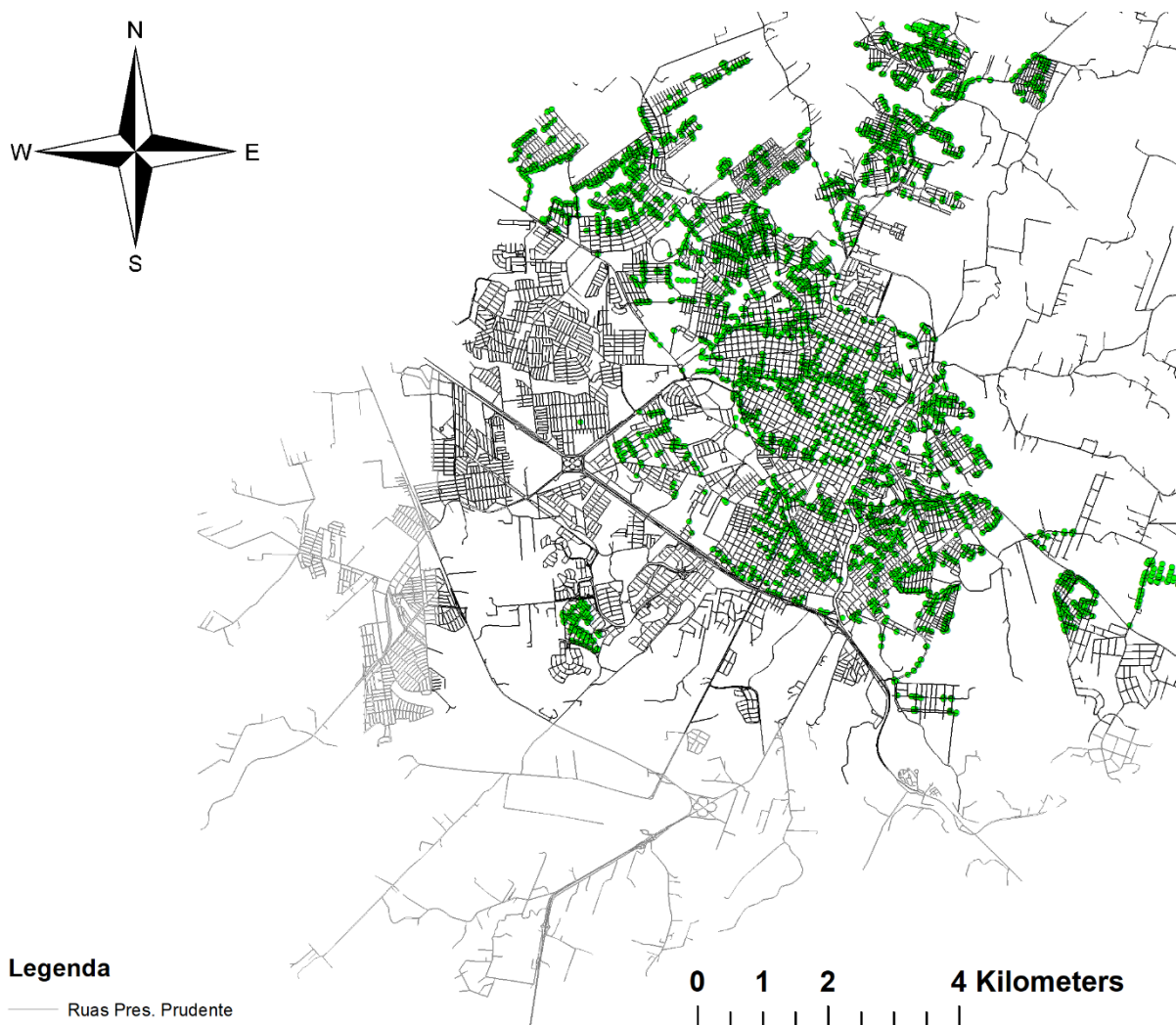


Fonte: Autor

Analisando a figura 9 podemos perceber que casos graves da doença são um evento raro, portanto trabalhar com todos casos positivos da doença parece uma decisão mais sensata.

Quando olhamos para as bocas de lobo, dois atributos se destacam, bocas com larvas (que não são necessariamente larvas da dengue) e bocas com presença de água (não necessariamente água parada), portanto, esses serão os vetores estudados nesse trabalho.

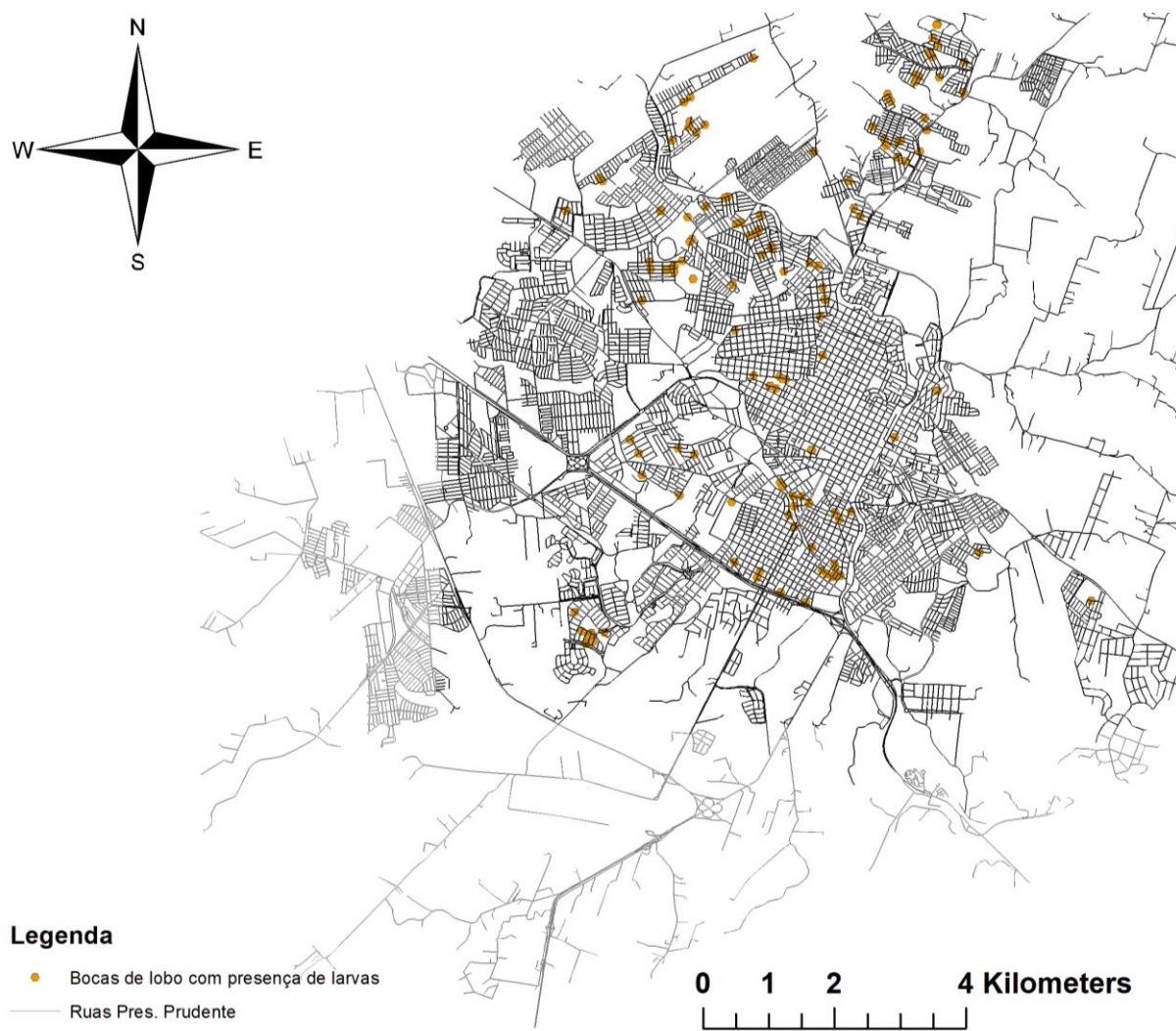
Figura 10 - Mapa de Presidente Prudente com bocas de lobo.



Fonte: Autor

As bocas de lobo se distribuem uniformemente pela cidade, visto que estão em todas as ruas, algumas regiões como o lado oeste da cidade não tiveram os dados coletados portanto o estudo se limitará a não estudar essas partes da cidade.

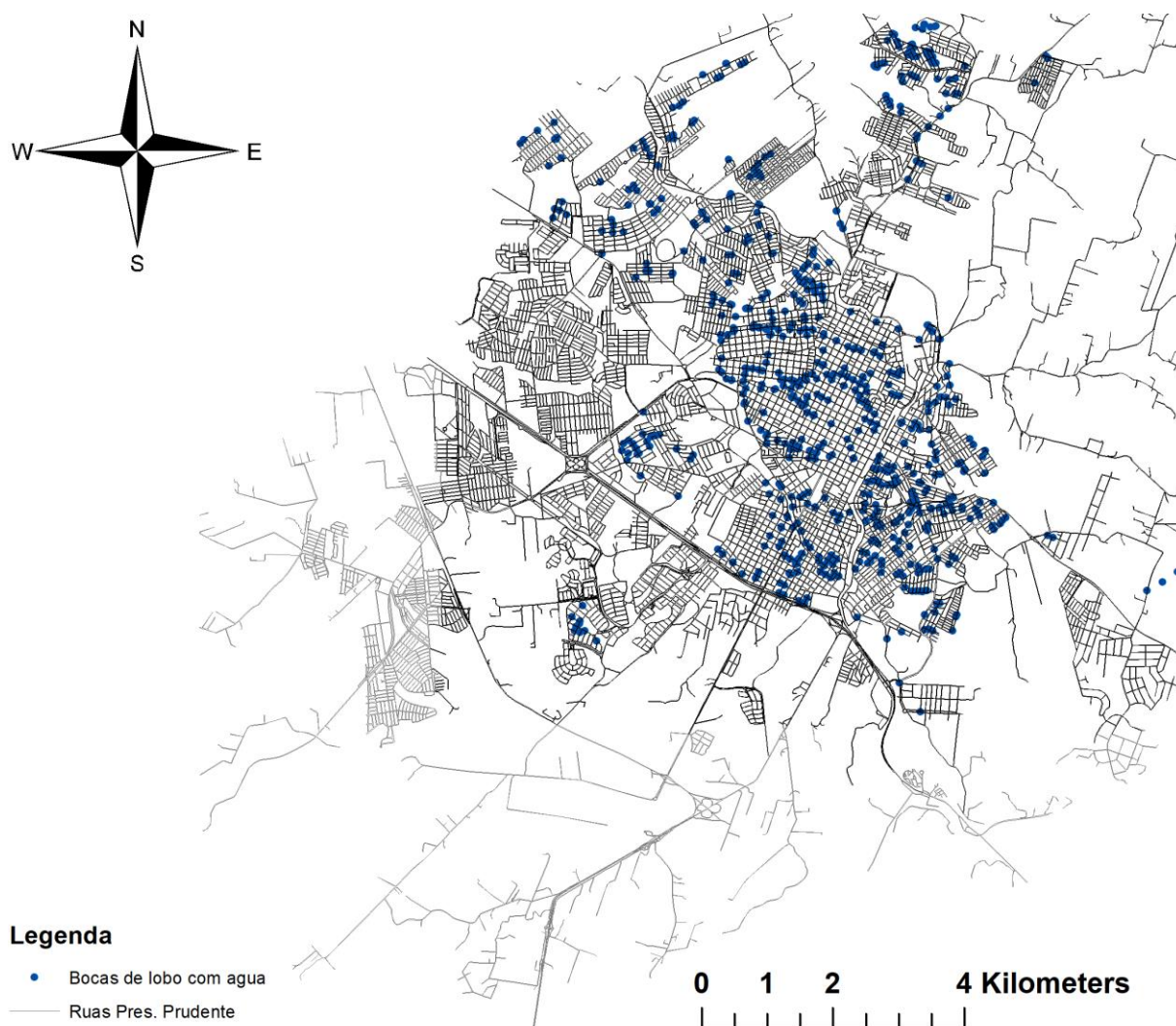
Figura 11 - Mapa de Presidente Prudente com bocas de lobo com larvas identificados em laranja.



Fonte: Autor

A taxa de bocas de lobo com presença de larvas é relativamente baixa, porém ainda tem potencial para trazer algum entendimento da causa dos casos de dengue.

Figura 12 - Mapa de Presidente Prudente com bocas de lobo com água.



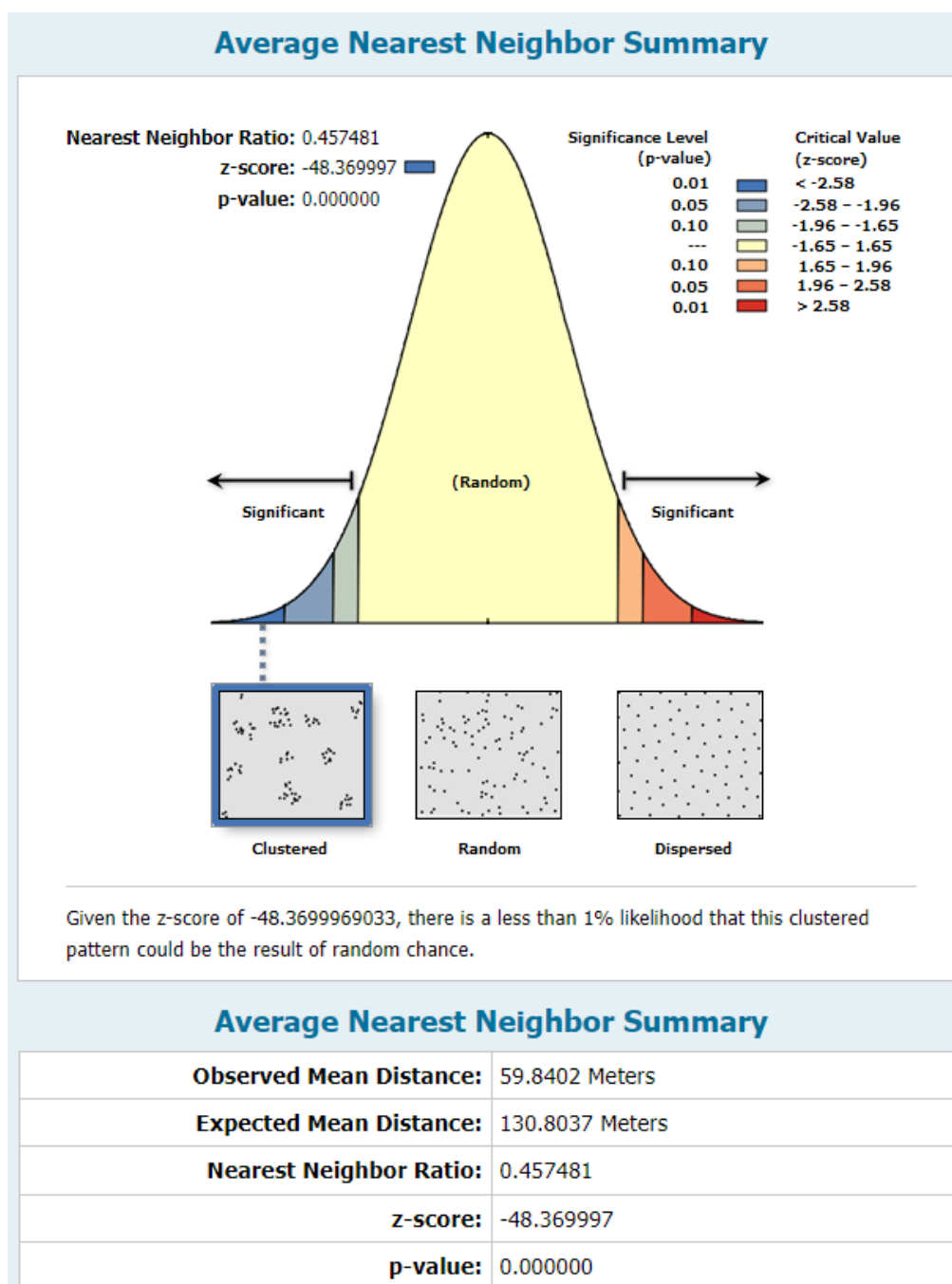
Fonte: Autor

Ao observar as bocas de lobo com presença de água, é possível observar sua presença em praticamente toda área analisada, porém com alguns locais com mais incidências que outros. Padrão também encontrado quando olhamos pra como se distribuem os casos positivos.

O objetivo, portanto, é entender se os dados estão agrupados, e caso estejam, como se distribuem esses agrupamentos dentro da cidade para que seja possível entender se há uma possível correlação entre as incidências.

Para saber se os dados estão agrupados, utilizamos o estimador do vizinho mais próximo:

Figura 13 - Relatório de análise do estimador do vizinho mais próximo para os casos positivos de dengue.

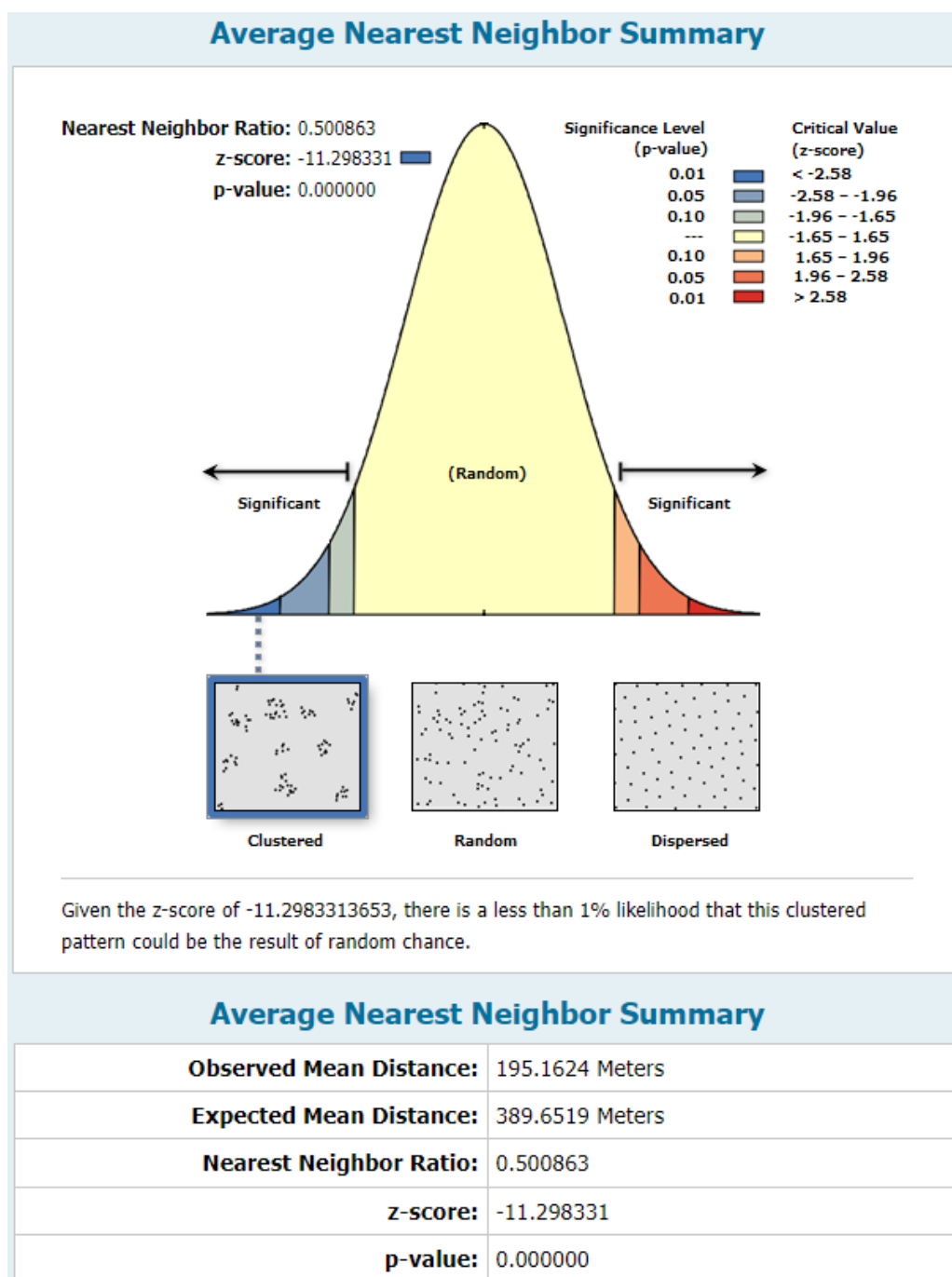


Fonte: Autor

A partir do relatório, podemos observar que os casos de dengue seguem um padrão agrupado, onde há menos de 1% de probabilidade de que esse padrão agrupado seja resultado de uma chance aleatória.



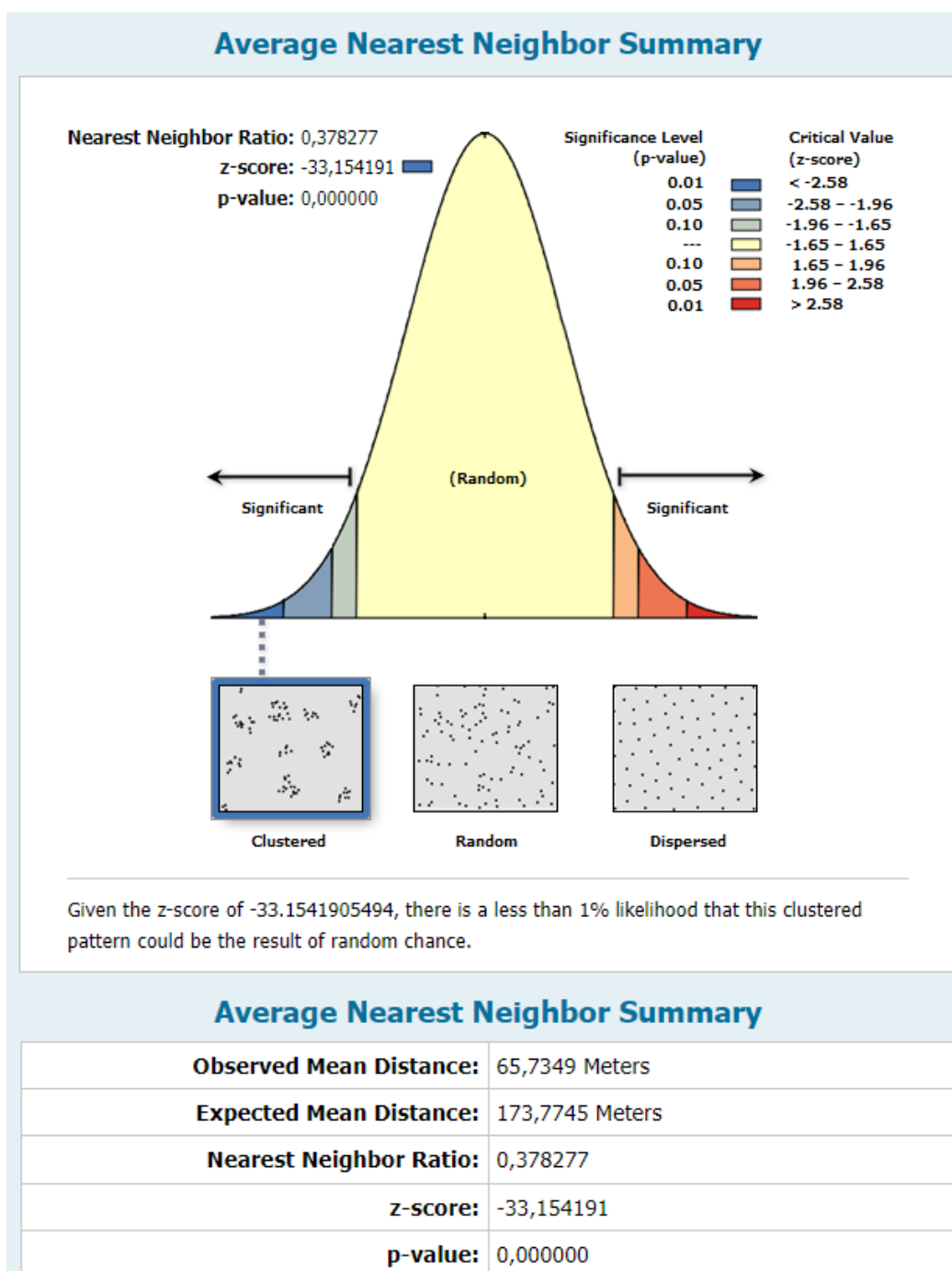
Figura 14 - Relatório de análise do estimador do vizinho mais próximo para as bocas de lobo com larvas.



Fonte: Autor

A partir do relatório, podemos observar que as bocas de lobo com presença de larvas seguem um padrão agrupado, onde há menos de 1% de probabilidade de que esse padrão agrupado seja resultado de uma chance aleatória.

Figura 15 - Relatório de análise do estimador do vizinho mais próximo para as bocas de lobo com água.

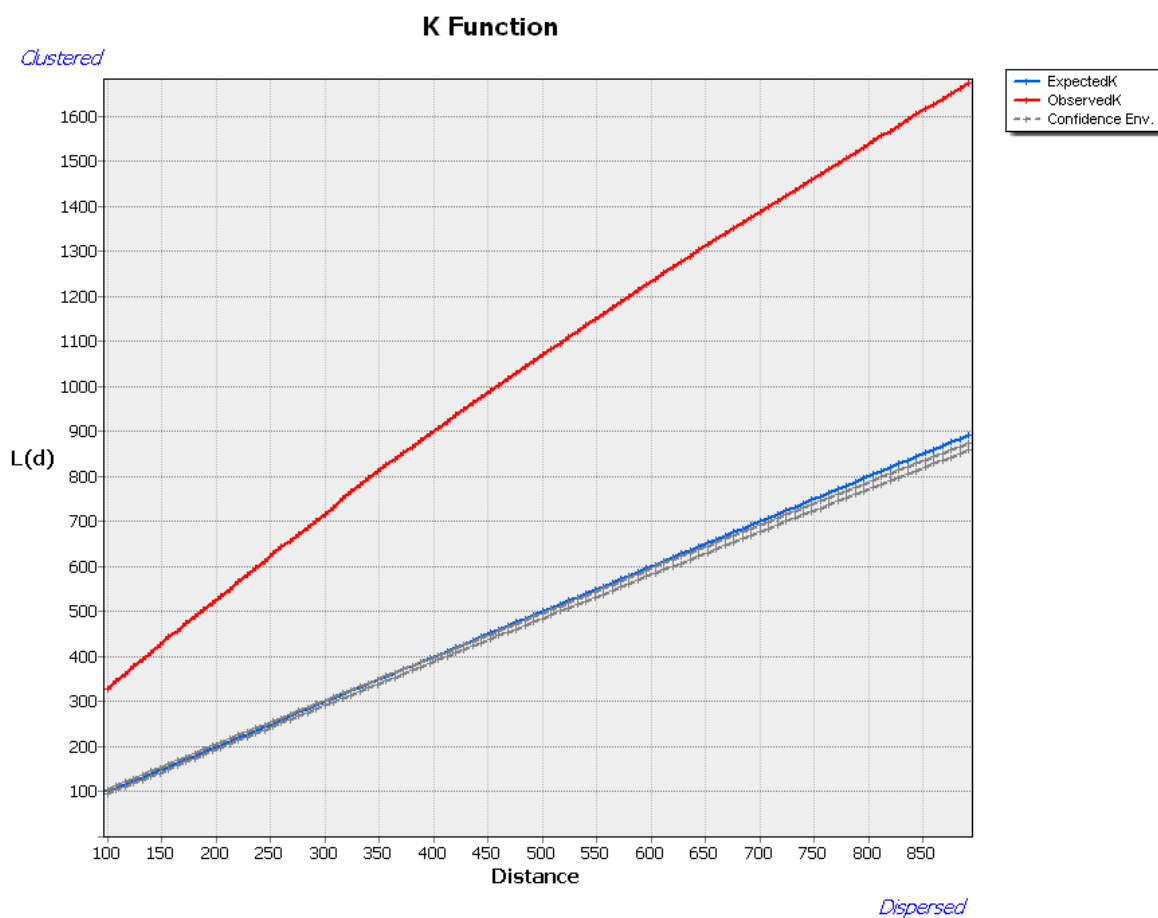


Fonte: Autor

Analisando as figuras 13,14 e 15 temos que nos 3 casos analisando o Z-Score e o p-valor, concluímos com 99% de confiança estatística que há menos de 01% de probabilidade de que esses dados agrupados, sejam resultados de uma chance aleatória.

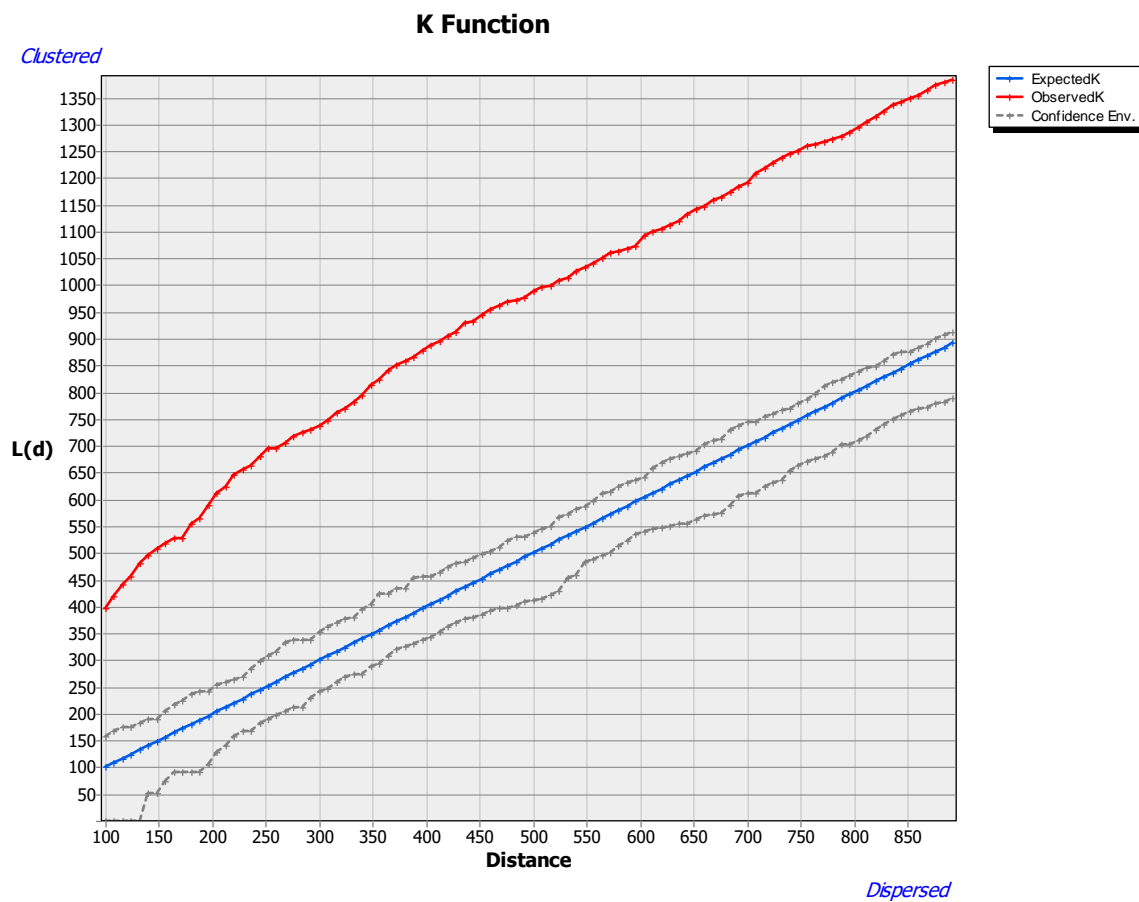
Como através do Estimador do vizinho mais próximo concluímos que os dados estão agrupados. O próximo passo é saber a qual distancia esses dados melhor se grupam e para saber qual a melhor distancia para utilizar na análise, para isso utilizaremos o estimador K.

Figura 16 - Relatório de análise do estimador K para os casos positivos de dengue.



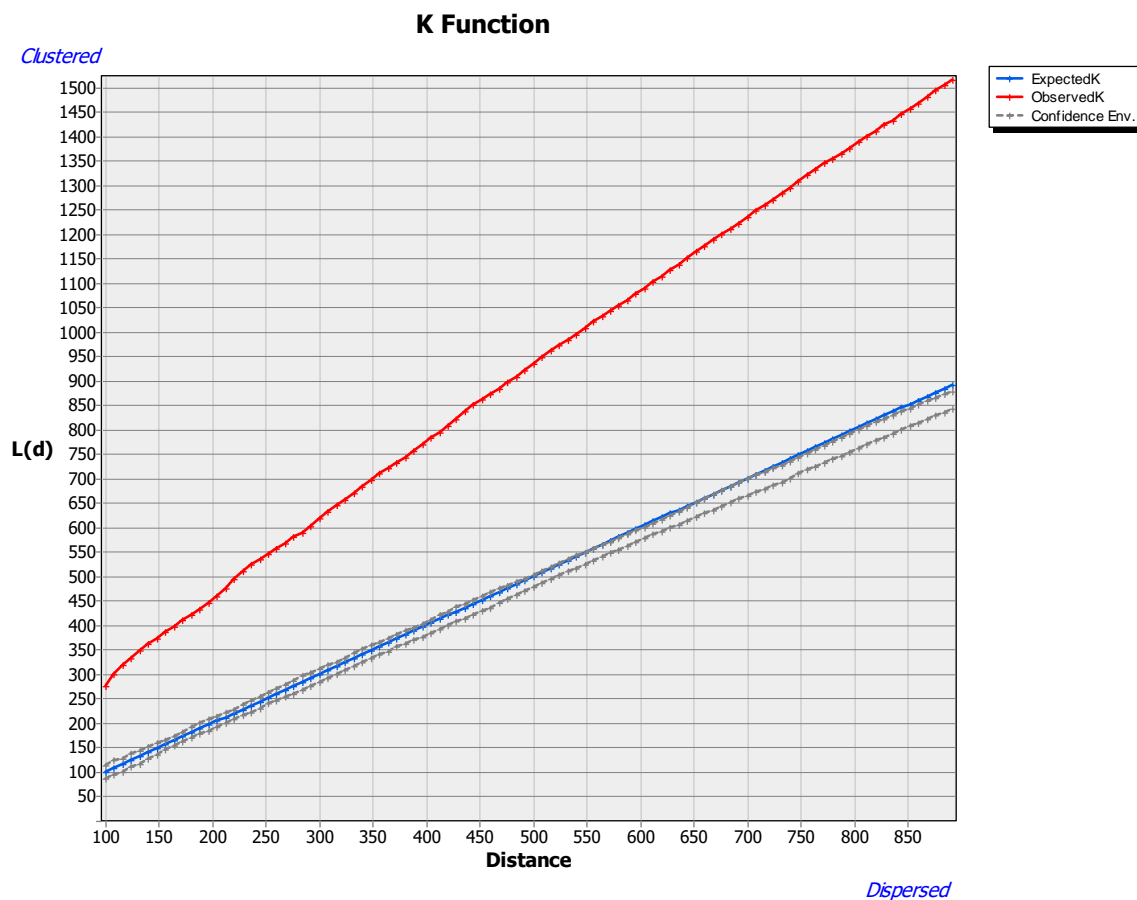
Fonte: Autor

Figura 17 - Relatório de análise do estimador K para as bocas de lobo com larvas.



Fonte: Autor

Figura 18 - Relatório de análise do estimador K para as bocas de lobo com água.

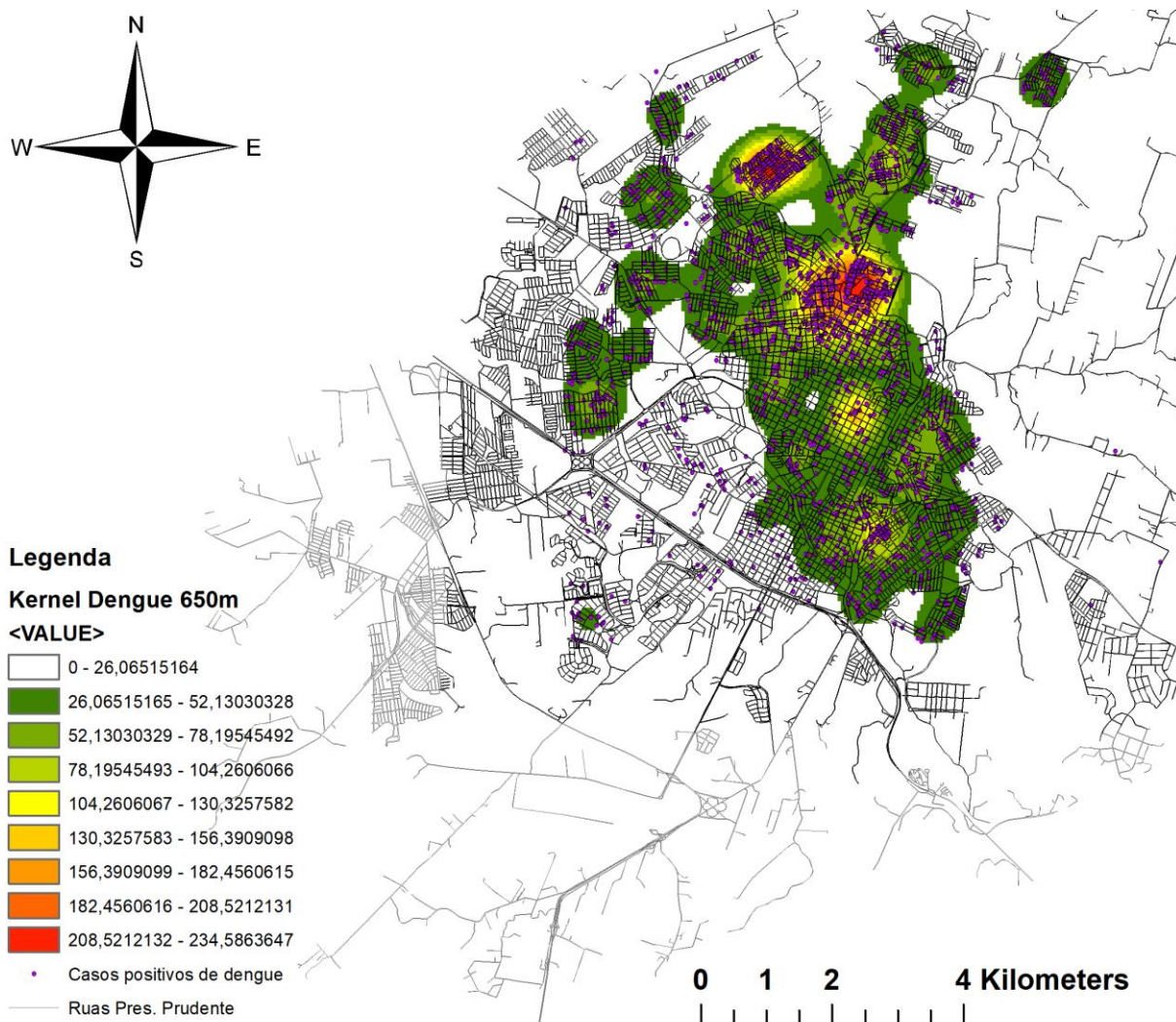


Fonte: Autor

Como podemos observar nas figuras 16, 17 e 18, para as distâncias de 100 a 900 metros os valores observados da estatística K sempre se encontram acima dos valores esperados de K e do intervalo de confiança superior, portanto podemos afirmar que nessa faixa de metros de distância os dados são agrupados.

Portanto foi definido um raio de estudo de 650 metros para os casos de dengue e para as bocas de lobo obtemos os seguintes mapas de Kernel

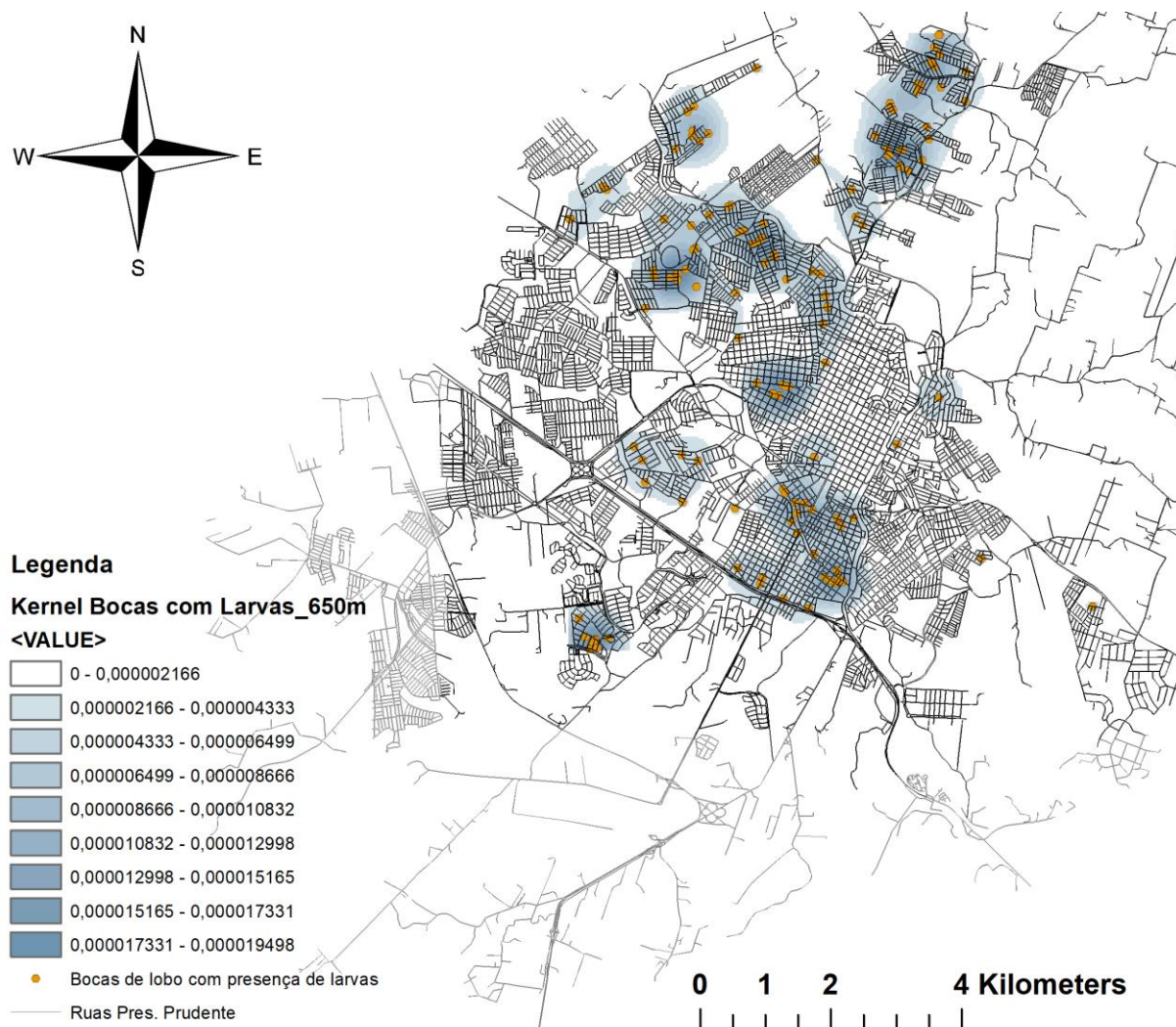
Figura 19 - Mapa de Kernel para os casos positivos de dengue na cidade de Presidente Prudente.



Fonte: Autor

Analisando o mapa de kernel podemos ver 2 áreas principais com grandes focos da doença nos bairros jardim guanabara e jardim Humberto Salvador. O que indica serem dois locais que portanto merecem receber uma atenção especial.

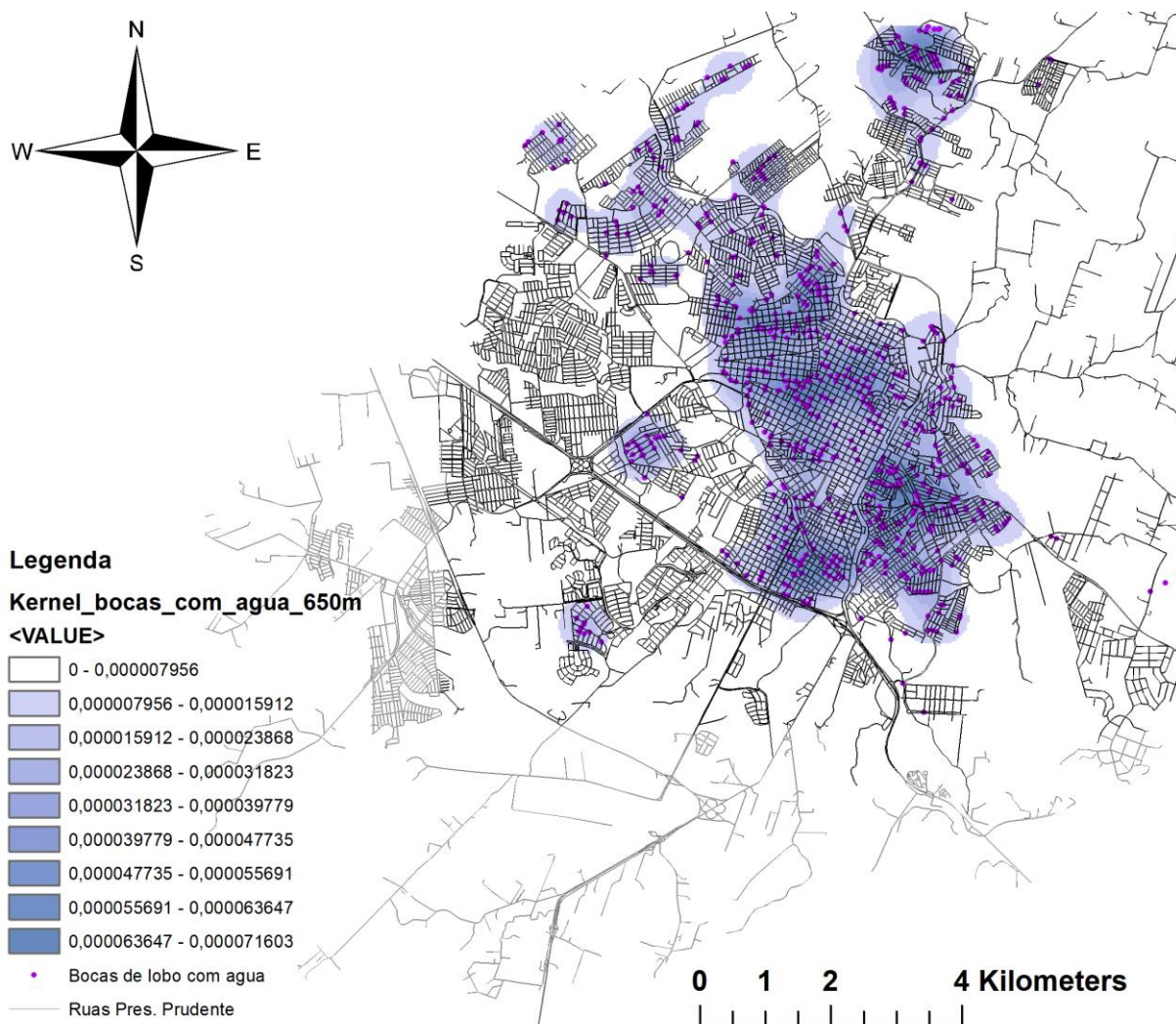
Figura 20 - Mapa de Kernel para as bocas de lobo com presença de larvas na cidade de Presidente Prudente.



Fonte: Autor

Não percebemos a presença de bocas de lobo com larvas nas duas principais zonas com casos de dengue, porém é possível notar alguns focos em seu entorno. Ou seja, a presença de larvas em bocas de lobo não parece estar diretamente relacionada com os casos locais da doença.

Figura 21 - Mapa de Kernel para as bocas de lobo com água na cidade de Presidente Prudente.

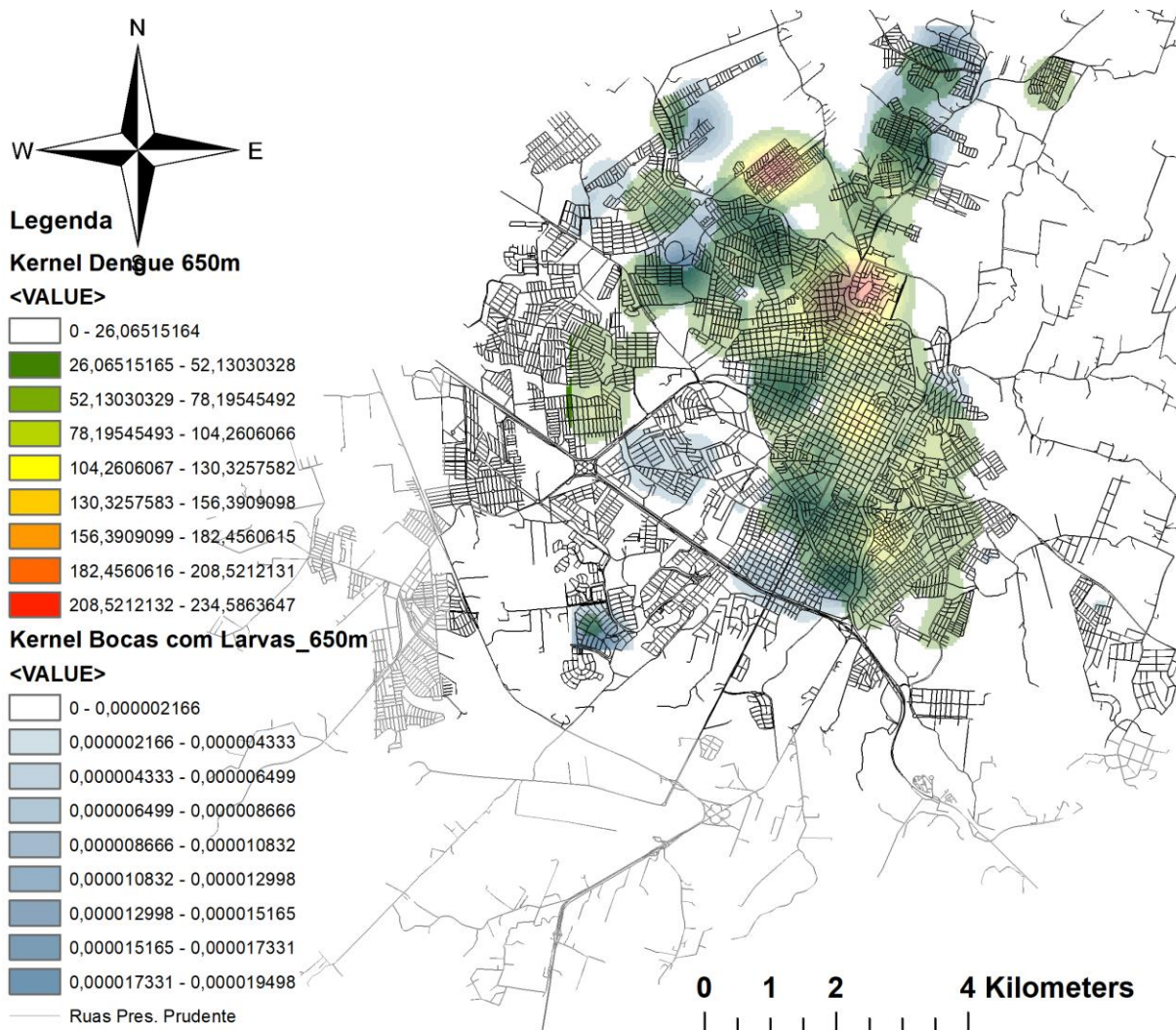


Fonte: Autor

Analisando as bocas de lobo com presença de água é possível notar que no jardim guanabara, um dos principais focos de dengue na cidade há uma grande concentração de bocas de lobo com presença de água. Porém outras regiões com muita concentração como o centro da cidade, não possuem tantos casos de dengue



Figura 22 - Mapas de Kernel para as bocas de lobo com presença de larvas e casos positivos de dengue sobrepostos na cidade de Presidente Prudente.



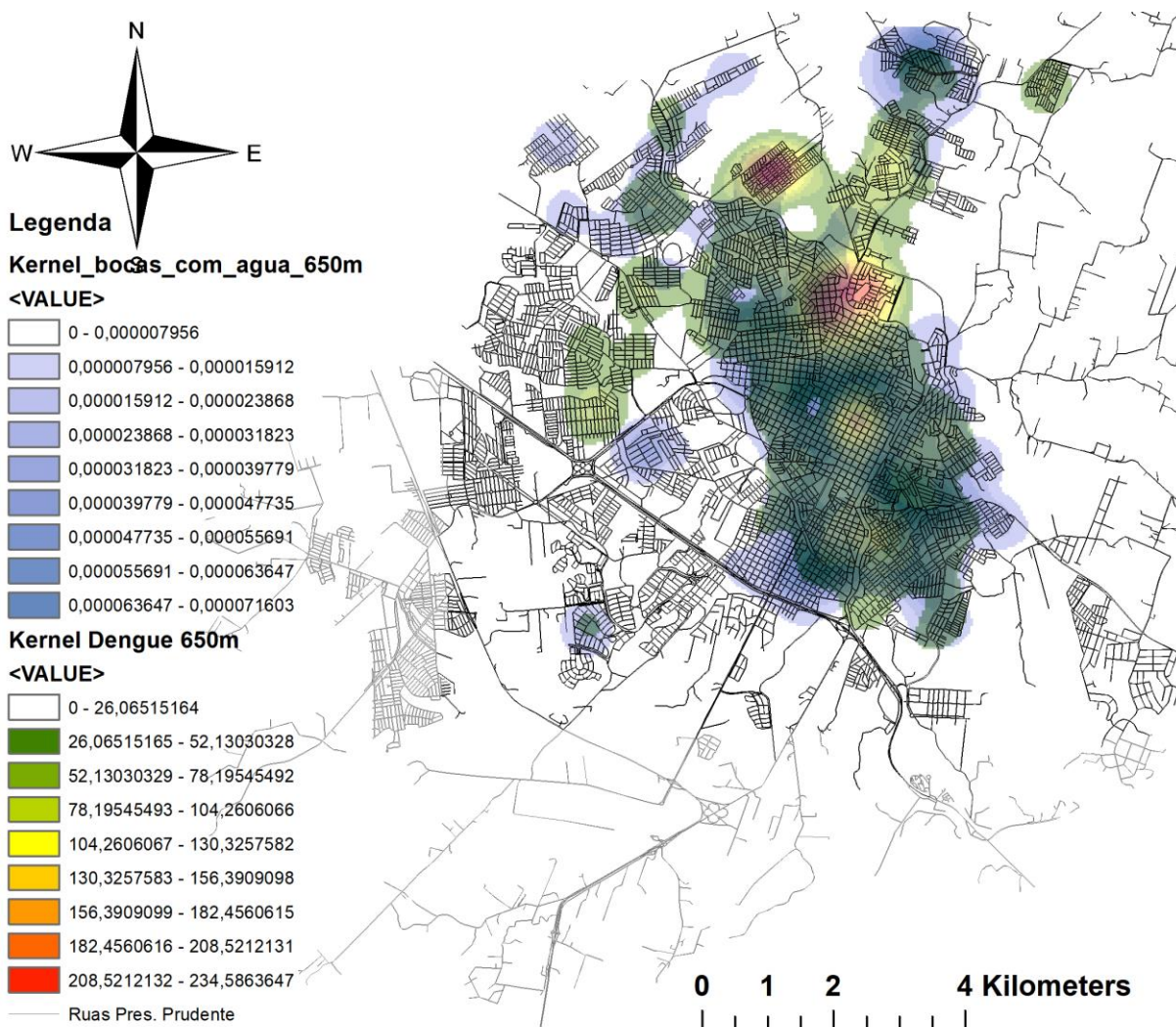
Fonte: Autor

Quadro 2 - Matriz de covariância entre dengue e bocas com larvas.

	Casos de dengue	Bocas com larvas
Casos de dengue	1,00000	0,38694
Bocas com larvas	0,38694	1,00000

Fonte: Autor

Figura 23 - Mapas de Kernel para as bocas de lobo com água e casos positivos de dengue sobrepostos na cidade de Presidente Prudente.



Fonte: Autor

Quadro 3 - Matriz de covariância entre dengue e bocas com água.

	Casos de dengue	Bocas com água
Casos de dengue	1,00000	0,57642
Bocas com água	0,57642	1,00000

Fonte: Autor

A partir do Quadro 2 vemos que não temos uma relação grande entre os casos de dengue e as bocas de lobo com presença de larvas visto que a correlação entre as duas camadas é de 0.38694, um valor relativamente baixo, e graficamente nas figuras 19,20 e 22 é possível confirmar essa falta de relação entre as camadas visto que as

duas zonas com maior concentração de caso de dengue não possuem bocas de lobo com presença de larvas.

Já no Quadro 3, quando analisamos os casos de dengue e as bocas de lobo com presença de água temos uma correlação maior, 0.57642, esse valor nos dá o indício de uma possível correlação entre as camadas, que pode ser confirmada graficamente ao observar as figuras 19,21 e 23, onde fica claro que, diferente dos casos de bocas de lobo com larvas, nos locais com focos de dengue também é possível se encontrar bocas de lobo com água e ainda é possível encontrar bocas de lobo com presença de água em grandes quantidades no entorno desses focos principais de dengue.

## **6. Considerações finais**

O presente estudo mapeou os casos confirmados de dengue na cidade de Presidente Prudente no ano de 2021, mapeou os vetores encontrados nas bocas de lobo no ano de 2019 e gerou os mapas de Kernel para ambos os dados. Também foi possível calcular a correlação entre eles para entender como os casos de dengue e as bocas de lobo se relacionam. Portanto, considerando as figuras e o valor das correlações foi possível concluir que as bocas de lobo com água não são a única forma dos mosquitos se reproduzirem, visto que é possível encontrar água parada de diversas outras formas na cidade, porém foi possível identificar uma relação entre ambos que traz à tona uma possível causa no aumento de casos da doença que pode ser combatida com o saneamento básico para que a saúde da população não seja colocada em risco.

Assim, conclui-se que a pesquisa contribuiu não só para o autor que pode obter uma maior compreensão da utilização de técnicas de análise espacial e do uso de softwares relacionados a área, como para a sociedade, pois os resultados alcançados tem grande importância, visto que, elucida um problema que deve ser resolvido pelos órgãos governamentais. Para as próximas pesquisas um tópico que vale a ressaltar, seria o uso de dados mais atualizados de forma que seja possível acompanhar o diagnóstico da situação atual.

## Referências

- ArcGIS**. Versão 10.3. Redlands, Califórnia. ESRI, 2009. Disponível em < <http://www.esri.com/software/arcgis/index.html/> >. Acesso em 15/07/2022.
- BAILEY, T. C.; GATRELL, A. C. **Interactive spatial data analysis**. New York: Longman, 1995.
- BHATT, Samir *et al.* **The global distribution and burden of dengue**. Nature, [S. l.], v. 496, n. 7446, 7 abr. 2013. DOI 10.1038/nature12060. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/23563266/>. Acesso em: 25 jan. 2022.
- BRASIL. Ministério da Saúde. **Introdução à Estatística Espacial para Saúde Pública**. 1 ed. Brasília: Editora do Ministério da Saúde, 2007.
- DRUCK, S., CARVALHO, M.S., CÂMARA, G. MONTEIRO, A.V.M. **Análise Espacial de Dados Geográficos**. Brasília: EMBRAPA, 2004.
- BAILEY, T. C.; GATRELL, A. C. **Interactive Spatial Data Analysis**. London: Longman, 1995. 413p.
- FERREIRA, L.M., SÁFADI, T., LIMA, R.R. **Técnicas da Estatística Espacial na Análise de Dados de Áreas no Estudo da Dengue**. Revista UNIVAP, São José dos Campos, v.24, n.44, p. 13-27, jul. 2018.
- FERREIRA, M. C. **Iniciação à análise geoespacial: teoria, técnicas e exemplos para geoprocessamento**. São Paulo: Ed. UNESP, 2014.
- GUZMÁN, M. G.; KOURÍ, G. **Dengue diagnosis, advances and challenges**. [S. l.] *Int. J. Infect. Dis.*, v.8, p.69-80, 2004.
- KINDHAUSER, M. K. **Dengue y fiebre hemorrágica dengue**. In: **Defensa Global ante la amenaza de Enfermedades Infecciosas** Genebra: Organização Mundial da Saúde, 2003. p.140-3.
- LAWSON, A.B. **Statistical Methods in Spatial Epidemiology**. Sussex: John Wiley & Sons, 2006.
- MARTÍNEZ TORRES, E. Dengue. **Estudos Avançados**, [S. l.], v. 22, n. 64, p. 33-52, 2008. Disponível em: <https://www.revistas.usp.br/eav/article/view/10347>. Acesso em: 17 jan. 2022.
- MATIAS, L.F. **Sistema de Informações Geográficas (SIG): Teoria e Método para Representação do Espaço Geográfico**. Tese (Pós-Graduação em Geografia

Humana) – Faculdade de Filosofia, Letras e Ciências Humanas. Universidade de São Paulo. São Paulo, 2001.

OSWALDO CRUZ, Fundação. Ministério da Saúde. **Introdução à Estatística Espacial para a Saúde Pública**. Brasília, 2007.

PAPLOSKI, I.A.D., RODRIGUES, M.S., MUGABE, V.A. *et al.* **Storm drains as larval development and adult resting sites for *Aedes aegypti* and *Aedes albopictus* in Salvador, Brazil**. *Parasites Vectors* **9**, 419 (2016). <https://doi.org/10.1186/s13071-016-1705-0>

TEIXEIRA, A.L.A., MORETTI, E., CHRISTOFOLETTI, A. **Introdução aos Sistemas de Informação Geográfica**. Rio Claro: Edição do autor, 1992.

ZARA, A. L. S. A.; et al. **Estratégias de controle do *Aedes aegypti*: uma revisão**. *Epidemiologia e Serviços de Saúde* [online]. 2016, v. 25, n. 2 [Acessado 25 de janeiro 2022], pp. 391-404. Disponível em: <https://doi.org/10.5123/S1679-49742016000200017>.