



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

**UM ESTUDO SOBRE A DISTRIBUIÇÃO
ESPAÇO-TEMPORAL DA DENGUE NA CIDADE DE
SÃO JOSÉ DOS CAMPOS USANDO SCAN STATISTICS**

Cátia Souza do Nascimento Sepetauskas

Trabalho de conclusão da disciplina SER 301 - Análise Espacial de Dados Geográficos.

RESUMO

A detecção de clusters espaço-temporais permite encontrar possíveis persistências espaço-temporais num conjunto de eventos. Neste trabalho, foi feito o estudo espaço-temporal de casos diários notificados de dengue na cidade de São José dos Campos entre 2013 e 2019, usando scan statistics com modelo de probabilidade Discreta de Poisson. Dois experimentos foram realizados considerando o critério de parada como 10% e 50% da população total em cada cluster. A partir das estatísticas calculadas é possível detectar zonas de risco mais alto de dengue durante os períodos estudados.

Palavras-chave: Dengue. São José dos Campos. Scan Statistics.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Mapa mostrando as populações na cidade de São José dos Campos dividida em 55 Zonas de Mobilidade entre os anos 2013 a 2019.	3
2.2 Fluxograma da Metodologia utilizada no trabalho.	6
2.3 Taxa de Incidência por 100000 habitantes para São José dos Campos entre os anos 2013 a 2019.	8
2.4 Taxa de Incidência por 100000 habitantes para São José dos Campos considerando todos os anos entre 2013-2018 agrupada por anos.	8
2.5 Taxa de Incidência por 100000 habitantes por Zona de Tráfego entre os anos 2013 a 2019.	9
2.6 Esquema de pesquisa e formação de clusters espaço-temporais através do SATScan. Fonte: (BIOMEDWARE,).	10
2.7 Dados de entrada para execução do SATScan.	11
3.1 Taxa de Incidência de Dengue por Ano.	13
3.2 Cluster gerado pelo SATScan considerando no máximo 50% da população total.	14
3.3 Estatísticas calculadas pelo SATScan considerando um máximo de população de 50% em cada cluster.	15
3.4 Cluster encontrado considerando no máximo 50% da população total.	16
3.5 Série temporal de casos por semana epidemiológica destacando o cluster detectado.	17
3.6 Clusters encontrados pelo SATScan ao executar considerando um máximo de 10% da população por cluster.	17
3.7 10 clusters encontrados considerando no máximo 10% da população total.	18
3.8 Estatísticas geradas após executar SATScan considerando uma população máxima de 10 % por cluster.	20
3.9 Série temporal de casos por semana epidemiológica destacando o primeiro cluster detectado.	20

LISTA DE TABELAS

	<u>Pág.</u>
2.1 Zonas de Tráfego definidas para a cidade de São José dos Campos.	5
3.1 Métricas calculadas para cada cluster encontrado.	19

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
2 MATERIAS E MÉTODOS	3
2.1 Estudo de Caso	3
2.2 Dados e Pré-Processamento	4
2.2.1 Fluxograma da Metodologia	6
2.2.2 Taxa de Incidência da Dengue	7
2.3 Scan Statistics	9
2.4 SatScan	10
3 ANÁLISE E RESULTADOS	13
4 CONCLUSÕES	21
REFERÊNCIAS BIBLIOGRÁFICAS	23

1 INTRODUÇÃO

Arboviroses como Dengue, Zika e Chicungunha causam impactos negativos em muitas cidades no Brasil, como, por exemplo, o aumento no número de hospitalizações e faltas no trabalho (DONALISIO et al., 2017; FREITAS CARLOS MACHADO DE, 2020). Isso implica num ponto de importante discussão sobre os desastres biológicos e epidemias e seus impactos diretos e indiretos na economia.

Considerando a taxa de incidência no Brasil, a região sudeste registrou no ano de 2020, entre as semanas epidemiológicas de 1 a 26 (período de maior incidência), uma taxa de incidência de 347,1 casos por 100.000 habitantes, sendo que o estado de São Paulo registrou uma taxa de incidência de 429,4 casos por 100.000 habitantes (SANITÁRIA, 2020).

A escala interurbana nos permite estudar como a propagação das epidemias acontece entre as cidades dada uma região. Existem diversos trabalhos desenvolvidos na escala interurbana (SABA H. E MORET, 2018; LANA et al., 2017), porém existe uma demanda por trabalhos na escala intraurbana. A escala intraurbana permite visualizar e estudar a pluralidade do espaço urbano, que é fragmentada, cada parte contendo diferente cobertura de terra, clima, sistema hidrológico e de esgoto (XAVIER DIEGO RICARDO, 2017).

A Estatística Espacial é um ramo da Estatística que contribui nos estudos epidemiológicos na parte de identificação, localização e visualização de ocorrências de doenças (ANDRADE et al., 2007). Depois de modelados os fenômenos, pode-se repensar em novas formas de organização do espaço e serviços públicos, por exemplo.

Este trabalho tem como objetivo ajudar na compreensão de dados registrados de dengue como distribuição espaço-temporal na cidade de São José dos Campos. A partir da ferramenta de varredura estatística SatScan (KULLDORFF, ; KULLDORFF, 2016; KULLDORFF, 1999) foram geradas métricas estatísticas e clusters que agregam regiões de maior atenção. A partir de evoluções de estudos como esse, é possível detectar clusters espaço-temporais e ainda verificar se eles são estatisticamente significativos, ou seja, que não foram gerados ao acaso.

2 MATERIAS E MÉTODOS

2.1 Estudo de Caso

A cidade de São José dos Campos localiza-se na Região do Vale do Paraíba, a uma distância de cerca de 90 km de São Paulo e 320 km do Rio de Janeiro. Possui uma população de 721.944 habitantes (IBGE, 2019) e área territorial de 1.099,4 km², dos quais 356 km² (33%) são de área urbana e 743,4 km² (67%) de área rural (ANTP, 2012).

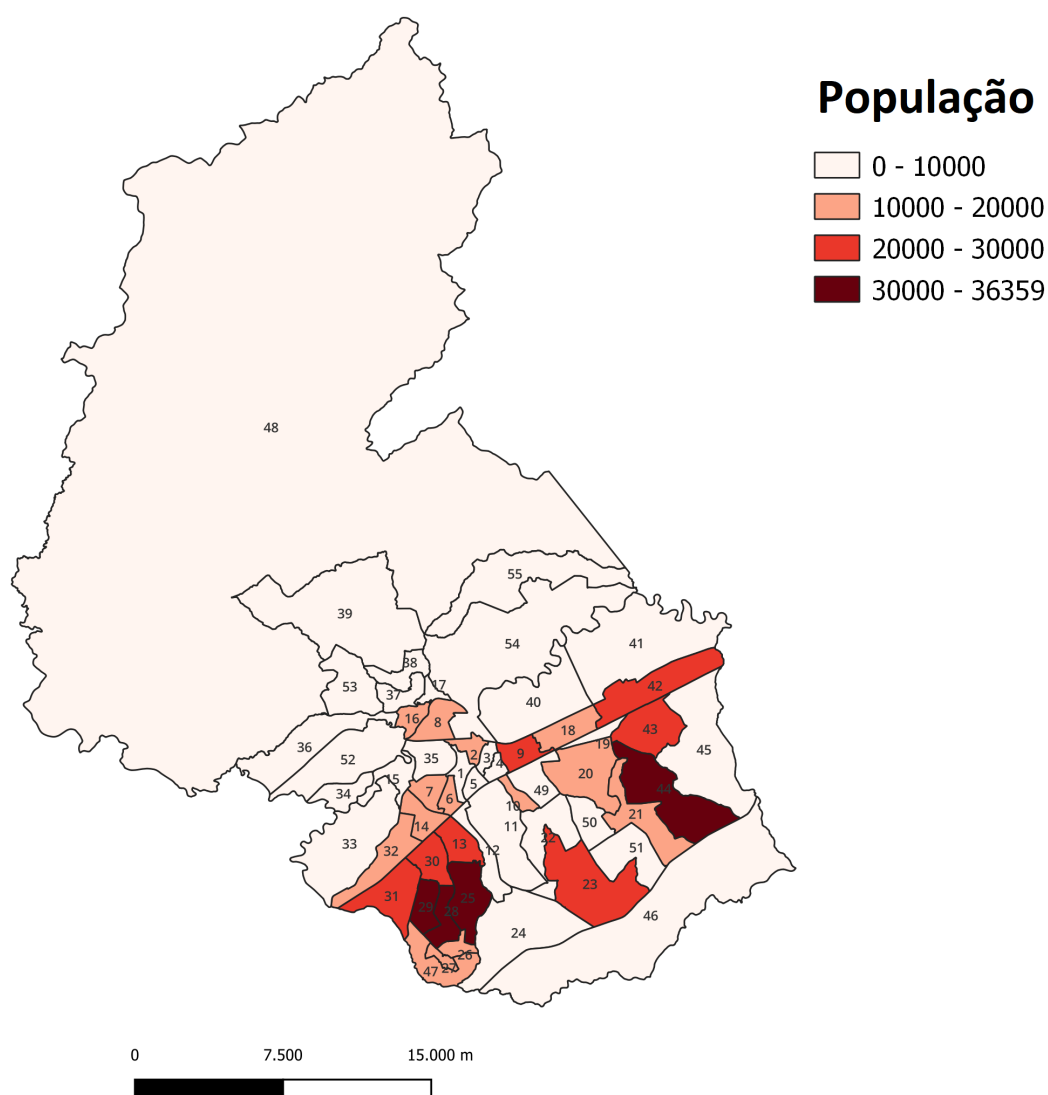


Figura 2.1 - Mapa mostrando as populações na cidade de São José dos Campos dividida em 55 Zonas de Mobilidade entre os anos 2013 a 2019.

2.2 Dados e Pré-Processamento

Os registros de casos de dengue foram fornecidos em formato CSV pelo Centro de Controle de Zoonoses (CCZ) do Departamento de Saúde da cidade de São José dos Campos e compreendem o período de 2013 a 2019. O arquivo recebido possui os seguintes campos: identificador, classificação dos casos, data, latitude, longitude e bairro. Os casos formam uma tabela contendo coordenadas geográficas (Latitude e Longitude). Em seguida, os dados foram agrupados delimitando-os por Zonas de Tráfego - ZT. Os dados foram agrupados em ZT para facilitar a comparação com dados de mobilidade, que utilizam dados definidos na pesquisa Origem-Destino (IPPLAN,). Algumas dessas zonas (36, 41, 50, 52 e 55) foram desconsideradas nos nossos cálculos por possuírem populações muito pequenas ou por não terem casos de dengue registrados. A tabela 2.1 mostra os identificadores e nomes de cada uma das ZT que foram apresentadas na Figura 2.1.

ZT	ZT_Name	ZT	ZT_Name
1	Centro	29	Cidade Morumbi e Residencial União
2	Vila Maria e Vila Santa Luzia	30	Parque Industrial e Jardim Oriente
3	Jardim Paulista e Monte Castelo	31	Palmeiras de São José e Rio Comprido
4	Vila São Pedro e Jardim Jussara	32	Jardim Alvorada e Jardim das Indústrias
5	Vila Piratininga e Jardim Augusta	33	Limoeiro
6	Vila Adyanna e Vila Betânia	34	Urbanova
7	Vila Ema e Jardim Esplanada	35	Banhado
8	Santana e Vila Rossi	36	Vazio urbano
9	Vazio urbano	37	Vila São Geraldo e Jardim Minas Gerais
10	Jardim da Granja e Jardim Souto	38	Vila Paiva e Jardim Boa Vista
11	CTA e INPE	39	Chácara dos Freitas e Chácara do Bonsucesso

12	Vila Nova Conceição e Vila São Bento	40	Residencial Vista Linda
13	Jardim Satélite e Floradas de São José	41	Vazio urbano
14	Jardim Aquarius e Jardim Colinas	42	Eugênio de Melo e Galo Branco
15	Condomínio Esplanada do Sol e Serimbura	43	Jardim Santa Inês e Residencial Frei Galvão
16	Jardim Telespark e Jardim Altos de Santana	44	Parque Novo Horizonte e Residencial Dom Bosco
17	Alto da Ponte e Vila Veneziani	45	Jardim Ebenezer e Chácara Sítio Jata
18	Vila Tesouro e Jardim Universo	46	Cajuru
19	Jardim Motorama e Jardim Nova Detroit	47	Jardim República e Vila das Flores
20	Jardim Americano e Vista Verde	48	São Francisco Xavier e Área Rural
21	Chácara Pousada do Vale e Campos de São José	49	Jardim Diamante e REVAP
22	Residencial Flamboyant e Jardim Uirá	50	Vazio urbano
23	Jardim do Lago e Vila Iracema	51	Santa Cecília e Jardim Helena
24	Parque Interlagos e Altos da Quinta	52	Paratehy e Residencial Jaguar
25	Bosque dos Eucaliptos e Quinta das Flores	53	Chácara Fazenda Caeté
26	Conjunto Habitacional Dom Pedro	54	Águas de Canindé e Chácara Havaí
27	Jardim Imperial e Jardim Colonial	55	Chácara Pedra D'água II
28	Bosque dos Ipês e Campo dos Alemães		

Tabela 2.1 - Zonas de Tráfego definidas para a cidade de São José dos Campos.

Após ter sido feito o agrupamento dos casos por ZT, foi obtido um arquivo com séries temporais de casos de dengue, onde cada coluna do arquivo possui uma série temporal para uma zona. Cada coluna representa uma ZT e cada linha, o número de casos registrados para um dia da semana entre 2013 e 2019.

2.2.1 Fluxograma da Metodologia

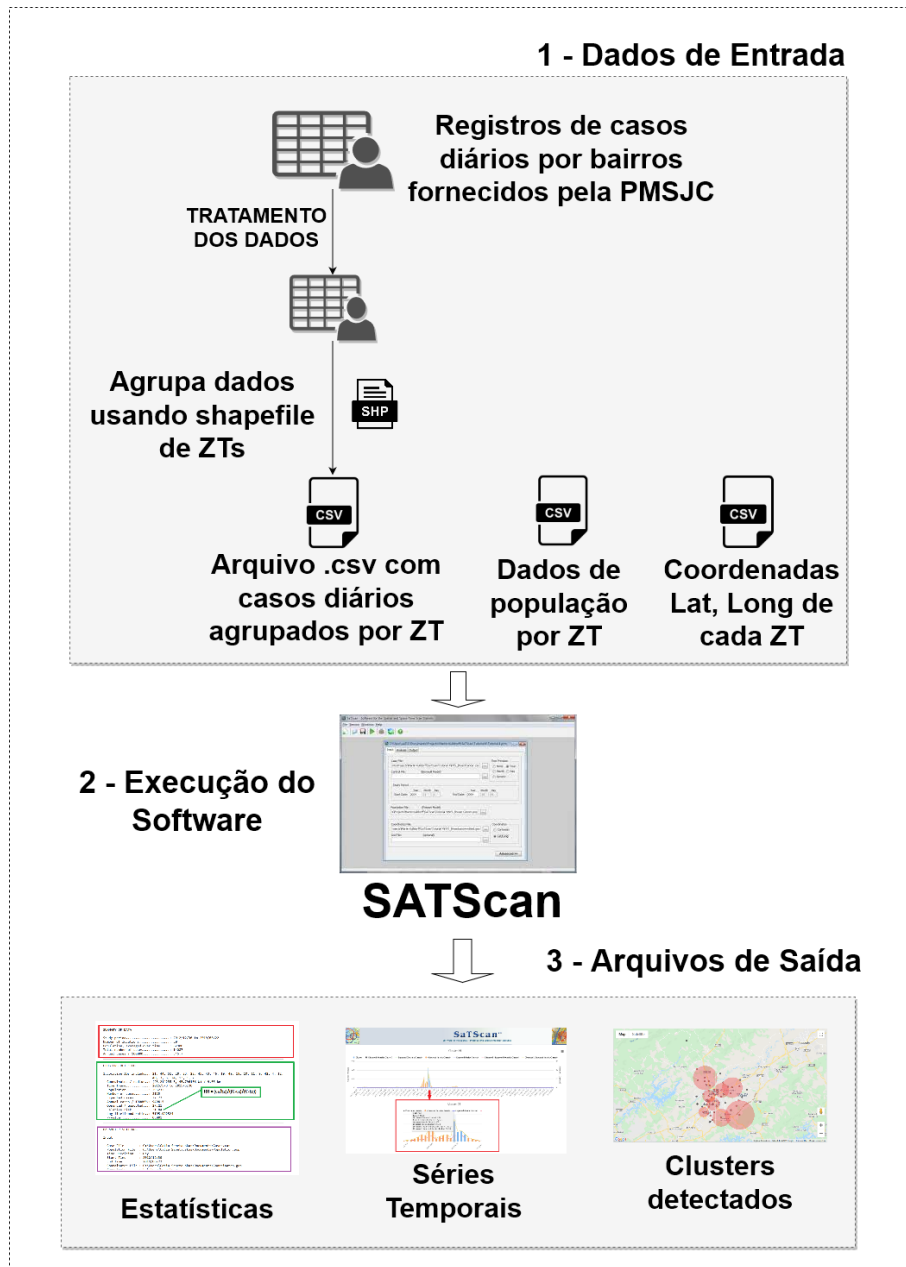


Figura 2.2 - Fluxograma da Metodologia utilizada no trabalho.

A Figura 2.2 mostra o fluxograma da metodologia usada nesse trabalho, desde o tratamento dos dados até a geração das saídas dos dados.

- 1 - Dados de Entrada - o primeiro tratamento dos dados de casos de dengue foi detalhado na Seção 2.2. Para ser usado como dados de entrada no SATScan, esse arquivo precisa estar organizado de forma que cada linha contenha: a data do registro, o identificador da ZT e a quantidade diária de casos, vide exemplo na Figura 2.7 b. Os dados de população vieram do Censo do IBGE 2010 e também foram agrupados de setores censitários em ZT.
- 2 - Execução do Software - após gerar os arquivos de entrada necessários, estes são usados como entrada no SATScan.
- 3 - Arquivos de Saída - na Figura 2.2 é possível ver algumas das saídas possíveis geradas pelo SATScan: estatísticas dos clusters, séries temporais e clusters detectados no mapa.

2.2.2 Taxa de Incidência da Dengue

A taxa de incidência de uma doença expressa o número de casos novos durante um período de tempo definido, numa população sob o risco de desenvolver a doença. A incidência mede o risco ou probabilidade de ocorrer um evento de uma doença na população exposta.

Neste trabalho, foi escolhido trabalhar com semanas epidemiológicas. O primeiro gráfico usando taxa de incidência é mostrado na Figura 2.3. Ele mostra a taxa de incidência por 100.000 habitantes na cidade de São José dos Campos por semana epidemiológica considerando os dados entre os anos de 2013 e 2019. É possível verificar uma elevação entre as semanas 106 e 131, que correspondem às semanas do ano 2015, onde o Brasil teve uma alta incidência de casos de Dengue, juntamente com Chikungunya e Zika ao mesmo tempo.

Já na Figura 2.4 mostra para cada semana epidemiológica dos anos entre 2013 e 2018 a taxa de incidência da dengue, sendo que o ano de 2019 não aparece nesse figura por causa da falta de dados para a segunda metade do ano. Pode-se verificar claramente que a curva em verde, que representa a incidência no ano de 2015, apresenta uma amplitude muito maior em relação às outras. Também é possível notar que as semanas entre 4 e 20 (aproximadamente entre janeiro e junho) apresentam valores maiores de incidência em relação às outras semanas.

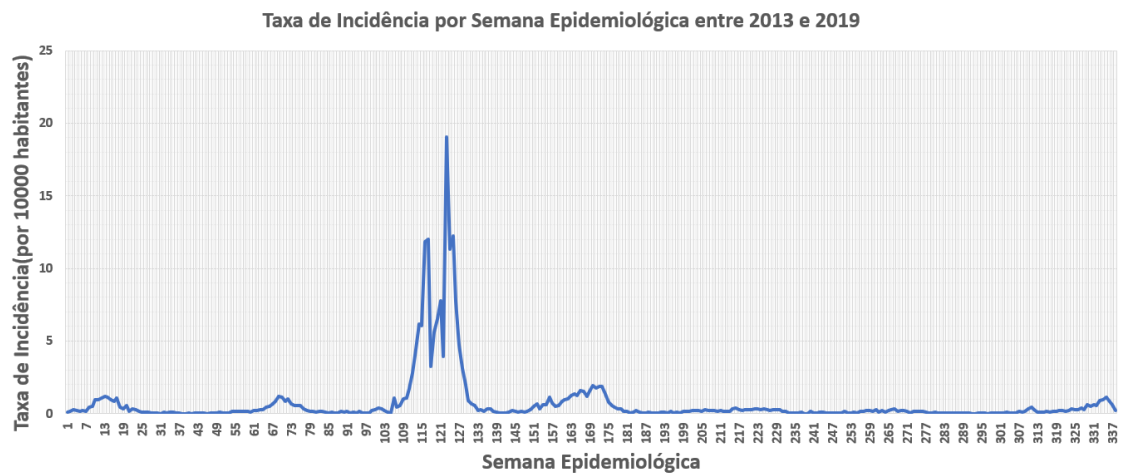


Figura 2.3 - Taxa de Incidência por 100000 habitantes para São José dos Campos entre os anos 2013 a 2019.

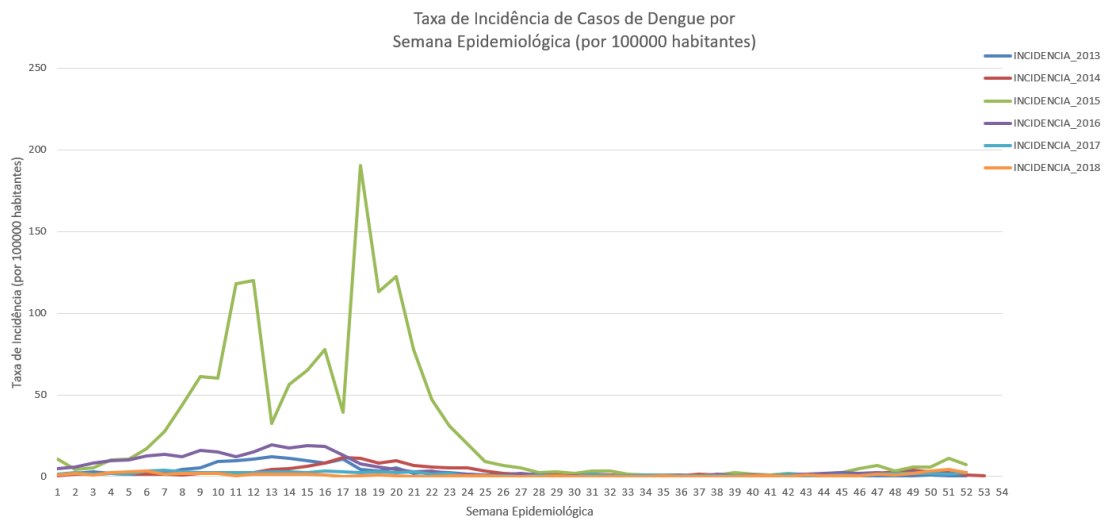


Figura 2.4 - Taxa de Incidência por 100000 habitantes para São José dos Campos considerando todos os anos entre 2013-2018 agrupada por anos.

Por fim, taxas de casos de dengue para cada ZT foram agrupadas durante todos os anos para identificar quais seriam as zonas com maiores taxas de incidência. Como podemos verificar na Figura 2.5, as zonas com maiores taxas de incidência da dengue foram $ZT = 3$ (Jardim Paulista e MonteCastelo) e $ZT = 52$ (Paratehy e Residencial Jaguary).

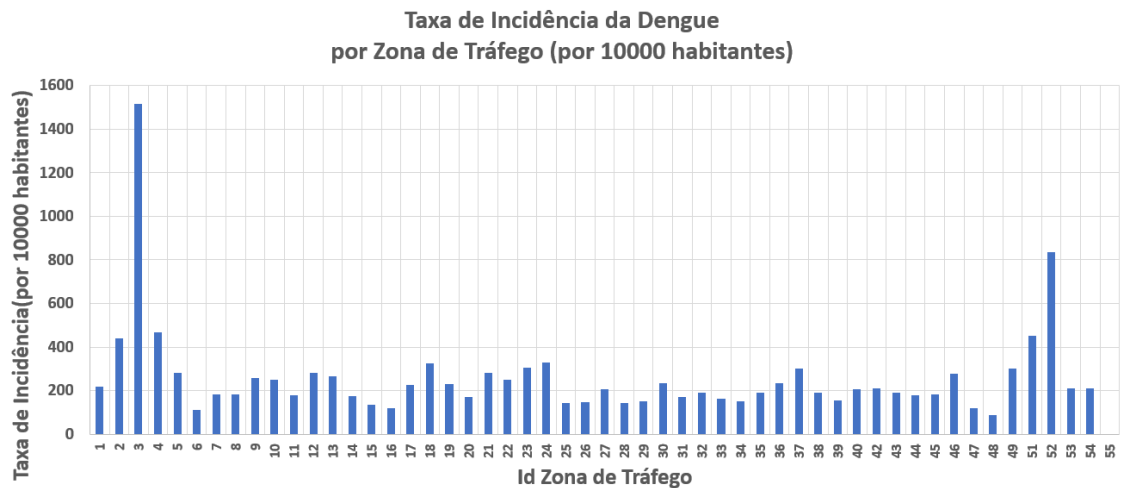


Figura 2.5 - Taxa de Incidência por 100000 habitantes por Zona de Tráfego entre os anos 2013 a 2019.

2.3 Scan Statistics

Scan Statistics é uma técnica estatística que consiste em varrer um mapa com alguma estrutura de janela que pode ser um círculo, uma elipse, etc, somando os eventos das regiões dentro de um raio até que um critério de parada seja atingido. Considerando a varredura espaço-temporal, essa região passa a incluir uma dimensão a mais, tendo então um cilindro, por exemplo. A Figura 2.6 mostra um exemplo de cilindros espaço-temporais usados na varredura estatística de eventos no mapa.

No método de Besag Newell, o critério de parada consiste em agregar regiões até que um número de eventos seja alcançado. Já para Kuldorff, o critério de parada é a população dentro de um cluster.

Para cada cluster gerado, é calculada uma métrica chamada Razão de Verossimilhança, que se baseia no número de casos observados dentro e fora do cluster. O cluster que possui o maior valor para essa métrica é chamado de cluster mais verossímil. Uma vez identificado esse cluster, são executadas simulações de Monte Carlo para cálculo da significância.

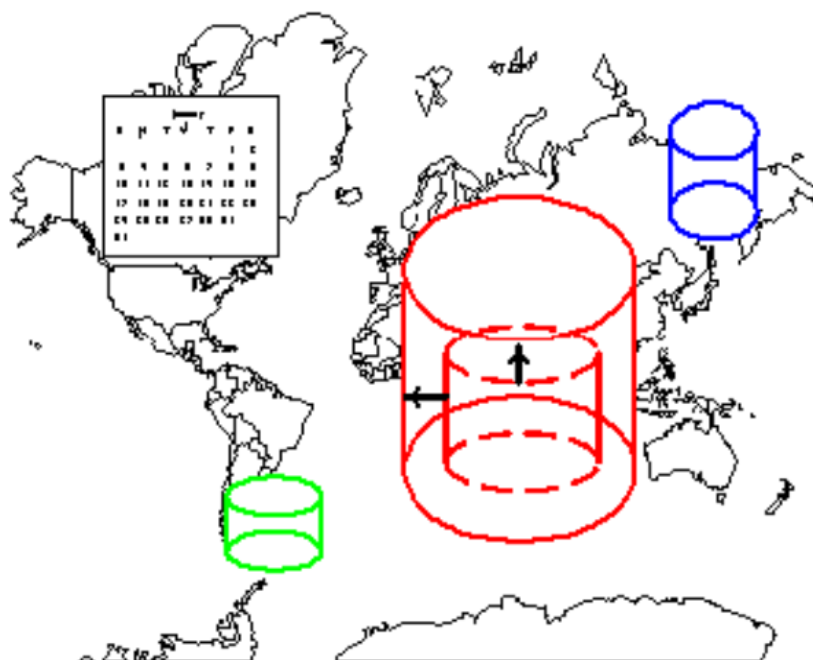


Figura 2.6 - Esquema de pesquisa e formação de clusters espaço-temporais através do SATScan. Fonte: (BIOMEDWARE,).

Neste trabalho, quanto maior o número de casos observados da doença dentro do cluster em relação aos casos observados fora do cluster, maior será a Razão de Verossimilhança do cluster.

2.4 SatScan

O SatScan é uma ferramenta de análise dados espaciais, temporais e espaço-temporais (KULLDORFF, ; KULLDORFF, 1999; KULLDORFF et al., 2007; KULLDORFF, 2016) a partir da aplicação da técnica de varredura estatística. Ele permite definir a janela de varredura como um intervalo de tempo e um círculo no espaço ou um cilindro para o caso de espaço-tempo.

Neste trabalho, foi feita a análise espaço-temporal de casos diariamente notificados de dengue na cidade de São José dos Campos entre 2013 e 2019, usando Scan Statistics com modelo de probabilidade Discreta de Poisson. Foram consideradas 50 das 55 ZT da base, pois são considerados como "Vazios Urbanos", então alguns deles não tinham casos registrados.

Como dado de entrada do SATScan para a análise de dados espaço-temporais, devemos passar ao menos três arquivos de entrada: um arquivo contendo as populações

de cada subregião com os identificadores, um arquivo contendo dados diários para cada uma das subregiões, um arquivo contendo as coordenadas geográficas (latitude e longitude) de cada subregião. Um exemplo de cada um desses arquivos, respectivamente, pode ser visto na Figura 2.7. Como a avaliação foi feita por semanas epidemiológicas, ao rodar o SATScan, é necessário informar ao programa a quantidade de dias que devem ser agrupados e também informar os limites mínimos e máximos de dias/semanas a serem considerados na janela. Nesse trabalho, foram considerados um mínimo de 1 semana epidemiológica e 4 semanas epidemiológicas.

a) Arquivos com populações

ZT	Pop_IBGE
1	8378
2	11615
3	8870
4	8122
5	5472
6	11995
7	14796
8	17425
9	21544
10	11252
11	2509
12	6328
13	27201
14	14967
15	1555

b) Arquivos com casos diários

DATE	ZT	DENGUE
30/12/2012	1	0
31/12/2012	1	0
01/01/2013	1	0
02/01/2013	1	0
03/01/2013	1	0
04/01/2013	1	0
05/01/2013	1	0
06/01/2013	1	0
07/01/2013	1	0
08/01/2013	1	0
09/01/2013	1	0
10/01/2013	1	0
11/01/2013	1	0
12/01/2013	1	0
13/01/2013	1	0
14/01/2013	1	0

c) Arquivos com coordenadas geográficas

ZT	LONGITUDE	LATITUDE
1	-458867015	-2319063757
2	-4588207028	-231780091
3	-4587374462	-2318293872
4	-4586932666	-2318890402
5	-4588035978	-2319587905
6	-4589225193	-2320321786
7	-4590217501	-2319970357
8	-4589806923	-2316530199
9	-4585707696	-231807981
10	-4585807174	-2320225179
11	-4586545731	-2322147156
12	-4587482021	-2323327146
13	-4588742879	-2322558912
14	-459062476	-2321243557
15	-4592069136	-2319350488
16	-4591089574	-2316347543
17	-4590138781	-2315125182
18	-4583361971	-2316929201
19	-4581931208	-2317145827
20	-4582485931	-2319072316

Figura 2.7 - Dados de entrada para execução do SATScan.

Para cada Zona de Tráfego, existem dados sobre o número observado de casos, da população em risco e das coordenadas geográficas do centróide, que é um ponto no interior da zona de tráfego. Dentro de um cluster c , considera-se que indivíduos em qualquer zona pertencente a c , a probabilidade de que ele venha a ser um caso de dengue como p_z e a probabilidade fora do cluster c como $p_{z'}$. Assim, temos as seguintes Hipóteses:

- H_0 : $p_z = p_{z'}$, qualquer que seja o cluster
- H_1 : existe ao menos uma zona em que $p_z > p_{z'}$

Assim, H_0 (hipótese nula) diz que existe a mesma probabilidade de alguém ter dengue estando dentro ou fora do cluster e H_1 (hipótese alternativa) diz que existe ao menos uma zona onde a probabilidade é maior dentro do cluster que fora. Então, através da estatística de varredura, é possível calcular a razão de verossimilhança que atribui a cada cluster um valor de acordo com a equação a seguir.

$$\lambda_z = \begin{cases} \left(\frac{c_z/n_z}{C/N}\right)^{c_z} \left(\frac{c_{z'}/n_{z'}}{C/N}\right)^{c_{z'}}, & se(c_z/n_z) > (c_{z'}/n_{z'}). \\ 1, & \text{caso contrário.} \end{cases},$$

onde:

- c_z - casos dentro do cluster
- n_z - população dentro do cluster
- C - casos totais no mapa
- N - população total no mapa
- $c_{z'}$ - casos fora do cluster
- $n_{z'}$ - população fora do cluster

Como foi explicado anteriormente, são gerados diversos agrupamentos a partir de cada zona de tráfego e comparados uns aos outros. O cluster que apresentar maior valor de verossimilhança é eleito o cluster mais verossímil (SAENGER, 2018).

Neste trabalho, foram realizados dois experimentos diferentes considerando a população máxima como sendo 50% da população por ser o padrão da ferramenta SatScan e a mais usada nos trabalhos e um outro experimento considerando 10% da população por cluster. Os resultados desses experimentos serão detalhados no capítulo de Análise e Resultados.

3 ANÁLISE E RESULTADOS

A Figura 3.1 mostra a taxa de incidência por ano para a cidade. Apesar da grande diminuição na taxa de 2015 até o ano de 2018, é possível perceber que o valor da taxa volta a aumentar no ano de 2019, sendo que para esse ano os dados estão registrados até o mês de junho, logo, ao fim do ano a tendência é que essa taxa estivesse ainda maior.

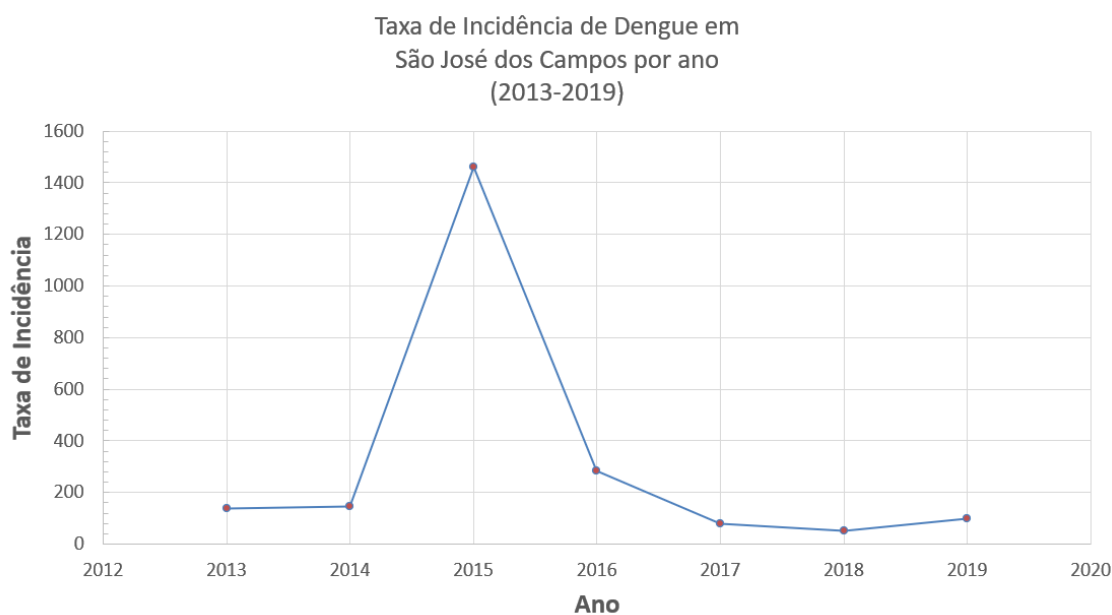


Figura 3.1 - Taxa de Incidência de Dengue por Ano.

Algumas características em relação aos dados registrados na cidade de São José dos Campos em geral são as seguintes:

- Número total de casos (entre 2013-2019): 14087
- Casos anuais por 100000 habitantes : 345,7
- Número máximo de casos semanais: 1086
- Semana com número máximo de casos: 05/05/2015-11/05/2015

A ferramenta SatScan foi utilizada para detectar clusters espaço-temporal nos nossos dados. Foi aplicado Scan Statistics Kulldorff's com modelo de probabilidade discreto de Poisson, devido à natureza dos dados de entrada que dados de contagem de

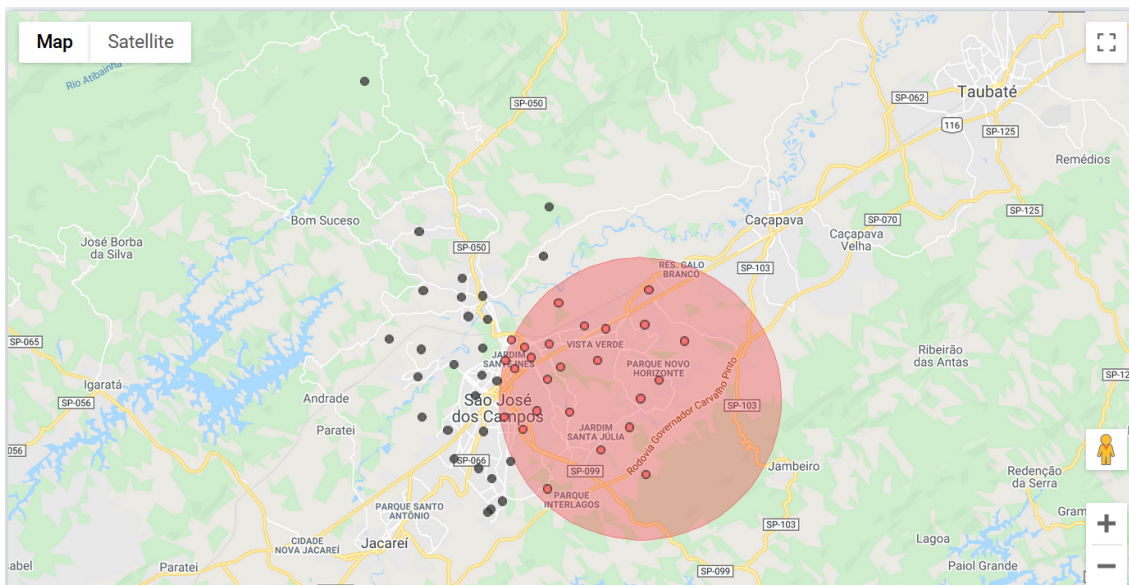


Figura 3.2 - Cluster gerado pelo SATScan considerando no máximo 50% da população total.

casos de dengue. Além disso, foram executadas simulações de Monte Carlo ($n=999$) para avaliação de significância estatística, onde foram considerados clusters com significância estatística ($p\text{-value} < 0,05$), sem sobreposição no espaço e que incluíam no máximo 50% da população de São José dos Campos. O resultado dessa execução pode ser visto na Figura 3.2. Além disso, as características desse cluster são:

- Ids das zonas: 21, 44, 51, 20, 23, 22, 45, 43, 46, 19, 49, 18, 10, 11, 9, 42, 4, 12, 40, 3, 5, 24, 13, 1, 2
- Período: Semanas entre 03/05/2015-30/05/2015
- População: 293232
- Número de casos: 2115
- Casos esperados: 77,72
- Casos anuais / 100000 habitantes: 9408,5
- Observados/Esperados: 27,21
- Risco Relativo: 31,84
- Razão de verossimilhança: 5105,992

Um dos arquivos de saída fornecido pelo próprio SATScan em formato txt, é o arquivo contendo estatísticas calculadas sobre o cluster, e uma parte desse arquivo pode ser vista na Figura 3.3. Nesse arquivo, é possível ver os dados gerais (destacado em vermelho): período estudado, número de regiões, população total, número total de casos e casos anuais por 100000 habitantes. Em seguida, para cada cluster encontrado, são detalhadas informações: identificadores das zonas; coordenadas do cluster e raio; período em que foi detectado este cluster; população dentro do cluster; casos esperados dentro do cluster; casos anuais por 100000 habitantes; casos observados/esperados; risco relativo; razão de verossimilhança; p-value. Os clusters encontrados são ordenados pela Razão de Verossimilhança em ordem decrescente. A razão de Verossimilhança leva em conta a quantidade de casos observados dentro e fora do cluster. Já o Risco Relativo (RR), mostrado na última coluna da tabela, é calculado comparando o risco numa zona em relação ao risco fora da zona, se esses riscos forem iguais, o Risco Relativo será igual a 1 (PRATES et al., 2014).

```

SUMMARY OF DATA
Study period.....: 2012/12/30 to 2019/06/22
Number of locations.....: 50
Population, averaged over time.....: 629005
Total number of cases.....: 14087
Annual cases / 100000.....: 345.7

CLUSTERS DETECTED
1.Location IDs included.: 21, 44, 51, 20, 23, 22, 45, 43, 46, 19, 49, 18, 10, 11, 9, 42, 4, 12,
                        40, 3, 5, 24, 13, 1, 2
Coordinates / radius..: (23.214235 S, 45.796156 W) / 9.65 km
Time frame.....: 2015/5/3 to 2015/5/30
Population.....: 293232
Number of cases.....: 2115
Expected cases.....: 77.72
Annual cases / 100000.: 9408.5
Observed / expected...: 27.21
Relative risk.....: 31.84
Log likelihood ratio..: 5105.992524
P-value.....: 0.001

PARAMETER SETTINGS
Input
-----
Case File       : C:\Users\Catia Sepetauskas\Documents\Cases.cas
Population File : C:\Users\Catia Sepetauskas\Documents\Population.pop
Time Precision  : Day
Start Time     : 2012/12/30
End Time      : 2019/06/22
Coordinates File : C:\Users\Catia Sepetauskas\Documents\Coordinates.geo
Coordinates    : latitude/longitude

```

Figura 3.3 - Estatísticas calculadas pelo SATScan considerando um máximo de população de 50% em cada cluster.

A partir dos identificadores de zonas dentro do cluster fornecidos no arquivo mostrado na Figura 3.3, um novo mapa foi gerado utilizando a ferramenta QGIS (QGIS,) para facilitar a visualização das regiões que estão incluídas no cluster encontrado. Desta forma, foi gerada a imagem da Figura 3.4. Pode-se verificar que a ZT com identificador 50 está colorida em verde (classificada com "Fora do cluster") no centro do cluster, pois ela foi uma das regiões que não foram consideradas no cálculo, como foi explicado anteriormente.

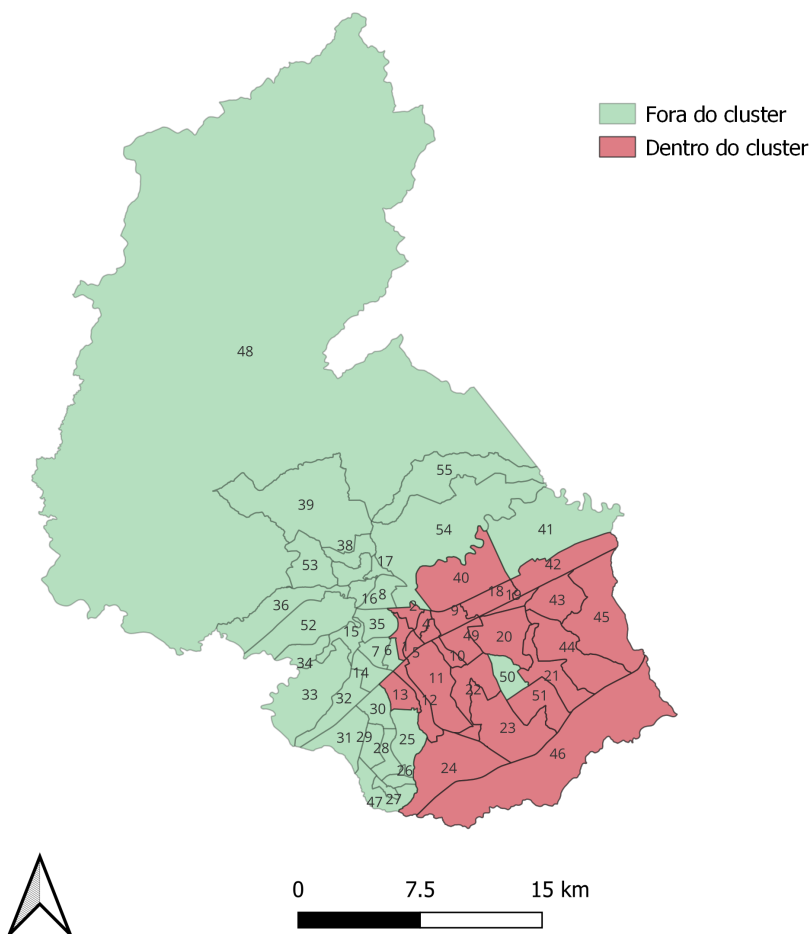


Figura 3.4 - Cluster encontrado considerando no máximo 50% da população total.

Um outro dado fornecido como saída pelo SATScan, é a série temporal mostrada na Figura 3.5, que mostra a série temporal de casos de dengue por semana epidemiológica, dando destaque para a semana em que foi detectado o cluster. É possível

verificar no destaque em azul quando o cluster foi detectado, e ao clicar sobre o gráfico é possível ver as informações referentes ao cluster. Esta visualização permite verificar os casos observados e casos esperados dentro e fora do cluster.

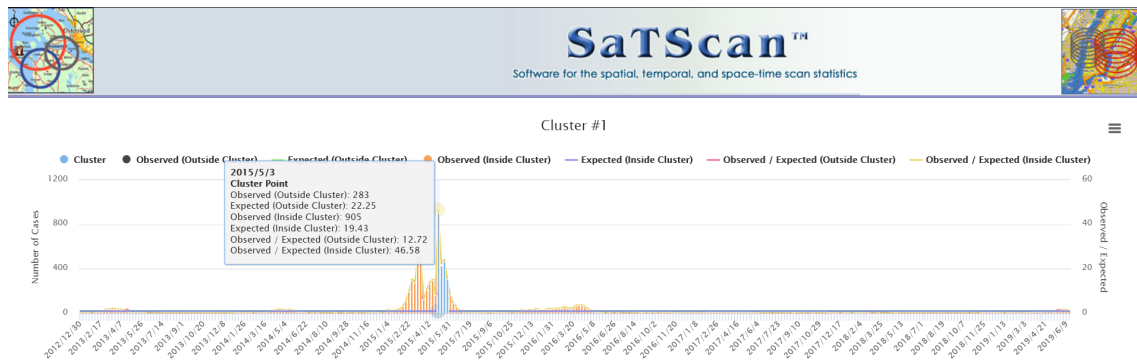


Figura 3.5 - Série temporal de casos por semana epidemiológica destacando o cluster detectado.

Depois desse experimento, fizemos algumas combinações de variações entre os limites temporais e espaciais para formação dos clusters. O segundo experimento, então, consiste em considerar os clusters que continham no máximo 10% da população da cidade, uma vez que algumas zonas possuem até 5% de toda a população da cidade. O resultado pode ser visto na Figura 3.6.

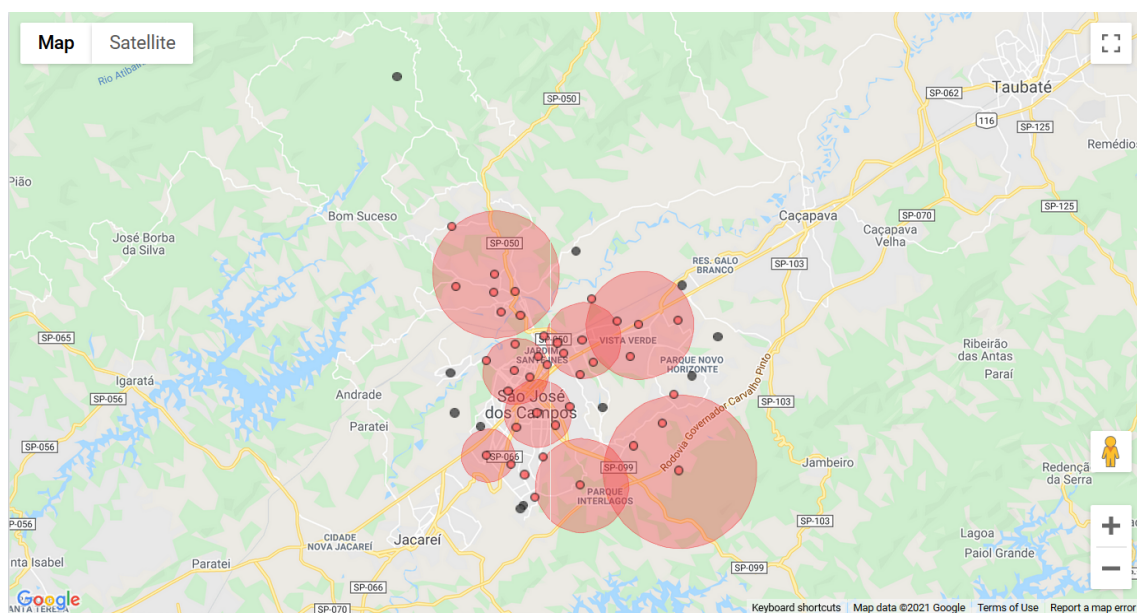


Figura 3.6 - Clusters encontrados pelo SaTScan ao executar considerando um máximo de 10% da população por cluster.

Em seguida, foi utilizada a ferramenta QGIS para gerar o mapa mostrado na Figura 3.7.

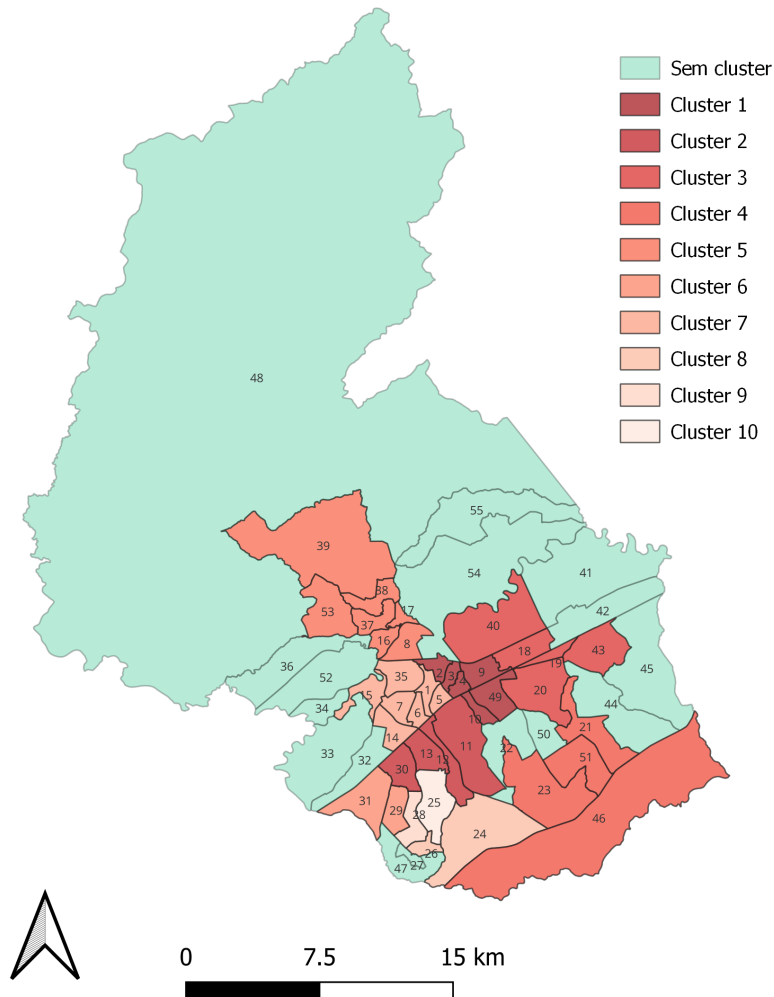


Figura 3.7 - 10 clusters encontrados considerando no máximo 10% da população total.

O SATScan detectou dez clusters espalhados pela cidade, sendo que foram encontrados 2 clusters contendo apenas uma ZT de tráfego com populações de 31955 e 36359, número de casos observados de 117 e 111, com taxa de risco relativo de 13,20 e 9,50. É uma particularidade do SATScan considerar um cluster como um agregado de apenas uma zona, ou seja, sem a necessidade de ter vizinhos para ser considerado um cluster. Todos os clusters foram detectados num período de tempo de 4 semanas epidemiológicas entre os dias 03/05/2015 a 30/05/2015.

Pode-se verificar que o primeiro cluster, ou seja, com maior razão de verossimilhança engloba zonas de tráfego mais ao centro da cidade de São José dos Campos. As zonas que fazem parte desse cluster são: 2,3,4,10,49 e 9, apesar de que, se verificarmos na Figura 2.1, pode-se ver que essas não são áreas com maior população da cidade. Isso pode significar que existe uma outra razão para o alto Risco Relativo desse cluster, por exemplo, considerar a mobilidade entre essas zonas centrais.

A Tabela 3.1 mostra as características dos clusters encontrados mostrados na Figura 3.7. Todos os clusters foram detectados no mesmo período de tempo em 4 semanas epidemiológicas entre 03/05/2015 a 30/05/2015. Nessa tabela, os clusters estão ordenados pela Razão de Verossimilhança.

N Zonas	Período	População	Observados	Esperados	RR
6	03/05/15 - 30/05/15	62466	991	16,56	64,31
4	03/05/15 - 30/05/15	58866	326	15,60	21,37
5	03/05/15 - 30/05/15	60419	278	16,01	17,69
4	03/05/15 - 30/05/15	38696	207	10,26	20,47
7	03/05/15 - 30/05/15	62102	204	16,46	12,56
2	03/05/15 - 30/05/15	60549	200	16,05	12,63
7	03/05/15 - 30/05/15	58475	157	15,50	10,23
2	03/05/15 - 30/05/15	26629	117	7,06	16,71
1	03/05/15 - 30/05/15	31955	111	8,47	13,20
1	03/05/15 - 30/05/15	36359	91	9,64	9,50

Tabela 3.1 - Métricas calculadas para cada cluster encontrado.

Quando diversos clusters são detectados, o SATScan gera um agrupamento de informações para cada cluster. Como pode ser visto na Figura 3.8, o cluster principal, primário aparece destacado em vermelho. Os próximos clusters estão apresentados em seguida como clusters secundários.

Por fim, a Figura 3.9 mostra a série temporal de casos entre o ano 2013 a 2019 considerando a população máxima em cada cluster em 10 %. Essa Figura mostra apenas a série para o Cluster primário encontrado.

CLUSTERS DETECTED

1. Location IDs included.: 9, 4, 49, 3, 10, 2
 Coordinates / radius.: (23.180798 S, 45.857077 W) / 2.57 km
 Time frame.....: 2015/5/3 to 2015/5/30
 Population.....: 62466
 Number of cases.....: 991
 Expected cases.....: 16.56
 Annual cases / 100000.: 20694.4
 Observed / expected...: 59.86
 Relative risk.....: 64.31
 Log likelihood ratio..: 3115.252769
 P-value.....: 0.001

Cluster Primário

2. Location IDs included.: 13, 12, 30, 11
 Coordinates / radius.: (23.225589 S, 45.887429 W) / 2.29 km
 Time frame.....: 2015/5/3 to 2015/5/30
 Population.....: 58866
 Number of cases.....: 326
 Expected cases.....: 15.60
 Annual cases / 100000.: 7224.0
 Observed / expected...: 20.90
 Relative risk.....: 21.37
 Log likelihood ratio..: 683.933145
 P-value.....: 0.001

Clusters Secundários

3. Location IDs included.: 19, 18, 20, 43, 40
 Coordinates / radius.: (23.171458 S, 45.819312 W) / 3.68 km
 Time frame.....: 2015/5/3 to 2015/5/30
 Population.....: 60419
 Number of cases.....: 278
 Expected cases.....: 16.01
 Annual cases / 100000.: 6002.0
 Observed / expected...: 17.36
 Relative risk.....: 17.69
 Log likelihood ratio..: 533.935306
 P-value.....: 0.001

4. Location IDs included.: 46, 51, 23, 21
 Coordinates / radius.: (23.261170 S, 45.792616 W) / 5.23 km
 Time frame.....: 2015/5/3 to 2015/5/30
 Population.....: 38506

Figura 3.8 - Estatísticas geradas após executar SATScan considerando uma população máxima de 10 % por cluster.

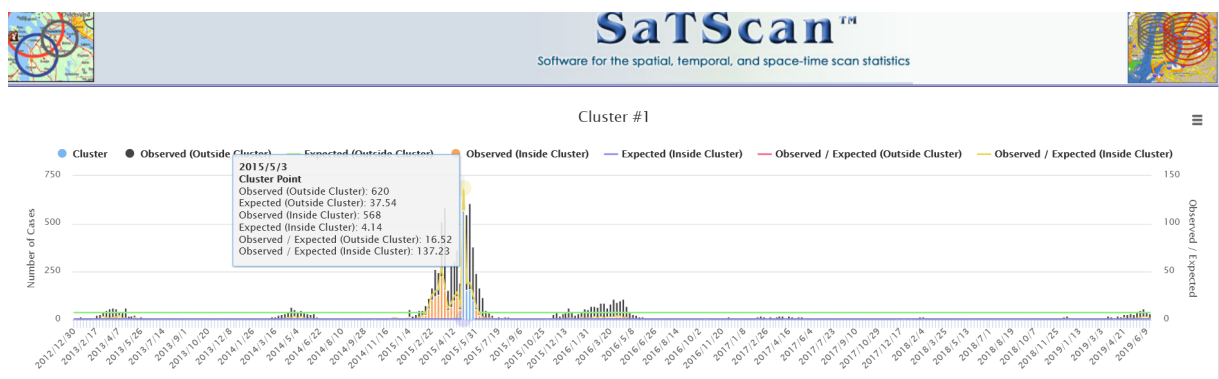


Figura 3.9 - Série temporal de casos por semana epidemiológica destacando o primeiro cluster detectado.

4 CONCLUSÕES

Este trabalho foi um trabalho descritivo sobre a distribuição espacial dos dados de dengue na cidade de São José dos Campos, utilizando como ferramenta o SATScan. Através do uso da ferramenta SatScan foram calculados clusters (agrupamentos) espaço-temporais que consideram, na formação dos clusters, informações espaciais e temporais. Foram realizados dois experimentos considerando como critério de parada os limites máximos de 50% e 10% na população dentro do cluster em relação à população total do mapa.

Como trabalhos futuros, pretende-se considerar outras janelas de tempo e espaço. Além disso, com o uso de grafos, podemos aplicar algoritmos de detecção de comunidades e verificar se há relações entre as comunidades encontradas e aquelas detectadas pelo SATScan.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, A. L.; MONTEIRO, A. M. V.; BARCELLOS, C.; LISBOA, E.; ACOSTA, L. M. W.; ALMEIDA, M. C. d. M.; BRITO, M. R. V.; CARVALHO, M. S.; SANTOS, M. A. d.; CRUZ, O. et al. Introdução à estatística espacial para a saúde pública. 2007. 1

ANTP. **Sistema de Informações da Mobilidade Urbana - Relatório Geral 2012**. 2012. Acessado em 16.12.2021. Disponível em: <<http://www.antp.org.br>>. 3

BIOMEDWARE. **About Kulldorff's Scan method**. Acessado em 20.12.2021. Disponível em: <https://www.biomedware.com/files/documentation/clusterseer/KScan/About_Kulldorffs_Scan.htm>. v, 10

DONALISIO, M. R.; FREITAS, A. R. R.; ZUBEN, A. P. B. V. Arboviroses emergentes no brasil: desafios para a clínica e implicações para a saúde pública. **Revista de saúde pública**, SciELO Brasil, v. 51, 2017. 1

FREITAS CARLOS MACHADO DE, e. a. Desastres naturais e seus custos nos estabelecimentos de saúde no brasil no período de 2000 a 2015. **Cadernos de Saúde Pública**, v. 36, 2020. 1

IPPLAN, P. e. **Prefeitura de São José dos Campos - PMSJC E Instituto de Pesquisa, Administração e Planejamento - IPPLAN. - Atlas da Pesquisa Origem e Destino - Panorama da Mobilidade em São José dos Campos. 2014**. Acessado em 16.12.2021. Disponível em: <https://www.sjc.sp.gov.br/media/56152/atlas_origem_destino_baixa_res.pdf>. 4

KULLDORFF, M. **SaTScan™ User Guide for version 9.6 [Internet]**. 2018. [S.l.]: Available. 1, 10

_____. Spatial scan statistics: models, calculations, and applications. In: **Scan statistics and applications**. [S.l.]: Springer, 1999. p. 303–322. 1, 10

_____. Satscantm manual do usuário. 2016. 1, 10

KULLDORFF, M.; MOSTASHARI, F.; DUCZMAL, L.; YIH, W. K.; KLEINMAN, K.; PLATT, R. Multivariate scan statistics for disease surveillance. **Statistics in medicine**, Wiley Online Library, v. 26, n. 8, p. 1824–1833, 2007. 10

LANA, R. M.; GOMES, M. F. d. C.; LIMA, T. F. M. d.; HONORIO, N. A.; CODEÇO, C. T. The introduction of dengue follows transportation infrastructure changes in the state of acre, brazil: A network-based analysis. **PLoS neglected tropical diseases**, Public Library of Science San Francisco, CA USA, v. 11, n. 11, p. e0006070, 2017. 1

PRATES, M. O.; KULLDORFF, M.; ASSUNÇÃO, R. M. Relative risk estimates from spatial and space–time scan statistics: are they biased? **Statistics in medicine**, Wiley Online Library, v. 33, n. 15, p. 2634–2644, 2014. 15

QGIS. **QGIS**. Acessado em 20.12.2021. Disponível em:
<<https://www.qgis.org/en/site/>>. 16

SABA H. E MORET, M. A. e. B. F. R. e. A. M. L. V. e. J. E. M. F. e. N. F. A. S. e. M. J. G. V. Relevance of transportation to correlations among criticality, physical means of propagation, and distribution of dengue fever cases in the state of bahia. **Science of the Total Environment**, v. 618, n. 23, p. 971–976, 2018. 1

SAENGER, C. C. Análise da distribuição espaço-temporal de intoxicação humana por agrotóxicos agrícolas e de metais pesados em águas superficiais no estado de goiás. 2018. 12

SANITÁRIA, S. de V. **Monitoramento dos casos de arboviroses urbanas transmitidas pelo Aedes (dengue, chikungunya e Zika), Semanas Epidemiológicas 1 a 4**. 2020. Disponível em:
<https://www.gov.br/saude/pt-br/assuntos/media/pdf/2020/dezembro/11/boletim_epidemiologico_svs_48.pdf>. Acesso em: 06 dez. 2021. 1

XAVIER DIEGO RICARDO, e. a. Spatial-temporal diffusion of dengue in the municipality of rio de janeiro, brazil, 2000-2013. **Cadernos de saude publica**, 2017. 1