

Um Estudo sobre a distribuição Espaço-Temporal da dengue na cidade de São José dos Campos usando Scan Statistics

Aplicando SATScan a Dados de Casos de Dengue

Cátia Nascimento

INPE - Instituto Nacional de Pesquisas Espaciais

19/11/2021



1 Introdução

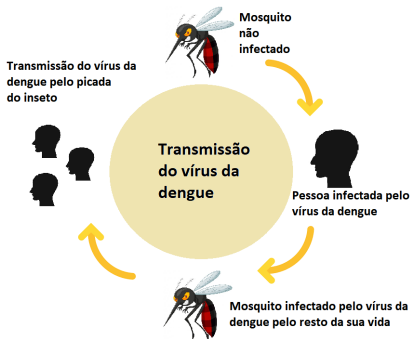
2 Materiais e Métodos

3 Conclusões

1 Introdução

2 Materiais e Métodos

3 Conclusões



- A dengue pode causar diversos impactos negativos, como prejuízos econômicos ao município.
 - É uma doença que pode levar à morte.
 - Intervenções governamentais podem ajudar no controle da propagação da doença.
- Taxa de incidência da dengue: Região Sudeste : 379,4/100.000 habitantes. No estado de São Paulo - 457/100.000 habitantes (Boletim Epidemiológico, 2021).

- Existem diversos trabalhos que consideram a escala interurbana, mas existe uma demanda por trabalhos na escala intraurbana.
- A escala intraurbana nos permite ver a pluralidade do espaço urbano, que é fragmentado.

- Trabalho apresentado no CNMAC 2021 com o objetivo de confrontar dados de mobilidade com séries temporais de casos de dengue no município de São José dos Campos - SP.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Correlação cruzada em séries temporais de casos de dengue - em busca do atraso que maximiza a correlação

Cátia S. N. Sepetushka¹
INPE (Instituto Nacional de Pesquisas Espaciais), São José dos Campos, SP
Livia R. Tomá²
Cemaden, São José dos Campos, SP
Davi Sanchez³
Cemaden/UNIFESP, São José dos Campos, SP
Luciana R. Lotufo⁴
Cemaden, São José dos Campos, SP
Leonardo H. L. Santos⁵
Cemaden, São José dos Campos, SP

Resumo. Medidas de correlação entre séries temporais tem sido aplicadas em diversas temas de pesquisa, dentre elas epidemiologia matemática. A correlação cruzada permite identificar o atraso que maximiza a correlação entre séries temporais, por exemplo, de casos de dengue em diferentes regiões de uma cidade. A distribuição desses atrasos é ainda uma questão em aberto, especialmente na escala intra-urbana, assim como sua relação com questões ambientais e sociais. Neste trabalho, confrontamos tais atrasos com dados de mobilidade urbana para a cidade de São José dos Campos/SP. Os resultados mostram que a medida para as correlações decresce frente à distância geográfica e cresce com a mobilidade de pessoas entre as regiões, e que o atraso máximo (correlação convencional) é o mais frequente entre os que maximizam a correlação.

Palavras-chave. Epidemiologia Matemática, Correlação cruzada, Mobilidade, Dengue.

1 Introdução

Medidas de correlação entre séries temporais tem sido aplicadas em diversos temas de pesquisa, em ordem de aumentar similaridades entre diferentes padrões [1]. Em particular, a área de epidemiologia matemática tem vários exemplos de pesquisas envolvendo tal abordagem [2,3]. Dentre as diferentes métricas de correlação está a correlação cruzada, que permite identificar o atraso que maximiza a correlação entre as séries temporais.

Nos últimos anos, doenças como Dengue, Zika e Chikungunya tem causado diversos impactos negativos em muitos municípios brasileiros, com aumento do número de internações e atendimentos no trabalho, além da dificuldade individual para realização das atividades cotidianas. Preocupação com propagação de epidemias veio desde doenças "reocorrentes", a pandemia mundial, como a atenuação pelo novo Coronavírus, tema de interesse global [4].

¹catsn.kit@gmail.com

²liviatoamas@gmail.com

³davisanchez@ufscar.br

⁴luciana.lotufo@cemaden.gov.br

⁵leonardo@inpe.com.br

- Este trabalho tem como objetivo estudar a distribuição espacial da dengue em SJC no período de 2013 a 2019.
- A Estatística é usada na área de saúde para ajudar a modelar a realidade - entender, estimar e modelar como o problema pode afetar a vida de saúde das pessoas.
- Verificar se existem persistências espaço-temporais nesse intervalo de tempo aplicando a técnica de varredura estatística (ScanStatistics).

1 Introdução

2 Materiais e Métodos

Área de Estudo

Banco de Dados

Método

Experimentos

3 Conclusões

1 Introdução

2 Materiais e Métodos

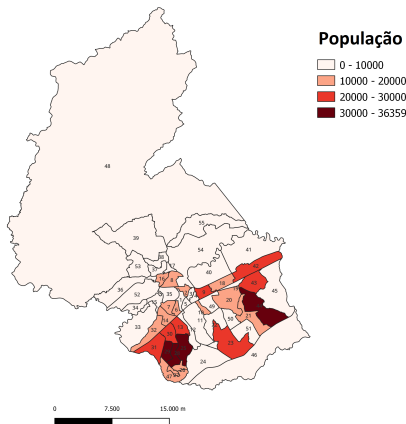
Área de Estudo

Banco de Dados

Método

Experimentos

3 Conclusões



- Cidade de São José dos Campos - região do Vale do Paraíba.
- 90 km da cidade de São Paulo e 320 km do Rio de Janeiro.
- População considerada foi de 629621 habitantes
- Área total de 1.099,4 km², sendo 356 km² (33%) de áreas urbanas e 743.4 km² (67%) de área rural.

Área de Estudo

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

ZT	ZT_Name	ZT	ZT_Name
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 Jaguarý
25	Bosque dos Eucaliptos e Quinta das Flores	53	Chácara Fazenda Caeté
26	Conjunto Habitacional Dom Pedro	54	Águas de Canindé e Chácara Havai
27	Jardim Imperial e Jardim Colonial	55	Chácara Pedra D'água II
28	Bosque dos Ipês e Campo dos Alemães		

1 Introdução

2 Materiais e Métodos

Área de Estudo

Banco de Dados

Método

Experimentos

3 Conclusões

- Dados de casos de dengue foram fornecidos pelo Centro de Zoonoses - Departamento de Saúde - São José dos Campos
- Dados de população do Censo do IBGE de 2010
- Mobilidade - Pesquisa Origem Destino - Atlas de Pesquisa Origem Destino, 2014.

1 Introdução

2 Materiais e Métodos

Área de Estudo

Banco de Dados

Método

Taxas de Incidência

Scan Statistics

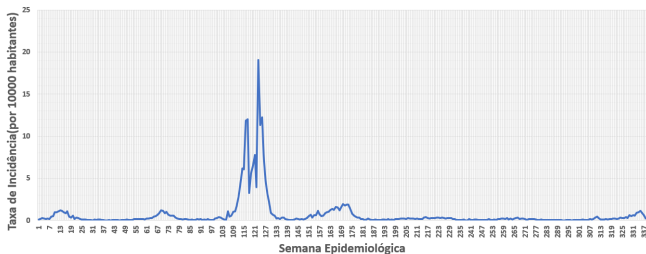
SATScan

SATScan

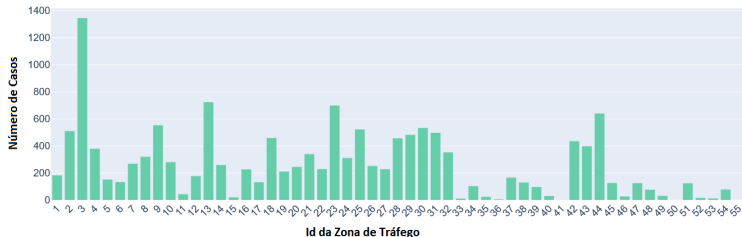
Experimentos

3 Conclusões

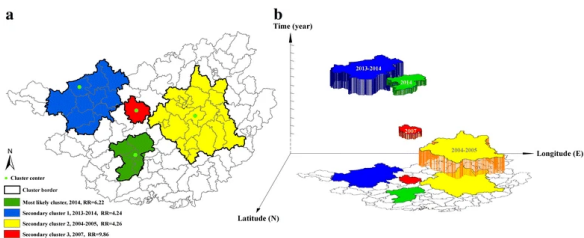
Taxa de Incidência por Semana Epidemiológica entre 2013 e 2019



Casos por Zona de Tráfego



- Aplicar o método de varredura de estatísticas para detecção de clusters espaço-temporais
- Clusters espaciais, temporais ou espaço-temporais



- Ferramenta SATScan

SaTScan™
Software for the spatial, temporal, and space-time scan statistics

Purpose
SaTScan™ is a free software that analyzes time series data for:

- The best geographical surveillance of disease
- Test whether a disease is randomly distributed
- Evaluate the statistical significance of clusters
- Predict the expected time-periodic disease

 The software may also be used for similar problems.

Data Types and Methods
SaTScan uses either a Poisson-based model, a permutation model, using only case data, an epidemic tract, zip code, county or other geographic grid provided by the user, as well as for temporal time series.

Developers and Funders
The SaTScan™ software was developed by:

- National Cancer Institute, Division of Cancer Control and Prevention, Division of Cancer Prevention and Control
- Alfred P. Sloan Foundation, through grant
- Centers for Disease Control and Prevention
- National Institute of Child Health and Human Development
- National Cancer Institute, Division of Cancer Control and Prevention
- National Institute of Clinical Medical Toxicology
- NIH, through grant to the

 These financial support is greatly appreciated. The

Input - Analysis - Output

Case File: [Browse] Time Period: [None] [Year] [Month] [Day]
 Control File: (Bernoulli Model) [Browse] [None] [Year] [Month] [Day]
 Study Period: [Start Date] [Year] [Month] [Day] [End Date] [Year] [Month] [Day]
 Population File: (Poisson Model) [Browse]
 Case-Cluster File: [Browse]
 Grid File: (optional) [Browse]

Display SaTScan Variables For: discrete Poisson model

SaTScan Variable	Source File Variable
Number of Cases	SAIT
Year (Time optional)	Year
Coverage1 (optional)	unassigned
Coverage2 (optional)	unassigned
Coverage3 (optional)	unassigned
Coverage4 (optional)	unassigned

Generated ID # One Count # DAY ID dengue
 location1 1 1 1 1
 location2 1 1 1 1
 location3 1 1 1 1
 location4 1 1 1 1

Fonte: <https://www.satscan.org/>.

- Dados de entrada

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	1556

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

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	-4582859211	-2319077216

1 Introdução

2 Materiais e Métodos

Área de Estudo

Banco de Dados

Método

Experimentos

3 Conclusões

- Experimento 1 - Considerar uma população de no máximo 50% para cada cluster encontrado.
- Experimento 2 - Considerar uma população de no máximo 10% para cada cluster encontrado.

1 Introdução

2 Materiais e Métodos

3 Conclusões

Resultados

Perspectivas Futuras

1 Introdução

2 Materiais e Métodos

3 Conclusões

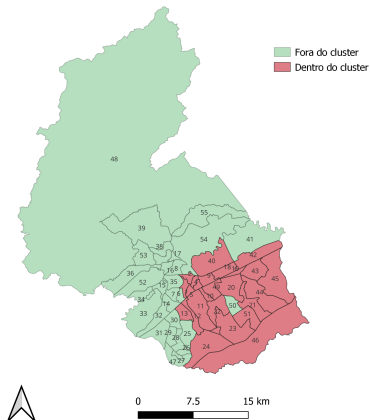
Resultados

Experimento 1

Experimento 2

Perspectivas Futuras

- Foi detectado 1 cluster Espaço-Temporal



• Estatísticas geradas:

Program run on: Fri Dec 17 17:31:10 2021

Retrospective Space-Time analysis
scanning for clusters with high rates
using the Discrete Poisson model.

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

$$RR = (cz/Ez)/((C-cz)/(C-Ez))$$

$$\lambda_z = \left(\frac{c_z/n_z}{C/N}\right) c_z \left(\frac{c_{z,t}/n_{z,t}}{C/N}\right) c_{z,t}$$

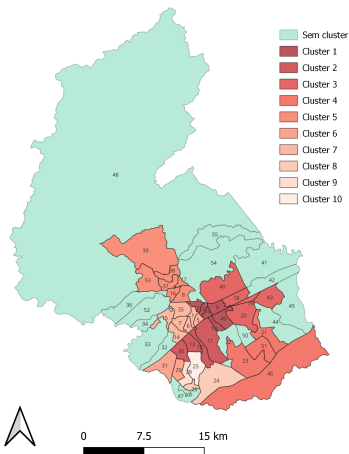
C - Casos no mapa
ni - Casos numa ZT
N - População Total do mapa
cz - número de casos no cluster
Ez - [C * nz]/N
nz - População no cluster

PARAMETER SETTINGS

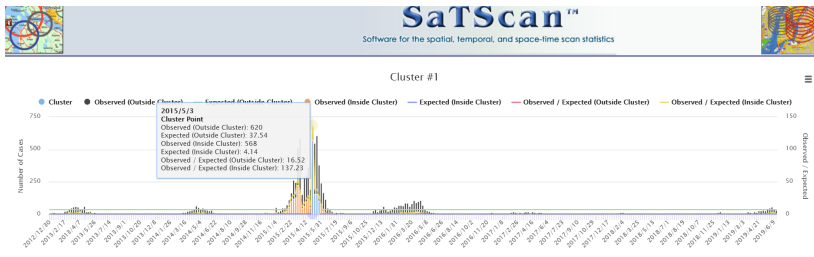
Input

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

- Foram detectados 11 clusters Espaço-Temporais



- Série temporal



- Estatísticas geradas:

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.: 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

1 Introdução

2 Materiais e Métodos

3 Conclusões

Resultados

Perspectivas Futuras

- Esse foi um trabalho descritivo. Temos regiões não-uniformes no espaço e diferentes entre elas.
- Como trabalhos futuros pretendemos relacionar os clusters identificados pelo SATScan com grafos de mobilidades entre as zonas e verificar se existem correlações entre os nós no grafo e as zonas no cluster.
- Fazer outras combinações de janelas espaço-tempo.
- Justificar cada cluster com base em informações complementares, não relativas ao número de casos. Ex: Mobilidade, Dados pluviométricos.

Obrigada!