



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

## **ANÁLISE DE UMA KRIGAGEM ESPAÇO-TEMPORAL EM UM ESTUDO DE CASO COM DADOS DE RADAR METEOROLÓGICO**

Aurelienne A. Souza Jorge

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

INPE  
São José dos Campos  
2019

# Sumário

<b>1. Introdução</b>	<b>2</b>
<b>2. Materiais e métodos</b>	<b>3</b>
2.1 Estudo de Caso	3
2.2 Dados e pré-processamento	6
2.3 Análise do semivariograma das amostras	8
2.4 Geração das previsões	10
2.4.1 Geração e ajustes do modelo de variograma	11
2.4.2 Krigagem espaço-temporal	14
2.4.3 Avaliação dos resultados	14
<b>3. Resultados</b>	<b>14</b>
<b>4. Conclusões e Perspectivas</b>	<b>16</b>
<b>Referências</b>	<b>17</b>

## Lista de Figuras

- Figura 1. Regiões afetadas pelas fortes chuvas em março/2019
- Figura 2. A região metropolitana de São Paulo e a bacia hidrográfica do rio Tamanduateí
- Figura 3. Precipitação total na Bacia do Tamanduateí durante o período do evento em estudo
- Figura 4. A cobertura do radar de São Roque e a região metropolitana de São Paulo
- Figura 5. Série temporal com dados de refletividade de radar para a área da bacia do Tamanduateí, no período de 10/03/2019 23h até 11/03/2019 03h (GMT)
- Figura 6. Exemplo de semivariograma experimental espacial
- Figura 7. Mapa do Semivariograma espaço-temporal para as amostras do estudo de caso
- Figura 8. Visualização em 3D do semivariograma espaço-temporal para as amostras do estudo de caso
- Figura 9. Metodologia implementada no R Studio.
- Figura 10. Variograma das amostras no *timelag* inicial
- Figura 11. Primeiro conjunto de amostras (à esquerda) e respectivas previsões geradas (à direita)
- Figura 12. Erros calculados para as previsões de 10 e 20 minutos
- Figura 13. Mapa do Erro Absoluto Médio acumulado

# 1. Introdução

Eventos extremos de precipitação podem deflagrar, por muitas vezes, ocorrências de ameaças como enxurradas, deslizamentos de terra e alagamentos, gerando sérios impactos econômicos e sociais nas áreas atingidas. Diversas regiões da grande São Paulo sofreram com esses efeitos após uma forte chuva que se iniciou no dia 10 de março deste ano e se estendeu até a madrugada do dia seguinte. Foram relatados doze óbitos e ao menos seis pessoas feridas em decorrência de alagamentos e deslizamentos de terra. [1]

Visando monitorar a ocorrência desses eventos extremos de precipitação, e antever as possíveis consequências, ferramentas de previsão a curtíssimo prazo têm sido desenvolvidas, e muitas delas utilizam dados de radares meteorológicos como base, por disporem de alta resolução espacial e temporal. [2]

O trabalho de Caseri et al. [3] propõe um método geoestatístico que gera uma simulação de campos de precipitação, com base em varreduras de radar e dados observacionais pluviométricos. O algoritmo, desenvolvido no Instituto de Pesquisa Francês da Ciência e Tecnologia para o Meio Ambiente e Agricultura (Irstea), simula chuva no espaço  $(x,y)$  e tempo  $(z)$  usando um modelo de variograma espaço-temporal aliado à uma técnica de advecção de direção e velocidade.

Ainda no âmbito geoestatístico, existem ferramentas como a krigagem espaço-temporal, também embasada em modelos de variograma, que através de uma interpolação nos domínios do espaço e tempo, possibilita que previsões sejam feitas para dadas localizações, visando os intervalos, ou até mesmo além, dos instantes observados em uma dada série temporal. [4]

A proposta do presente trabalho consiste em gerar e analisar simulações do deslocamento de sistemas de precipitação, aplicando a técnica da krigagem espaço-temporal em dados de varredura de radar meteorológico.

## 2. Materiais e métodos

### 2.1 Estudo de Caso

Um caso mais recente de evento extremo de precipitação acometeu várias áreas na região metropolitana de São Paulo em março deste ano (2019), com chuvas que iniciaram na noite do dia 10 e se estenderam até a madrugada do dia seguinte. A Figura 1 apresenta as regiões mais afetadas e os óbitos decorrentes de alagamentos e deslizamentos de terra.

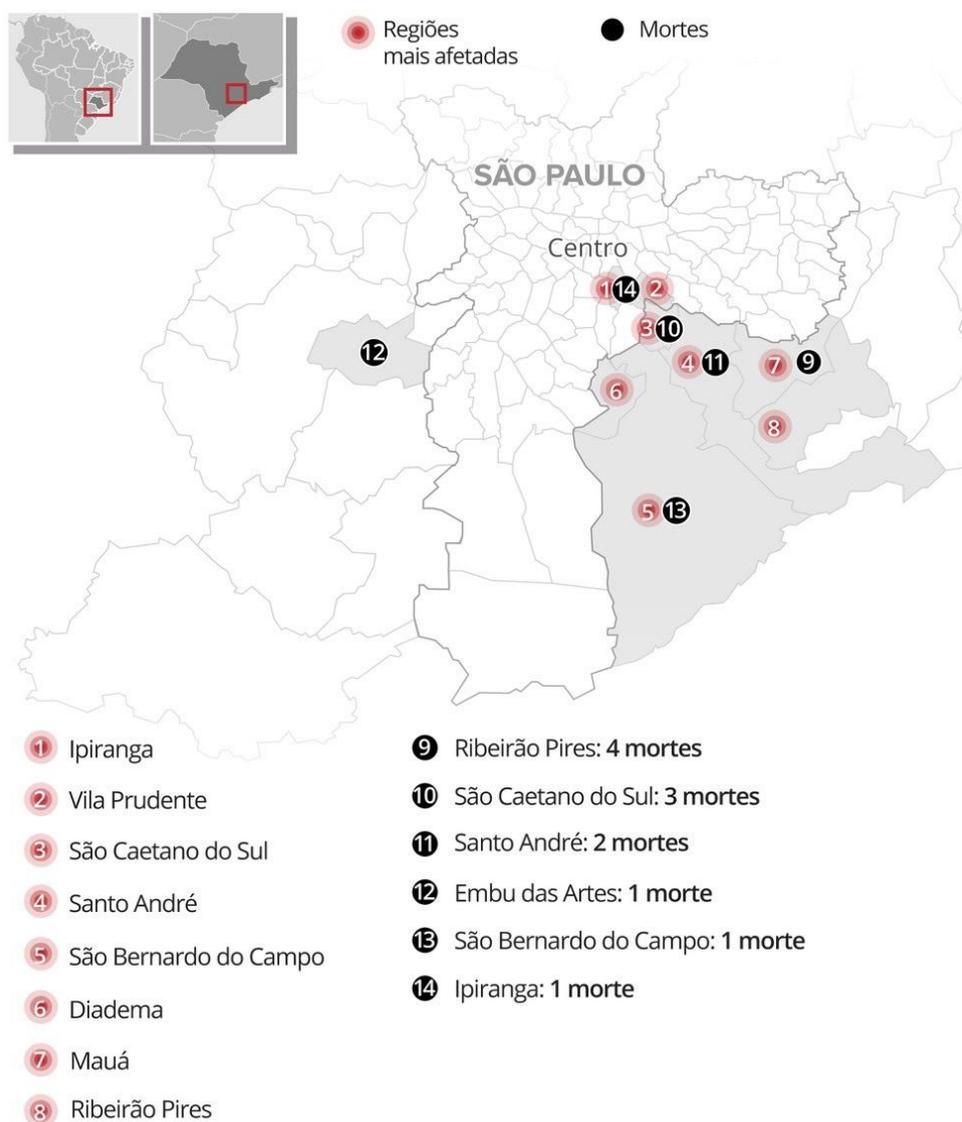
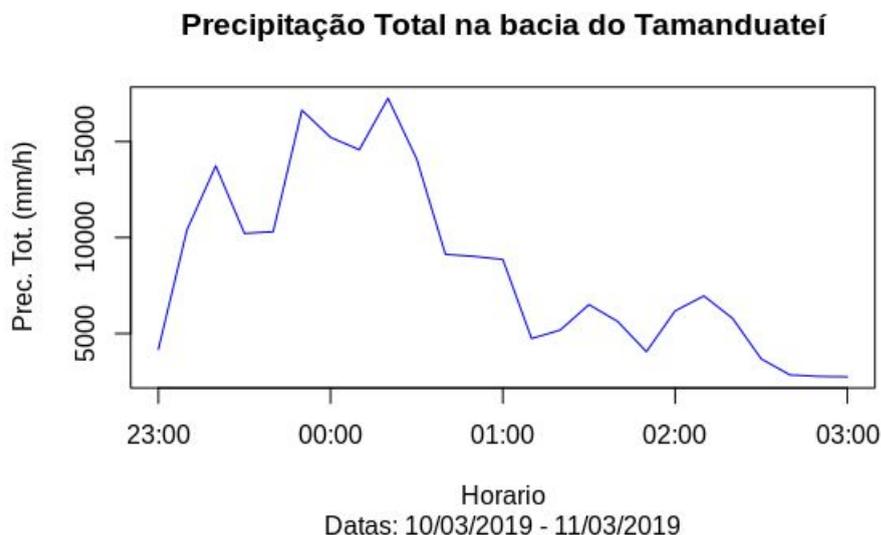


Figura 1. Regiões afetadas pelas fortes chuvas em março/2019

Fonte: G1





*Figura 3. Precipitação total na Bacia do Tamanduateí durante o período do evento em estudo*

## 2.2 Dados e pré-processamento

Por fornecerem alta resolução temporal e espacial, dados de radar meteorológico foram utilizados como base para o presente trabalho. O radar de São Roque, operado pelo Departamento de Controle do Espaço Aéreo, foi o escolhido por possuir boa cobertura da área de estudo definida, conforme ilustra a Figura 4. Os dados oriundos de suas varreduras volumétricas contam com resolução espacial de 1 quilômetro e temporal de 10 minutos. [6]

Dentre os diversos dados que podem ser obtidos com as varreduras dos radares meteorológicos, o produto denominado CAPPI (Constant Plan Position Indicator) na altitude de 3 quilômetros foi o adotado por fornecer os dados numa projeção em um plano horizontal constante pré-definido [7].

Os arquivos encontram-se no formato binário e contém valores de refletividade (ecos de retorno dos alvos) numa grade, que podem ser convertidos em taxas de precipitação estimada. Para o presente trabalho, foi necessário realizar um recorte espacial e temporal conforme delimitação do estudo de caso adotado.

A Figura 5 apresenta a série temporal do evento na área da bacia do Tamanduateí utilizando-se dos dados pré-processados. Os valores apresentados estão em escala de refletividade.

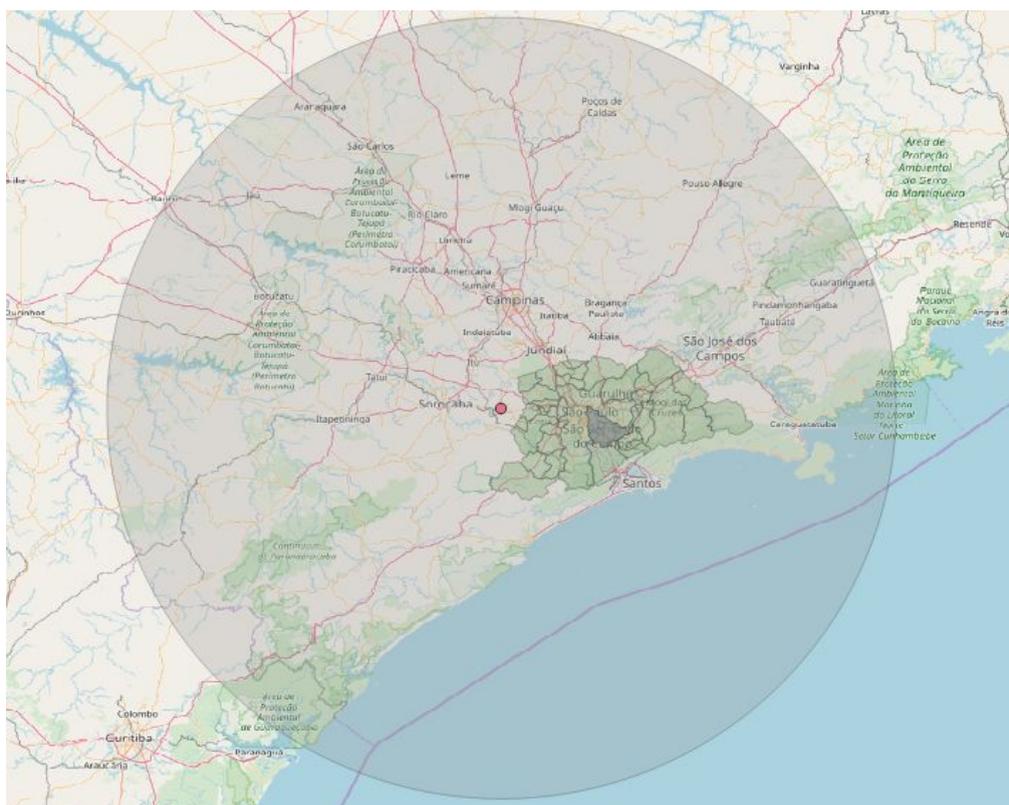


Figura 4. A cobertura do radar de São Roque e a região metropolitana de São Paulo

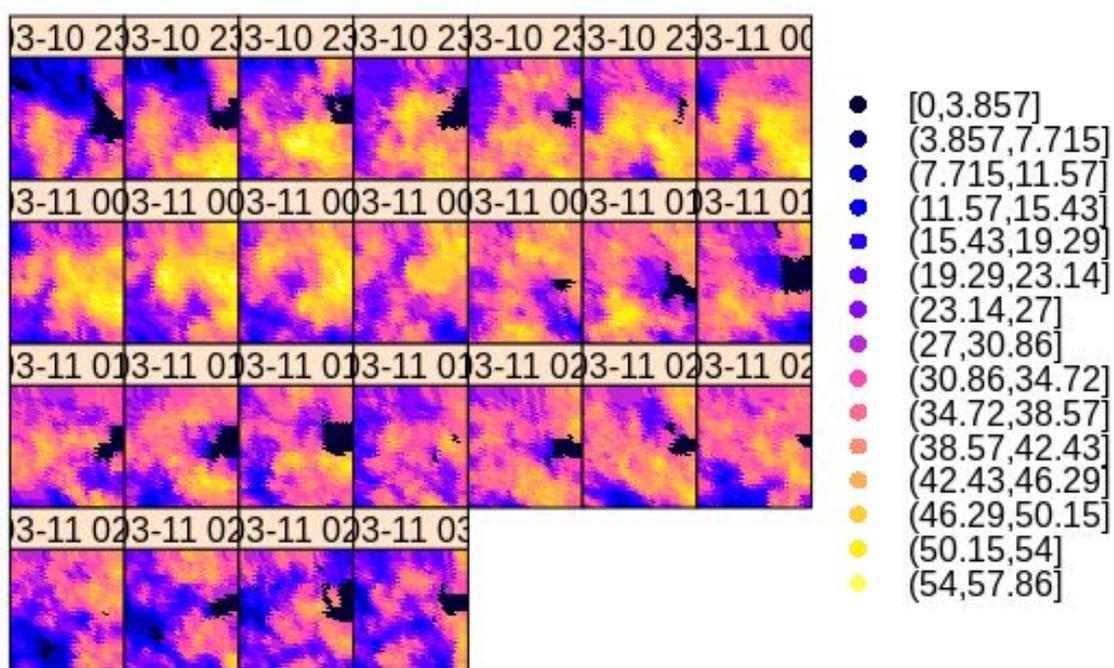


Figura 5. Série temporal com dados de refletividade de radar para a área da bacia do Tamanduateí, no período de 10/03/2019 23h até 11/03/2019 03h (GMT)

## 2.3 Análise do semivariograma das amostras

Um semivariograma, ou variograma, é um ferramenta da geoestatística que permite representar quantitativamente a variação de um dado fenômeno regionalizado no espaço. [8] A Figura 6 contém um exemplo de variograma experimental cujo padrão representa o que se espera, intuitivamente, de dados amostrais de campo, isto é, que as variâncias ( $\hat{\gamma}$ ) entre as amostras decresçam à medida que a distância ( $h$ ) que as separa também decresce, indicando que as observações mais próximas geograficamente tenham maior similaridade entre si do que as que se encontram separadas por maiores distâncias. Na mesma figura, podem também ser observados os parâmetros do semivariograma, descritos a seguir. [9]

- Alcance ( $a$ ): distância dentro da qual nota-se uma correlação espacial entre as amostras. No exemplo, o alcance é próximo de 25 metros; depois deste ponto, as variâncias apresentam independência em relação à variação das distâncias.
- Patamar ( $C$ ): é o valor da semivariância correspondente ao alcance ( $\hat{\gamma}(a)$ ).
- Efeito pepita ( $C_0$ ): idealmente igual a zero, porém, na prática, à medida que  $h$  tende a zero,  $\hat{\gamma}(h)$  tende a um valor positivo, que revela a descontinuidade do semivariograma para distâncias menores do que a menor distância entre as amostras; pode ser decorrente de erros de medição ou da variabilidade de pequena escala não captada pela amostragem realizada.
- Contribuição ( $C_1$ ): diferença entre os valores do patamar e do efeito pepita.

O pacote R dispõe de uma função para o cálculo de variogramas espaço-temporais a partir de amostras, que além de fornecer as semi-variâncias de acordo com os intervalos de distância  $h$  (*space lags*), considera também diferentes janelas de tempo  $t$  (*time lags*) dentro da série temporal das amostras. Para a geração de variogramas em um dado intervalo de distância ( $h$ ) e intervalo tempo ( $t$ ), o variograma pseudo-cruzado  $(Z_i(s) - Z_{i+t}(s+h))^2$  é calculado sobre a média de todos os conjuntos de observações  $Z_i$  e  $Z_{i+t}$  disponíveis (ponderando em relação ao número de pares envolvidos). [10]

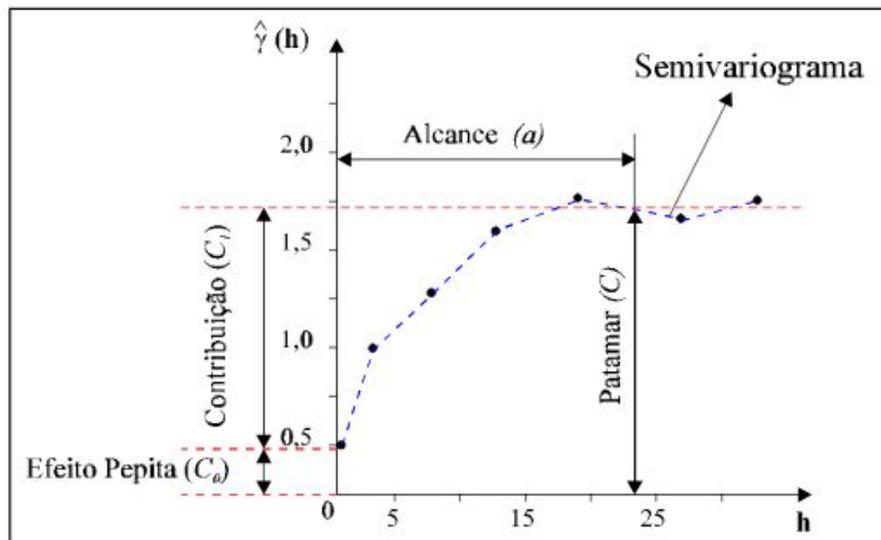


Figura 6. Exemplo de semivariograma experimental espacial  
Fonte: Camargo [9]

As Figuras 7 e 8 apresentam o variograma espaço-temporal calculado sobre as amostras do referido estudo de caso, através da função *variogramST* do pacote R, em diferentes formatos de visualização. Para a geração do variograma, o *time lag* foi incrementado de 10 em 10 minutos e o *space lag* de 1 em 1 km, de forma a seguir a resolução temporal e espacial das amostras.

Observando o mapa do variograma na Figura 7, pode-se notar que a maior dependência espaço-temporal (menores valores de semivariância) encontra-se nos *time lags* de 0 a 40 minutos e nos *space lags* de 0 a aproximadamente 12 km. Pode-se observar também que há alguma similaridade entre as observações realizadas num intervalo de tempo de 100 a 150 minutos dentro de curtas distâncias (aproximadamente 6 km). Mas de forma geral, pode-se dizer que as variâncias tendem a aumentar conforme também crescem as janelas temporais e espaciais consideradas.

Buscando identificar alguma tendência direcional na dependência espacial, foram geradas várias versões do variograma experimental alterando-se o parâmetro *alpha*, que indica a direção na qual os pares de amostras serão agrupados no plano espacial. Porém, não foi identificada uma tendência direcional significativa, e portanto, o parâmetro *alpha* foi desconsiderado, gerando então um variograma experimental omnidirecional.

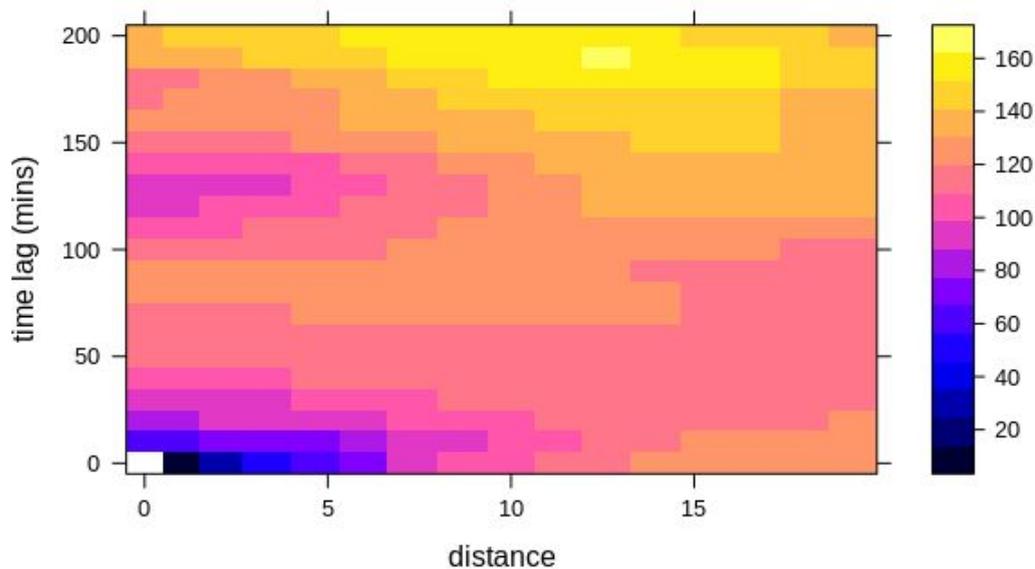


Figura 7. Mapa do Semivariograma espaço-temporal para as amostras do estudo de caso

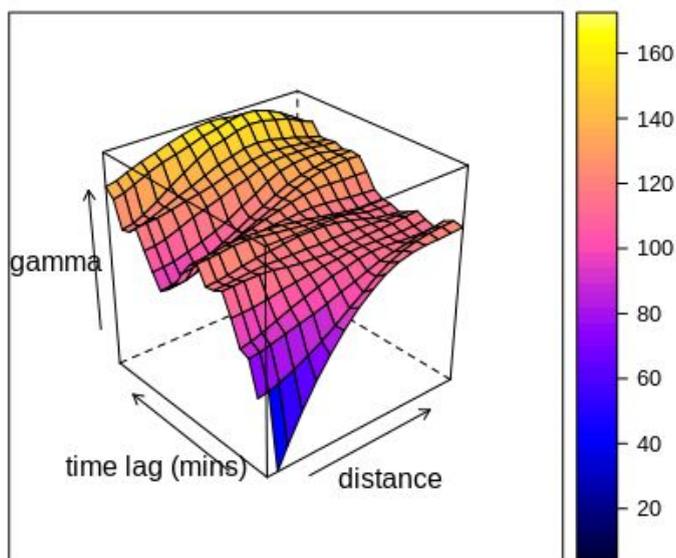


Figura 8. Visualização em 3D do semivariograma espaço-temporal para as amostras do estudo de caso

## 2.4 Geração das previsões

Com base nas análises realizadas no semivariograma das amostras, a metodologia adotada para a geração das previsões foi definida conforme fluxograma apresentado na Figura 9. O algoritmo foi todo implementado no R Studio, fazendo

uso de várias de suas bibliotecas, com destaque para a *gstat*, que dispõe de todas as funções geoestatísticas necessárias para o projeto.

Como notou-se uma maior dependência temporal num intervalo de até 40 minutos, a proposta foi definida de forma a gerar previsões para os próximos 10 e 20 minutos, a partir de um modelo ajustado com base nos últimos 40 minutos de dados observados. Toda a série temporal das amostras é percorrida para a geração das previsões, sempre com base nos últimos 5 campos amostrais. Dessa forma, como são 4 horas de duração do evento (10/03 23h a 11/03 03h), a primeira previsão de 10 e 20 minutos é feita, respectivamente, para as 23h50 do dia 10/03 e 00h00 do dia 11/03, com base nos dados de 23h00 até 23h40; e assim segue de 10 em 10 minutos, até que a última previsão de 20 minutos seja feita para as 3h00 do dia 11/03 (último dado observado).

É importante ressaltar que o modelo é sempre ajustado a cada amostra de 40 minutos de dados observados. E com base nesse modelo ajustado é feita a previsão dos próximos 10 e 20 minutos através de uma krigagem espaço-temporal. A construção e ajustes do modelo utilizado estão descritos na seção 2.4.1. A metodologia adotada para avaliação dos resultados gerados, é discutida na seção 2.4.2.

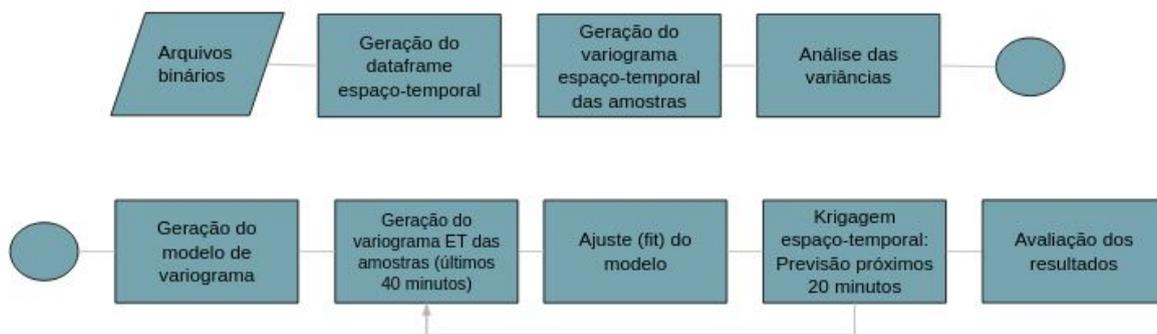


Figura 9. Metodologia implementada no R Studio.

### 2.4.1 Geração e ajustes do modelo de variograma

O pacote “*gstat*” do R disponibiliza a função “*vgmST*” para a geração de modelos de variogramas espaço-temporal. Dentre as opções de modelos disponíveis, a “*sumMetric*” foi a escolhida por ser a mais flexível, permitindo compor o modelo final com base em uma combinação de 3 modelos: um espacial, um temporal e um

terceiro que representa a junção do espaço e tempo (*joint*). Além disso, é possível informar um fator de anisotropia na construção do modelo.

Para a construção de cada um desses 3 modelos, foi utilizada a função “*vgm*” do pacote *gstat* que requer basicamente 4 parâmetros principais: *sill* (patamar), *range* (alcance), *nugget* (efeito pepita) e tipo de modelo (exponencial, esférico, gaussiano, linear, etc.). [11]

A definição dos parâmetros para o modelo de variograma espacial foi feita com base no variograma das amostras no *timelag* inicial, ilustrado na Figura 10 a seguir, que apresenta um alcance de aproximadamente 14 km, um patamar de 120, e uma curva que aparentemente se aproxima de uma função exponencial. O efeito pepita foi desconsiderado. Dessa maneira, o modelo de variograma espacial foi definido da seguinte forma: **`vgm(psill=120,"Exp", range=14, nugget=0)`**

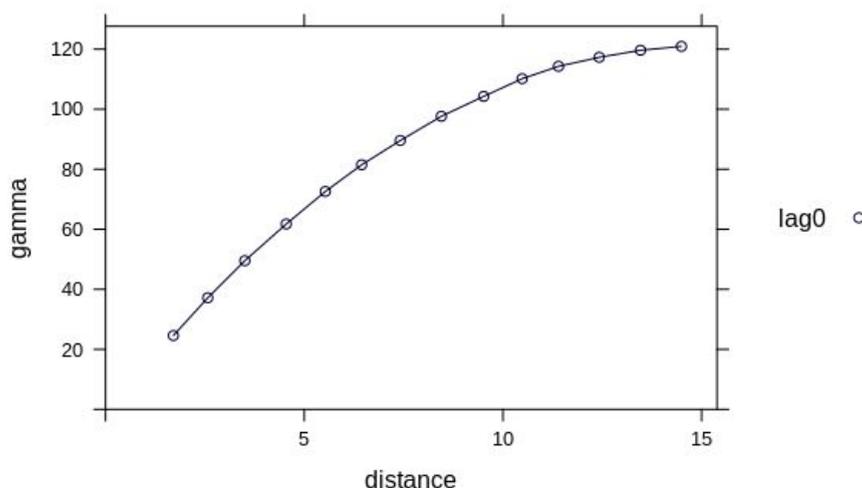


Figura 10. Variograma das amostras no timelag inicial

O *range* do modelo temporal (Time) foi definido em 40 minutos, pois é o intervalo de subagrupamento das amostras seguindo a dependência temporal observada. Da mesma forma, o *range* do modelo de junção espaço-temporal (Joint) também foi limitado a 40. Já o parâmetro *sill* de ambos os modelos foi definido após algumas tentativas com combinações de valores, optando-se pela combinação com menor Erro Médio Quadrático (MSE - Mean Square Error) considerando-se a média entre todas as subamostras de 40 minutos e o respectivo modelo ajustado. Foram feitas algumas tentativas também utilizando o modelo esférico ao invés do exponencial,

porém, resultaram em erros relativamente maiores. A Tabela 1 apresenta as combinações de parâmetros testadas e o respectivo erro resultante, sendo a linha destacada em azul a que contém a combinação com menor erro, e portanto, a que foi adotada na construção do modelo final proposto.

Space			Time			Joint			MSE (média das subamostras)
Função	Sill	Range	Função	Sill	Range	Função	Sill	Range	
Exp	120	14	Exp	110	40	Exp	20	40	93,7
Exp	120	14	Exp	110	40	Exp	30	40	93,36
Exp	120	14	Exp	110	40	Exp	50	40	80,32
Exp	120	14	Exp	110	40	Exp	70	40	80,15
Exp	120	14	Exp	110	40	Exp	100	40	80,11
Exp	120	14	Exp	50	40	Exp	100	40	70,24
Exp	120	14	Exp	50	40	Exp	90	40	<b>69,56</b>
Exp	120	14	Exp	50	40	Exp	80	40	71,75
Exp	120	14	Exp	40	40	Exp	100	40	72,87
Exp	120	14	Exp	50	40	Exp	50	40	78,79
Exp	120	14	Exp	80	40	Exp	100	40	76,56
Exp	120	14	Exp	60	40	Exp	80	40	71,95
Exp	120	14	Exp	50	40	Exp	40	40	75,68
Sph	120	14	Exp	50	40	Exp	90	40	71,37
Sph	120	14	Sph	50	40	Exp	90	40	71,63

Tabela 1. Combinação de parâmetros testados e os respectivos valores médios de MSE

O modelo final de variograma espaço-temporal é apresentado a seguir. O parâmetro “stAni”, que limita o alcance da anisotropia a ser aplicada, foi definido com o mesmo valor do alcance do modelo espacial (14 km), com o intuito de indicar que a anisotropia está presente em toda a área considerada como espacialmente correlacionada.

```
sumMetric <- vgmST("sumMetric",
  space = vgm(psill=120,"Exp", range=14, nugget=0),
  time = vgm(psill=50, "Exp", range=40, nugget=0),
  joint = vgm(psill=90, "Exp", range=40, nugget=0),
  stAni=14)
```

### 2.4.2 Krigagem espaço-temporal

O pacote *gstat* do R também fornece a função “krigeST”, que possibilita a interpolação espaço-temporal com base em um modelo de variograma, também definido nos domínios do espaço e do tempo.

Através da linha de código demonstrada abaixo, é feita uma krigagem ordinária com base em um conjunto de amostras (“data”), aplicando-se um modelo de variograma espaço-temporal (“modelList”), e gerando um conjunto de dados com a simulação realizada nas localizações e instantes temporais desejados (“newdata”).

```
pred <- krigeST(bindata~1, data=stf_sample, modelList=sumMetric_Vgm,  
newdata=new_stf)
```

### 2.4.3 Avaliação dos resultados

Para avaliação do acerto das previsões, foram calculados o Erro Absoluto Médio (MAE - Mean Absolute Error) e a Raiz do Erro Médio Quadrático (RMSE - Root Mean Square Error) entre a previsão simulada (próximos 10 e 20 minutos) e o dado real do radar correspondente ao instante da previsão. Ambas as métricas de erro expressam o erro médio na estimativa da variável de interesse, porém o RMSE, por elevar os erros ao quadrado no cálculo, acaba atribuindo um peso mais alto a erros maiores.

Como base de comparação, para avaliar a significância dos erros calculados, foi adotado um modelo de persistência de chuva, seguindo a metodologia adotada por Caseri et al. [3], na qual o último dado observado é replicado como previsão para os próximos 10 e 20 minutos, simulando a persistência da chuva nesse período.

Os resultados dessa avaliação são apresentados na seção 3, a seguir.

## 3. Resultados

A Figura 11 apresenta espacialmente o primeiro subconjunto de amostras, compreendendo o período de 23h00 até 23h40 do dia 10 de março, e as previsões estimadas para os próximos 10 e 20 minutos, respectivamente, 23h50 de 10 de março e 00h00 de 11 de março.

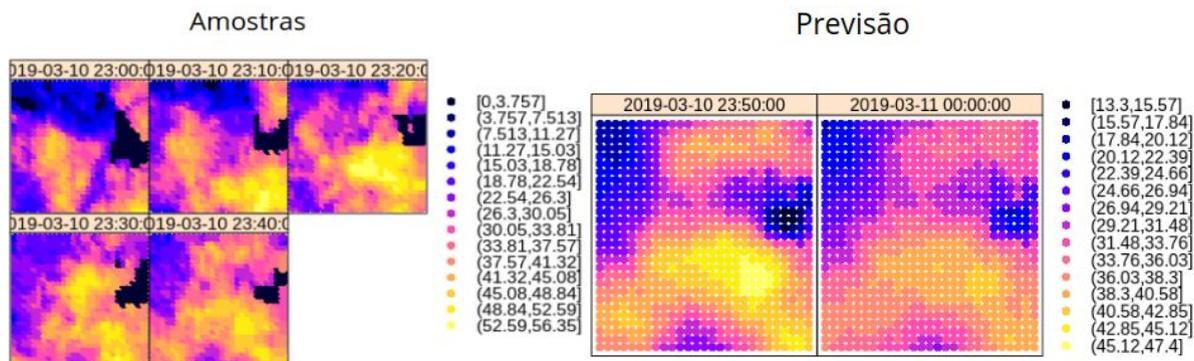


Figura 11. Primeiro conjunto de amostras (à esquerda) e respectivas previsões geradas (à direita)

Conforme a metodologia de avaliação dos resultados definida e descrita na seção 2.4.3, foram calculados o Erro Absoluto Médio e a Raiz do Erro Médio Quadrático de cada previsão gerada em relação aos dados reais do radar no instante de tempo correspondente. Como base de comparação, também foram calculados os erros em uma modelo de precipitação persistida. Os resultados são apresentados nos gráficos da Figura 12. Pode-se observar que, para praticamente 100% das previsões geradas, o erro médio foi inferior em comparação ao erro inerente no método de persistência da precipitação. Sendo esse resultado apontado tanto pelos valores do Erro Médio Absoluto, quanto pela Raiz do Erro Médio Quadrático.

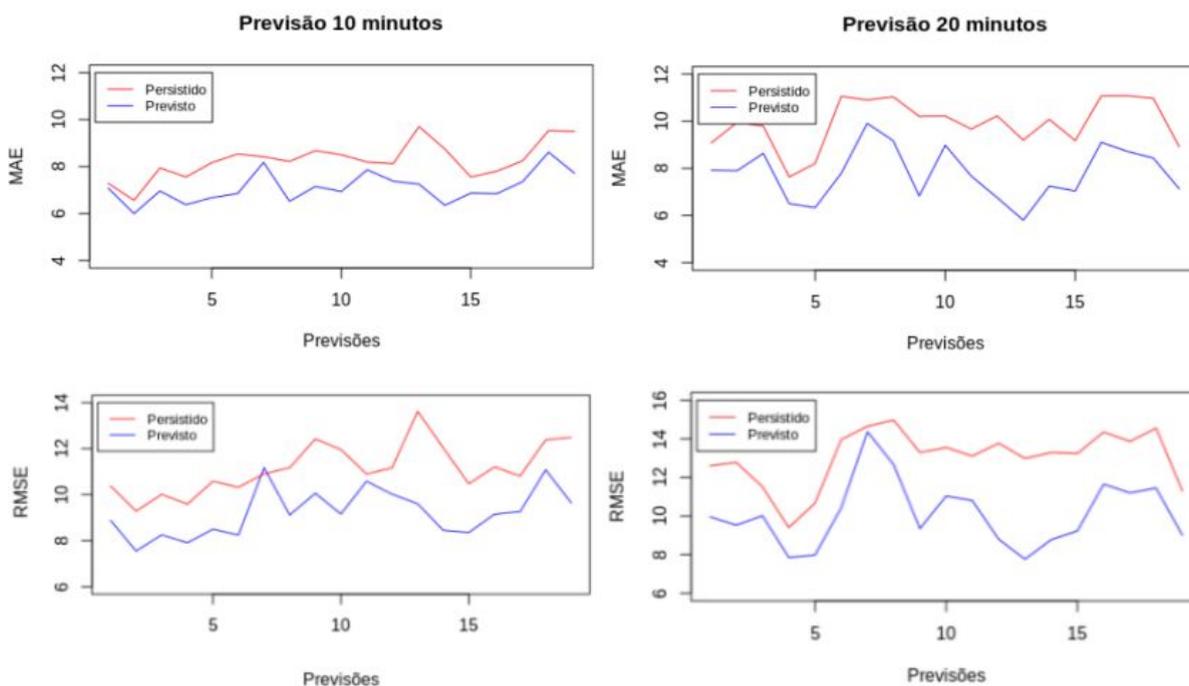


Figura 12. Erros calculados para as previsões de 10 e 20 minutos (acima, o erro médio absoluto e, abaixo, a raiz do erro médio quadrático). As linhas em azul representam os erros dos dados previstos e as linhas em vermelho representam os erros dos dados persistidos.

A Figura 13 contém o mapa do Erro Absoluto Médio acumulado para todas as previsões de 10 minutos simuladas, no intuito de indicar espacialmente como os erros ficaram distribuídos. Apesar de as previsões apresentarem um erro aparentemente significativo, nota-se visualmente que o modelo de precipitação persistida obteve erros distribuídos em maiores áreas.

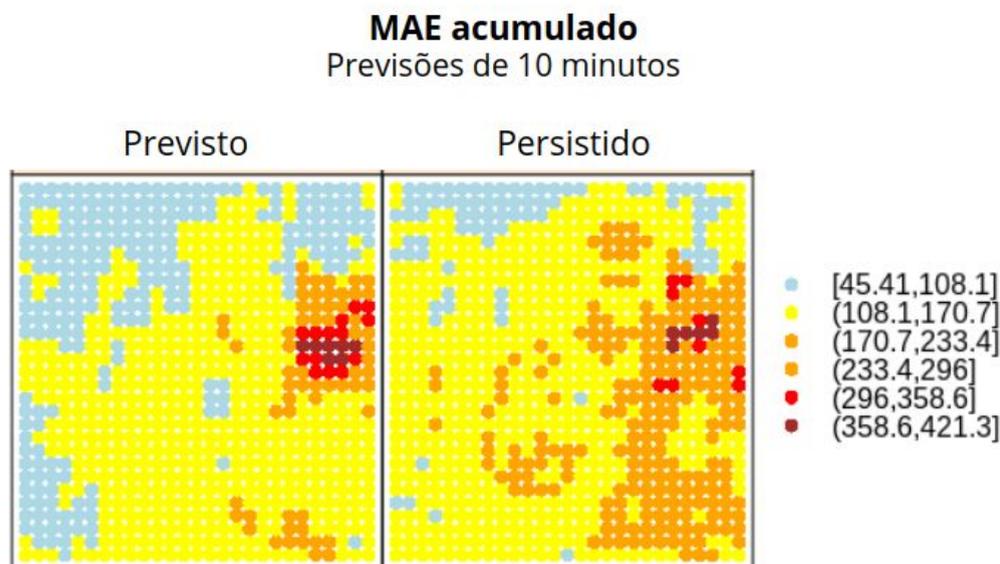


Figura 13. Mapa do Erro Absoluto Médio acumulado para todas as previsões de 10 minutos simuladas

## 4. Conclusões e Perspectivas

O semivariograma espaço-temporal mostrou-se uma ferramenta importante na análise das relações espaciais e temporais dos dados, podendo indicar os intervalos temporais e espaciais com maior autocorrelação.

Apesar da dificuldade na geração de previsões com boa precisão espacial e temporal, devido à natureza caótica do sistema atmosférico, a krigagem, como ferramenta para gerar estimativas nos domínios do espaço e do tempo, obteve algum sucesso na comparação dos resultados com um modelo de precipitação persistida.

Em trabalhos futuros, há o interesse da aplicação da metodologia em outros estudos de caso, até mesmo fora da área da meteorologia, para comparação de resultados e aprofundamento na aplicação das ferramentas.

## Referências

- [1] Forte chuva isola cidades, provoca alagamentos e desabamentos e deixa 12 mortos na Grande SP. G1, 2019. Disponível em: <https://g1.globo.com/sp/sao-paulo/noticia/2019/03/11/forte-chuva-isola-areas-na-grande-sao-paulo-e-causa-alagamentos.ghtml>. Acesso em: 10 de nov. de 2019.
- [2] Smith J. A., Baeck M. L., Meierdiercks K. L., Miller A. J., and Krajewski W. F. (2007). Radar rainfall estimation for flash flood forecasting in small urban watersheds. *Advances in Water Resources*, 30, 2087-2097.
- [3] A. Caseri, M. Ramos, P. Javelle, E. Leblois, and R. M-h, “A space-time geostatistical approach for ensemble rainfall nowcasting”, *EDP Sci.*, vol. 7, p. 18001, 2016.
- [4] Graler, B., Pebesma, E., Heuvelink, G. “Spatio-Temporal Interpolation using gstat.”, 2016. Disponível em: <https://cran.r-project.org/web/packages/gstat/vignettes/spatio-temporal-kriging.pdf>. Acesso em: 15 de nov. de 2019
- [5] Ramalho, D. RIO TAMANDUATEÍ – NASCENTE À FOZ: PERCEPÇÕES DA PAISAGEM E PROCESSOS PARTICIPATIVOS. Disponível em: [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=19&ved=2ahUKewi4\\_K7HksjmAhVJErkGHXxPCJMQFjASegQIBBAC&url=https%3A%2F%2Fwww.revistas.usp.br%2Fpaam%2Farticle%2Fdownload%2F85723%2F88496%2F&usg=AOvVaw02ilNWXS1682Kyl\\_hFFtP](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=19&ved=2ahUKewi4_K7HksjmAhVJErkGHXxPCJMQFjASegQIBBAC&url=https%3A%2F%2Fwww.revistas.usp.br%2Fpaam%2Farticle%2Fdownload%2F85723%2F88496%2F&usg=AOvVaw02ilNWXS1682Kyl_hFFtP). Acesso em: 01 de dez. de 2019.
- [6] Departamento de Controle do Espaço Aéreo - Comando da Aeronáutica. MCA 105-13: “Manual de Procedimentos Operacionais do Radar Meteorológico”. 2017. Disponível em: [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&ved=2ahUKEwj-xYmbu8nmAhUSHLkGHR8dCPAQFjAEegQIBxAC&url=https%3A%2F%2Fpublicacoes.decea.gov.br%2Fdownload.cfm%3Fd%3D4638&usg=AOvVaw219VrHsZPWQVVwkd\\_SW-v2](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&ved=2ahUKEwj-xYmbu8nmAhUSHLkGHR8dCPAQFjAEegQIBxAC&url=https%3A%2F%2Fpublicacoes.decea.gov.br%2Fdownload.cfm%3Fd%3D4638&usg=AOvVaw219VrHsZPWQVVwkd_SW-v2). Acesso em: 01 de dez. de 2019

- [7] REDEMETS. Como identificar os produtos dos radares meteorológicos. Disponível em: <https://www.redemet.aer.mil.br/?i=blog&id=2390>. Acesso em: 05 de dez. de 2019.
- [8] HUIJBREGTS, C.J. (1975) -Regionalized variables and quantitative analysis of spatial data. In: DAVIS, J.C. & MC CULLAGH, M. J. (ed.) Display and analysis of spatial data. John Wiley, p.38 -53.
- [9] Camargo, E. C. G. Geoestatística: Fundamentos e Aplicações. Geoprocessamento em Projetos Ambientais, cap.5. Disponível em: [http://www.dpi.inpe.br/gilberto/tutoriais/gis\\_ambiente/5geoest.pdf](http://www.dpi.inpe.br/gilberto/tutoriais/gis_ambiente/5geoest.pdf). Acesso em: 20 de nov. de 2019.
- [10] RDocumentation. VariogramST. Disponível em: <https://www.rdocumentation.org/packages/gstat/versions/2.0-3/topics/variogramST>. Acesso em: 20 de nov. de 2019.
- [11] RDocumentation. Vgm. Disponível em: <https://www.rdocumentation.org/packages/gstat/versions/2.0-3/topics/vgm>. Acesso em: 20 de nov. de 2019.