

ROTINAS PARA ANÁLISE GEOESTATÍSTICA UTILIZANDO A LINGUAGEM R: UM EXEMPLO COM DADOS PLUVIOMÉTRICOS DA BACIA HIDROGRÁFICA SOROCABA E MÉDIO TIETÊ

Andréa Pavan Perin

andrea.perin@fatec.sp.gov.br

Elias Silva de Medeiros

eliasmedeiros@ufgd.edu.br

José Carlos Ferreira

josé.ferreira35@fatec.sp.gov.br

Guilherme Barbosa de Oliveira

guilherme.oliveira@fatec.sp.gov.br

Fatec Itapetininga

RESUMO: O presente artigo tem como objetivo apresentar as rotinas geoestatísticas em R para a construção de variogramas que foram aplicadas nos dados de pluviométricos, bem como discutir o melhor ajuste para a interpolação de uma informação em um dado ponto não amostrado. Os dados utilizados correspondem à precipitação total de 28 estações meteorológicas no ano 2022 coletados pelo SIMET (Sistema de Informações Meteorológicas) e a região analisada é conhecida como UGRHI-10 ou Unidade de Gerenciamento de Recursos Hídricos de Sorocaba/Médio Tietê. Para a realização da interpolação dos dados e demais cálculos, foram utilizados o *Software R* e *RStudio*. Os resultados indicam que o variograma de modelo exponencial foi o que apresentou melhor ajuste para interpolação, mostrando-se eficiente para estimar o volume de chuva em um raio de 19 quilômetros.

Palavras-chave: Variograma. Precipitação. Bacia Hidrográfica Sorocaba e Médio Tietê.

OUTINES FOR GEOSTATISTICAL ANALYSIS USING R LANGUAGE: AN EXAMPLE WITH RAINWATER DATA FROM THE SOROCABA AND MEDIUM TIETÊ RIVER BASIN

ABSTRACT: This paper aims to present the geostatistical routines in R for the construction of variograms that were applied to rainfall data, as well as discuss the best fit for interpolation information at a given point unsampled point. The data used corresponds to the total precipitation of the year 2022 collected by SIMET (Meteorological Information System),

and the analyzed region is known as UGRHI-10 or Water Resources Management Unit of Sorocaba/Médio Tietê. The R and RStudio software were used for data interpolation and other calculations. The results indicate that the exponential model variogram provided the best fit for interpolation, proving to be efficient in estimating rainfall volume within a radius of 19 kilometers.

Keywords: Variogram. Rainfall. Sorocaba and Médio Tietê Watershed.

1 INTRODUÇÃO

O R é uma linguagem de programação voltada para a resolução de problemas estatísticos e para a visualização gráfica de dados. O código base do R foi inspirado na linguagem S, inicialmente criada no laboratório da Bell/AT&T por John Chambers e seus colegas. Esse código foi reaproveitado por dois acadêmicos, Ross Ihaka e Robert Gentleman, resultando na plataforma de programação que temos atualmente (PERLIN, 2021).

O *software R* é sinônimo de programação voltada à análise de dados com larga base de usuários e funções bem estabelecidas. Pesquisadores de áreas diversas, desde a Economia até Biologia, ou

mesmo do campo musical encontram no *software* uma quantidade significativa de códigos que facilitam suas análises. No campo empresarial, grandes empresas como *Google* e *Microsoft* já o adotaram como linguagem interna para a análise de dados.

Por ser uma ferramenta gratuita, duradoura e confiável fez com que milhares de usuários utilizassem o R como *software* estatístico para análises. A análise de dados espaciais também é muito difundida em R e tem ganhado um número significativo de recursos nos últimos tempos. Particularmente, em geoestatística, rotinas que por muito tempo estiveram disponíveis em *softwares comerciais* e de alto valor no mercado são disponibilizados em diversas bibliotecas.

A geoestatística é baseada na teoria das variáveis regionalizadas e tem a finalidade de realizar caracterização espacial de uma variável a partir de um conjunto amostral. Desse modo, a análise geoestatística visa determinar o modelo de correlação espacial dos pontos de dados, que será usado nos processos de estimativas geoestatísticas (YAMAMOTO, 2020). No *software* R encontramos o pacote *sp* (PEBESMA & BIVAND, 2005) o qual apresenta um conjunto de classes e métodos específicos para análise de dados espaciais.

Desse modo, o presente artigo tem como objetivo apresentar as rotinas geoestatísticas em R para a construção de variogramas que foram aplicadas nos dados de pluviométricos totais das estações localizadas na bacia hidrográfica Sorocaba e Médio Tietê do ano de 2022, bem como discutir o melhor ajuste para

a previsão de uma informação em um dado ponto.

2 METODOLOGIA

As operações geoestatísticas foram realizadas utilizando o programa R versão 4-3.1 (R Core Team, 2023), e as bibliotecas *sp* (PEBESMA & BIVAND, 2005) e *gstat* (GRÄLER et al., 2016).

Os dados referem-se aos índices pluviométricos totais do ano de 2022 de 28 estações localizadas na bacia Hidrográfica do Sorocaba e Médio Tietê ou estações muito próximas a ela.

Na tabela 1 são apresentadas as 28 estações, suas respectivas coordenadas geográficas e a precipitação registrada no ano de 2022.

Tabela 1 - 28 estações, suas respectivas coordenadas geográficas e a precipitação registrada no ano de 2022

Estação	Total (mm ³)	Latitude ° ' "s	Longitude ° ' "w
Tatuí	908	23 18 21	47 07 65
Cerquilha	905,8	23 10 03	47 44 51
Itu	806,6	23 15 57	47 17 57
Mairinque	814	23 32 50	47 11 03
Piedade	923,2	23 42 41	47 25 08
Itu Fatec	698,8	23 29 03	47 29 85
Mairinque ETEC	1010,2	23 55 10	47 18 69
Fatec Sorocaba	1179,09	23 30 22	47 27 21
Cabreúva	1254,6	23 18 21	47 07 65
Ibiúna	1225,49	23 39 25	47 13 33
Laranjal Paulista	855,76	23 03 01	47 50 14
Pereiras	1111,22	23 04 34	47 58 33
Porto Feliz	1124,83	23 12 55	47 31 26
Salto	1077,1	23 12 03	47 17 13

Sarapuí	1150,45	23 38 26	47 49 29
Tietê	1159,92	23 06 44	47 32 22
Vargem Gd. Paulista	1464,49	23 36 20	47 01 34
Cabreúva	1331,13	27 24 63	47 05 27
Cerquilha - ETEC	705,00	23 16 29	47 76 21
Laranjal Paulista	653,97	23 05 24	47 83 84
Piedade - ETEC	797,40	23 71 58	47 44 40
Salto	950,80	23 16 79	47 27 49
São Roque	299,77	23 52 09	47 12 68
Sarapuí	1097,32	23 61 55	47 82 39
Sorocaba	1104,54	23 47 99	47 52 67
Tatuí - ETEC	965,12	23 28 51	47 89 78
Tatuí - FATEC	784,80	23 32 89	47 89 89
Tietê Ciagro	977,21	23 06 88	47 72 25

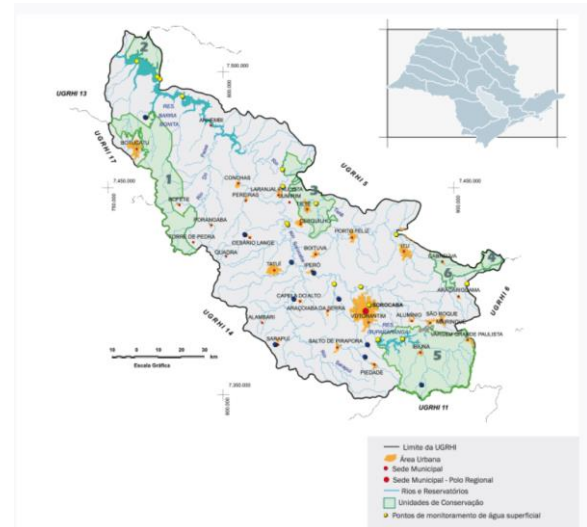
Fonte: elaboração própria (2023). Dados retirados de: <https://portal.inmet.gov.br/dadoshistoricos>

A tabela 1 pode ser editada por meio de uma planilha eletrônica, gravada, lida e importada pelo *software* R empregando-se a função *read.table*.

A região analisada, figura 1, é conhecida como UGRHI-10 ou Unidade de Gerenciamento de Recursos Hídricos de Sorocaba/Médio Tietê. Esta área abrange:

Alambari, Alumínio, Anhembi, Araçariguama, Araçoiaba da Serra, Bofete, Boituva, Botucatu, Cabreúva, Capela do Alto, Cerquilha, Cesário Lange, Conchas, Ibiúna, Iperó, Itu, Jumirim, Laranjal Paulista, Mairinque, Pereiras, Piedade, Porangaba, Porto Feliz, Quadra, Salto, Salto de Pirapora, São Roque, Sarapuí, Sorocaba, Tatuí, Tietê, Torre de Pedra, Vargem Grande Paulista, Votorantim, ou seja, são 34 municípios abrangidos por um conjunto de políticas responsáveis pela gestão dos Recursos hídricos de uma porção de 11.829 km² [...] tendo como principais rios: Alambari, Araçua, Capivara, Guarapó, Jundiuvira, Macacos, Murundu, Pirajibu, Ribeirão do Peixe, Sarapuí, Sorocaba, Sorocabuçu, Sorocamirim, Tatuí e Tietê. (FERREIRA, p.16)

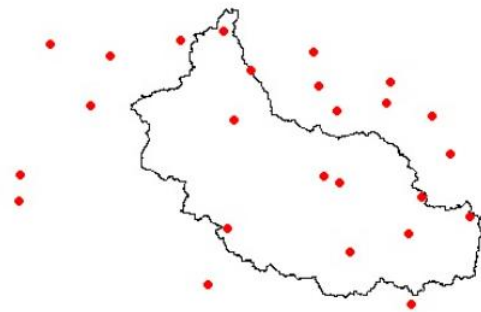
Figura 1 - Bacia Hidrográfica do Sorocaba e Médio Tietê



Fonte: <https://sigrh.sp.gov.br/cbhsmt/apresentacao>

A figura 2 mostra a Bacia Hidrográfica do Sorocaba Médio Tietê e os pontos em vermelho representam a localização das estações apresentadas na tabela 1.

Figura 2 - Bacia Hidrográfica do Sorocaba e Médio Tietê e a localização das estações meteorológicas.



Fonte: Elaboração própria (2023).

O clima da região é caracterizado por chuvas com médias anuais entre 1.250 mm e 1.400 mm e estações bem definidas com invernos secos e verões chuvosos. A vegetação é característica da Mata Atlântica com traços do Cerrado.

3 REFERENCIAL TEÓRICO

A precipitação é uma variável meteorológica de fundamental importância para estudos climáticos das diversas regiões do Brasil. Há grande preocupação de pesquisadores em analisar a dinâmica das chuvas em todo o território brasileiro a fim de, por exemplo, planejar as atividades urbanas e administrar os recursos hídricos (COSTA et al., 2021).

Os autores supracitados explicam que o estudo da distribuição das chuvas se torna fundamental para subsidiar ações no sentido de amenizar prejuízos e tentar administrar os inconvenientes causados pelas intempéries e da natureza. Destacam ainda que as técnicas da geoestatística têm-se mostrado eficientes no estudo da distribuição da precipitação e podem gerar informações essenciais por meio da Krigagem¹ e construção de mapas que descrevem o comportamento de variáveis de interesse.

O estudo geoestatístico tem como ponto de partida um conjunto de observações que constituem uma amostra. As observações, de natureza quantitativa ou qualitativa, são usadas para inferir as propriedades do fenômeno espacial em estudo. O objetivo da prática é a caracterização espacial de uma variável de interesse por meio do estudo de sua distribuição e variabilidade espaciais com determinação das incertezas associadas. O fenômeno espacial é o conjunto de todos os valores possíveis da variável de interesse, que

define a distribuição e variabilidade espaciais dessa variável dentro de um dado domínio em 2D ou 3D. Esta representa, portanto, em termos estatísticos, a população que é o conjunto de todos os valores do qual uma amostra pode ser extraída (YAMAMOTO & LANDIM, 2013).

Mello e Sampaio (2019) explicam que precipitação pluviométrica apresenta dependência espacial e, por ser conhecida a partir das informações coletadas em pontos amostrais, a estimativa de seus valores em locais não amostrados é feita por processos de interpolação. O método geoestatístico possui o diferencial de permitir que seja realizada uma análise da correlação espacial entre as amostras (autocorrelação), permitindo avaliar a pertinência de se estimar valores em locais não amostrados, bem como o erro associado aos valores estimados.

Ferramentas como o variograma e a covariância são empregadas para descrever e avaliar a dependência espacial entre pontos. Yamamoto e Landim (2013) explicam que a função variograma mede a variância entre pontos separados por uma distância h . Assim, para pontos próximos, a diferença é pequena e, portanto, a variância é pequena. Ao aumentar a distância, os valores dos pontos tornam-se mais diferentes e, conseqüentemente, a variância aumenta. Já a função covariância mede o grau de relacionamento entre pontos numa direção.

No mesmo sentido, Mello e Sampaio (2019) afirmam que é o variograma, que

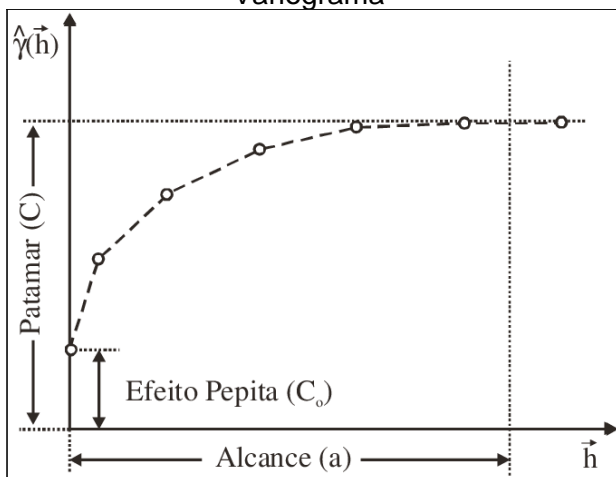
¹ É uma técnica de interpolação da geoestatística que realiza a predição de medições e observações em

localidades desconhecidas com base em dados previamente coletados.

juntamente com a covariância que fornecerão informações relevantes acerca do alcance espacial dos dados, ou seja, indicarão a distância mínima necessária entre os locais amostrais (estações de coleta), para que a estimativa em áreas não amostradas seja válida.

Portanto, pode-se dizer que o variograma representa de forma quantitativa a variação de um fenômeno regionalizado no espaço em função da distância entre as amostras. Na figura 3 ilustra-se os parâmetros de um variograma:

Figura 3 - Parâmetros de um Variograma



Fonte: Yamamoto (2020, p. 124)

Alcance (a): é a distância entre os pontos medidos em que o patamar atinge seu valor máximo. Também recebe o nome de “range”.

Patamar (C): mostra o nível de variabilidade dos valores medidos até alcançar uma medida estabilizada. Também é chamado de “Sill”.

Variância Espacial (C – C₀): representa as diferenças espaciais entre os pontos com distâncias cada vez maiores até atingir o patamar máximo.

Efeito Pepita (C₀): conforme a distância h tende a 0 o variograma deveria apresentar um valor C nulo; porém, devido a erros acumulados de medição nas amostragens ou até mesmo variações não detectadas pelo instrumento, causam esta descontinuidade da origem da curva do variograma. Este fenômeno também pode ser chamado de “Nugget Effect” e, de forma geral, podemos dizer que quanto menor for seu valor, melhor será a qualidade dos dados.

Na equação 1 representa-se a função de um variograma:

Equação 1 – Equação do Variograma

$$Y(h) = \frac{1}{2 \cdot N(h)} \sum_{i=1}^{N(h)} (Z_{xi} - Z_{xi+h})^2$$

Fonte: Yamamoto (2020, p. 125)

Em que:

$Y(h)$ = semivariância amostral,

$N(h)$ = número de pares de pontos separados pela distância h ,

Z_{xi} = valor do atributo no ponto x_i ,

Z_{xi+h} = valor do atributo no ponto distante h do atributo x_i ,

X_i = local do plano onde foi medido o atributo.

Sobre o emprego de tais métodos, Mello e Sampaio (2019) afirmam que estes vêm sendo largamente utilizados para caracterizar a precipitação pluviométrica em diferentes contextos climáticos no globo. Contudo, a acurácia e conveniência dos mapas e modelos produzidos nem sempre têm sido foco de análise. Este cuidado se faz necessário, uma vez que a interpolação em áreas de baixa densidade amostral, ainda que por processos geoestatísticos, pode resultar em mapas e

modelos com grandes resíduos para os valores estimados ou até mesmo em representações inadequadas do fenômeno cartografado.

4 RESULTADOS E DISCUSSÕES

Inicialmente, utiliza-se o comando `setwd("C:/GEOESTAT")` o qual dá o caminho para a busca dos arquivos que serão utilizados para análise. Em seguida, com o comando `dados <- read.csv2("dados.csv")`, criamos a variável "dados" e solicitamos a leitura do arquivo dados.csv. Este arquivo corresponde à tabela 1 exibida no presente arquivo.

A fim de obter as estatísticas descritivas utilizamos a rotina `summary(dados$Prec_2022)`, a qual nos retornou as informações expostas na tabela 2.

Tabela 2 – Estatísticas descritivas da variável precipitação (mm)

Mínimo	299,8
1º Quartil	812,1
Mediana	971,2
Média	976,3
3º Quartil	1131,2
Máximo	1464,5

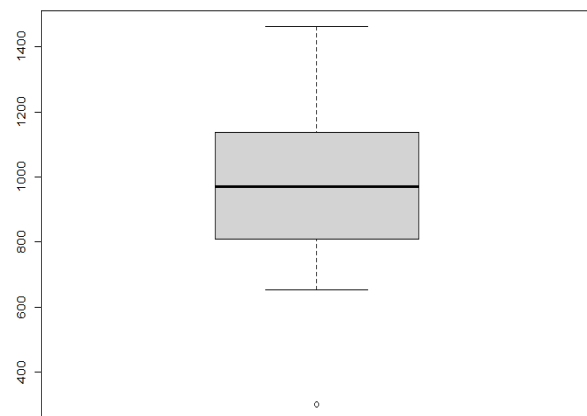
Fonte: elaboração própria (2023)

Por meio dos dados exposto na tabela 2, vê-se que o menor volume de chuva registrado no ano de 2002 foi de 299,8 mm, o qual foi registrado no município de São Roque. Já o município com maior volume é Vargem Grande Paulista com 1464,5 mm nesse mesmo período. Vê-se também que a média (971,2

mm) e a mediana (976,3 mm) são valores próximos.

Afim de visualizar a distribuição dos dados e a presença ou não de valores discrepantes (*outliers*) construímos o *boxplot*, (figura 4) fazendo uso do comando `boxplot(dados$Prec_2022)`.

Figura 4 – Box-plot da variável precipitação (mm)



Fonte: elaboração própria (2023)

Analisando-se as informações contidas nesse gráfico pode-se verificar que os dados são praticamente simétricos, têm uma pequena assimetria positiva. O município de São Roque com um volume de chuva equivalente a 299,77 mm revelou-se com um *outlier*, pois foi o que apresentou um índice pluviométrico muito abaixo das demais.

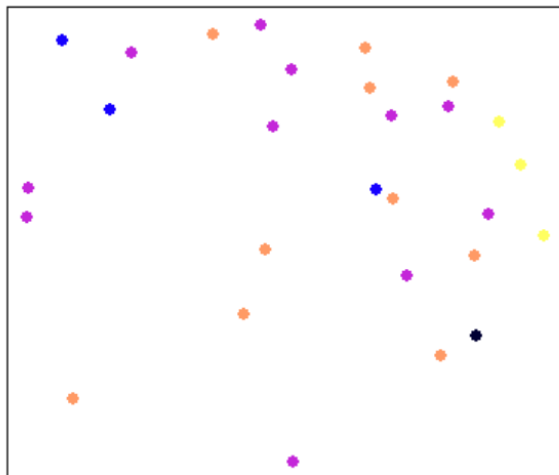
Os trechos a seguir referem-se a variografia executando, respectivamente, o cálculo do variograma a partir de uma amostra de dados, ajusta um modelo teórico ao variograma amostrado a partir de valores informados e constrói o gráfico juntamente com a linha de ajuste.

No entanto, primeiramente, há a necessidade de usar os dois comandos que

seguem de forma a tornar os dados derivados da classe *Spacial*.

```
> coordinates(dados) =~ Long+Lat
> spplot(dados['Prec_2022'])
```

Figura 5 – Mapa de pontos da variável precipitação (mm)



- [299.8,532.7]
- [532.7,765.7]
- [765.7,998.6]
- [998.6,1232]
- [1232,1464]

Fonte: elaboração própria (2023)

Na figura 5 vê-se que foi elaborado um mapa de pontos (cada ponto corresponde a um município localizado de acordo com suas respectivas coordenadas geográficas). Além disso, cada cor corresponde à uma classe de precipitação em mm³.

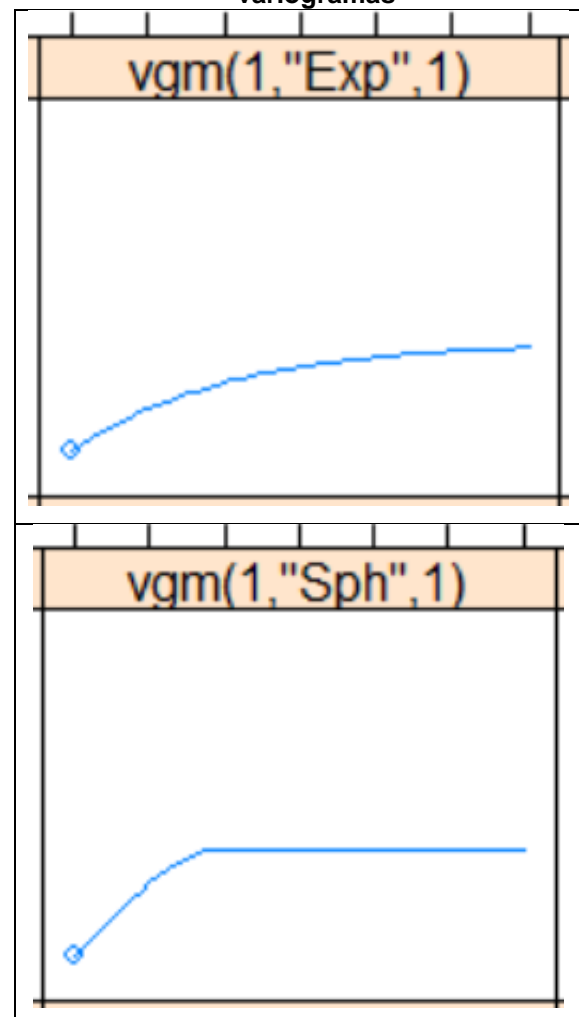
A fim de construir o variograma há a necessidade de transformar as coordenadas geográficas em utm (KM), ou seja, com suas respectivas distâncias dadas em quilômetros. Essa transformação foi realizada utilizando os seguintes comandos:

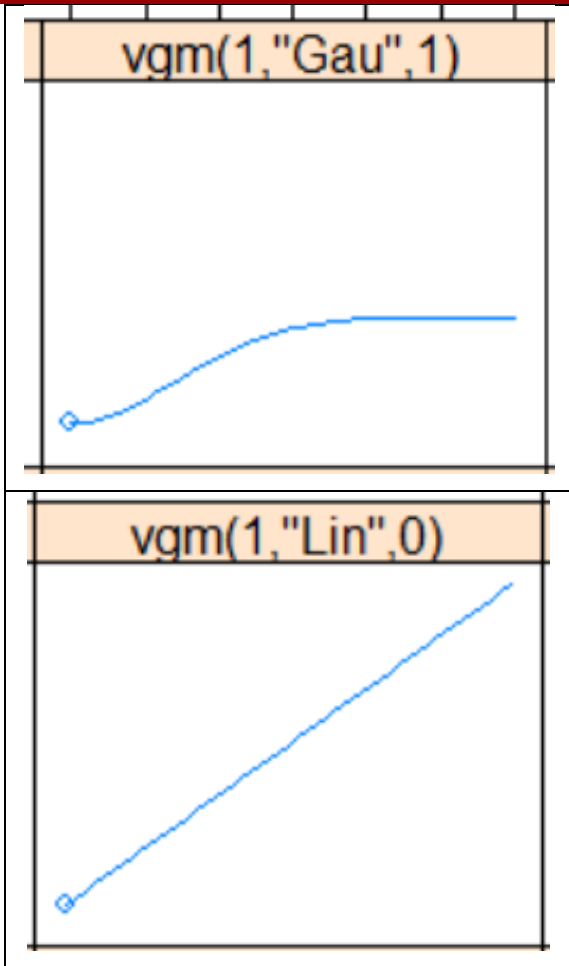
```
> proj4string(dados) <- CRS("+proj=longlat
+ellps=WGS84+towgs84=-205.57,168.77,-
4.12,0,0,0,0")
```

```
> ozone.UTM <- spTransform(dados,
CRS("+proj=utm +zone=22 +south
+ellps=intl +towgs84=-205.57,168.77,-
4.12,0,0,0,0 +units=km +no_defs"))
> ozone.UTM
```

Feita esta transformação, os comandos exibidos na sequência foram empregados para a construção do variograma. Os modelos mais comuns encontrados na literatura são o exponencial, e o esférico, o gaussiano e o linear. Na figura 6 são exibidos tais modelos teóricos.

Figura 6 – Modelos mais comuns de variogramas





Fonte: elaboração própria (2023)

Para a construção do variograma de modelo exponencial, foram empregados os seguintes comandos.

```
vv <- variogram(Prec_2022 ~ 1, ozone.UTM[
c(4,8,23),], cutoff = 70, width = 70/6)
> vv
> plot(vv)
> v.fit = fit.variogram(vv, vgm("Exp"))
> v.fit
```

Tais comandos retornaram as estimativas dos parâmetros (tabela 3).

Tabela 3 – Elementos do Modelo exponencial do variograma

model	psill	range
1 Nug	0.00	0.0000

2 Exp	44164.03	19.4821
-------	----------	---------

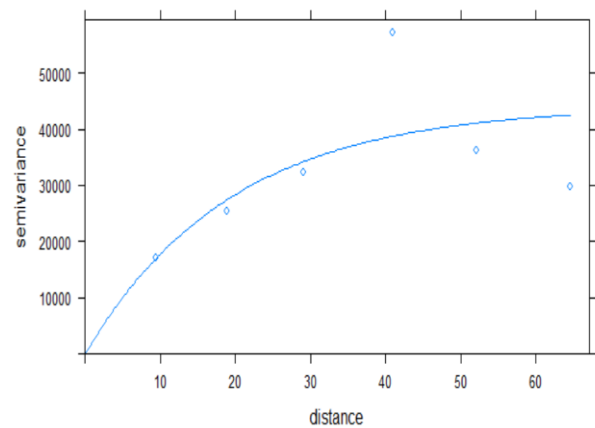
Fonte: Elaboração Própria

Os dados mostram que o alcance espacial no Modelo exponencial foi de aproximadamente 19,5 km, sugerindo que locais separados maiores que este limiar apresenta baixa dependência espacial entre si. O patamar, também chamado de sill, que mede o nível de variabilidade dos valores medidos até alcançar uma medida estabilizada, foi de aproximadamente 44164.

Na sequência, mostra-se o comando utilizado para a construção do gráfico do ajuste e, na figura 7, o gráfico de ajuste do modelo exponencial

```
> plot(vv,v.fit)
```

Figura 7 – Gráfico do modelo de ajuste do variograma de modelo exponencial



Fonte: elaboração própria (2023)

Abaixo são representados os comandos para determinar a eficiência do variograma de modelo exponencial.

```
> require(hydroGOF)
```



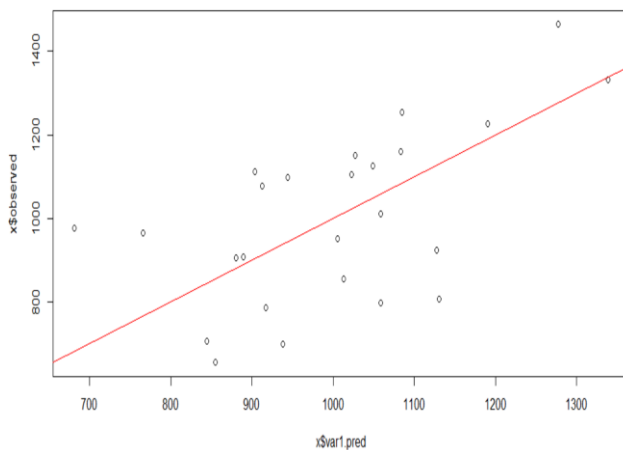
```
> x <- krige.cv(Prec_2022 ~ 1, ozone.UTM[-
c(4,8,23),], v.fit, nmax=2)
```

```
> gof(x$var1.pred, x$observed, digits=4)
> plot(x$var1.pred, x$observed)
> abline(0,1, col='red')
```

Como resultado, foi obtido que o valor do RMSE (*root mean-square error* em inglês ou raiz do erro quadrático) é de 167, isso significa que em média o modelo está cometendo um erro de 167 mm em relação aos valores observados.

Na figura 8 exibimos a reta de regressão ($y = x$), a qual relaciona valores observados e valores estimados.

Figura 8– Gráfico de pontos e reta de regressão ($y = x$) entre valores observados e estimados



Fonte: elaboração própria (2023)

O mesmo procedimento foi empregado para construir o variograma de modelo esférico

```
> v.fit1 = fit.variogram(vv, vgm("Sph"))
> v.fit1
> plot(vv,v.fit1)
```

Tais comandos retornaram os dados exibidos na tabela 4.

Tabela 4 – Elementos do Modelo esférico do variograma

model	psill	range
1 Nug	1023.712	0.0000
2 Eph	31767.957	3.724573

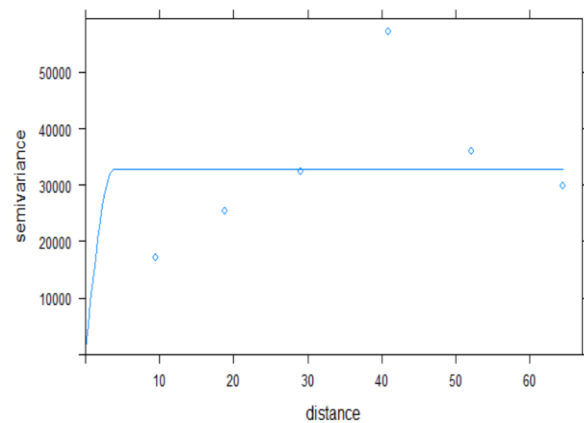
Fonte: elaboração própria (2023)

Para este modelo verifica-se que o alcance foi de aproximadamente 3,7 km, o que é significativamente inferior ao alcance do modelo exponencial que foi de 19,5 km.

O comando a seguir foi empregado para a construção do gráfico do modelo esférico, figura 9

```
> plot(vv,v.fit1)
```

Figura 9 – Gráfico do modelo de ajuste do variograma de modelo esférico



Fonte: elaboração própria (2023)

Por motivo deste modelo apresentar um alcance pouco significativo considerando-se o objetivo do presente estudo, não foi dado seguimento das análises no que tange a verificação da eficiência do modelo e construção da reta de regressão.

Os modelos gaussiano e linear, que são bastante comuns na literatura, também foram construídos, mas por apresentarem resultados estatísticos insignificantes para este estudo não são exibidos neste artigo.

5 CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo testar as rotinas do R para aplicações tecnológicas de apoio a Sistemas de Informações Meteorológicas com a intenção de efetivar preenchimentos de falhas em casos de danos em sensores. O foco principal de interesse será futuramente comparar a eficiência e precisão da interpolação de dados por krigagem comparadas a outras técnicas de interpolação, bem como discutir qual o melhor ajuste para a realização de informações em pontos que não há coleta de dados.

O *software* R mostrou-se como uma ferramenta prática e eficaz para análises geoestatísticas, uma vez que importa, armazena analisa e permite a visualização dos resultados. Revelou também que para o emprego e execução das rotinas é necessário um conhecimento intermediário, dado que não se trata de um *software* intuitivo, o qual o usuário é encaminhado através de janelas a cada tarefa executada.

No que tange as análises estatísticas preliminares, verificou-se que a região da Bacia Hidrográfica do Sorocaba e Médio Tietê apresentou uma média de 976,3 mm de chuvas anuais no ano de 2022, um valor um pouco abaixo do habitual da região que está entre 1.250mm e 1.400mm anuais.

Quanto ao variograma, percebeu-se que o melhor ajuste foi dado pelo modelo exponencial, o qual apresentou um alcance de aproximadamente 19,5 km, sugerindo que locais separados maiores que este limiar apresenta baixa dependência espacial entre si. O patamar, também chamado de *sill* foi de aproximadamente 44164. Além disso, RMSE para este modelo foi de 167 mm/ano, indicando que em média o modelo está cometendo um erro de 167 mm/ano em relação aos valores observados

REFERÊNCIAS BIBLIOGRÁFICAS

COSTA, I. S.; MEDEIROS, E. S.; SILVA, A. Q.; MEDEIROS, C. C. B.; OLIVEIRA, L. A.. Mapeamento da precipitação em uma Região do Nordeste Brasileiro por meio de técnicas geoestatística. Revista Ibero Americana de Ciências Ambientais, v.12, n.5, p.761-769, 2021. DOI: <http://doi.org/10.6008/CBPC2179-6858.2021.005.0059>

FERREIRA, Rhuan Henrique. Estudo das variações de precipitação na bacia do Sorocaba e Médio Tietê (UGRHI-10) nos períodos de verão entre os anos de 2000 e 2015. 2016. 78 f. Trabalho de conclusão de curso (bacharelado - Geografia) - Universidade Estadual Paulista Julio de Mesquita Filho, Câmpus Experimental de Ourinhos, 2016. Disponível em: <<http://hdl.handle.net/11449/155788>>. Acesso em: 22 mar. 2023.

GRÄLER, B.; PEBESMA, E.; HEUVELINK, G. Spatio-Temporal Interpolation using gstat, Journal R., 8, 204–218. 2016.

MELLO, Y. R.; SAMPAIO, T. V. M. Análise Geoestatística Da Precipitação Média Para O Estado Do Paraná. Revista Brasileira de Climatologia. V.25, p. 643 – 660, 2019

PEBESMA, Edzer; BIVAND, Roger S. S classes and methods for spatial data: the sp package. R news, v. 5, n. 2, p. 9-13, 2005.

PERLIN, M.S. Análise de dados financeiros e econômicos com o R. 3ª. Ed. Porto Alegre: Marcelo S. Perlin (publicações independentes), 2021.

R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

YAMAMOTO, J.K. **Estatística, análise e interpolação de dados geoespaciais.** São Paulo: Paulo's Editora, 2020.

YAMAMOTO, J.K.; LANDIM, P. M. B. **Geoestatística: conceitos e aplicações.** São Paulo: Oficina de Textos, 2013.