



**UNIVERSIDADE ESTADUAL PAULISTA**  
*FACULDADE DE CIÊNCIAS E TECNOLOGIA – PRESIDENTE  
PRUDENTE*

Programa de Pós-Graduação em Ciências Cartográficas

**CARLOS ALBERTO JÁCOMO**

**MODELOS GEOESTATÍSTICOS ESPAÇO-TEMPO  
APLICADOS A DADOS PLUVIOMÉTRICOS NO  
OESTE DO ESTADO DE SÃO PAULO**



**Presidente Prudente  
2011**

**CARLOS ALBERTO JÁCOMO**

**MODELOS GEOESTATÍSTICOS ESPAÇO-TEMPO  
APLICADOS A DADOS PLUVIOMÉTRICOS NO OESTE  
DO ESTADO DE SÃO PAULO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Cartográfica da Universidade Estadual Paulista – Faculdade de Ciências e Tecnologia – Campus de Presidente Prudente, como parte de requisitos para a obtenção do título de Mestre em Ciências Cartográficas.

Orientadora: Profa. Dra. Vilma Mayumi Tachibana.

Coorientador: Prof. Dr. Edilson Ferreira Flores.

**PRESIDENTE PRUDENTE  
2011**

J18m Jácomo, Carlos Alberto.  
Modelos geoestatístico espaço-tempo aplicados a dados pluviométricos  
no oeste do Estado de São Paulo / Carlos Alberto Jácomo. - Presidente  
Prudente : [s.n.], 2011  
109 f. : il.

Orientador: Vilma Mayumi Tachibana  
Coorientador: Edilson Ferreira Flores  
Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de  
Ciências e Tecnologia  
Banca: Paulo Milton Barbosa Landim, Nilton Nobuhiro Imai  
Inclui bibliografia

1. Geoestatística. 2. Modelos espaço-tempo. 3. Análise de agrupamento  
I. Tachibana, Vilma Mayumi. II. Flores, Edilson Ferreira. III. Universidade  
Estadual Paulista. Faculdade de Ciências e Tecnologia. III. Título.

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento de Informação – Serviço  
Técnico de Biblioteca e Documentação – UNESP – Presidente Prudente.

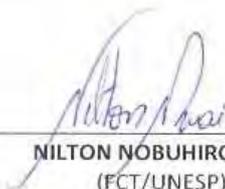
BANCA EXAMINADORA



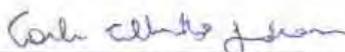
PROFA. DRA. **VILMA MAYUMI TACHIBANA**  
(ORIENTADORA)



PROF. DR. **PAULO MILTON BARBOSA LANDIM**  
(UNESP/BOTUCATU)



**MILTON NOBUHIRO IMAI**  
(FCT/UNESP)



**CARLOS ALBERTO JÁCOMO**

PRESIDENTE PRUDENTE (SP), 17 DE NOVEMBRO DE 2011

RESULTADO:

Aprovado

## **Agradecimentos**

Em primeiro lugar a Deus, pela saúde, sabedoria para encerrar essa trajetória seguida pela minha família, especialmente meu Pai, por estar sempre incentivando a estudar e correr atrás de meus objetivos.

À professora Dra. Vilma Mayumi Tachibana e o professor Dr. Edilson Ferreira Flores pela orientação, confiança no meu trabalho, sempre irei ter como exemplos de vida como pessoas e profissionais que são.

Ao Programa de Pós-Graduação em Ciências Cartográficas (PPGCC) pela oportunidade que foi dada e pela total infra-instrutora.

À Coordenação de Aperfeiçoamento de Pessoal do Ensino Superior (CAPES) pela bolsa concedida.

Aos amigos do PPGCC, os que já defenderam e os que entraram esse ano, pelos grandes momentos que tive especialmente na sala de permanência. Desculpe por não citar todos, pois tenho medo de esquecer algum.

Ao Adilson pelas caronas até minha casa, pela hospedagem no momento que voltei para encerrar minha dissertação, pelos conselhos, você sempre será lembrado não somente por mim, mas por todos os amigos do PPGCC.

## Resumo

O regime pluviométrico pode ser compreendido pelos modelos espaço-temporal. A busca, cada vez mais crescente, por modelos espaço-temporal se deve a abundância de dados armazenados pelo tempo e espaço disponibilizados por muitos institutos e órgãos governamentais e o conhecimento desse fenômeno pode ser importante para os planejamentos agrícolas, urbanos e ambientais. Este trabalho teve por objetivo estudar o comportamento espaço-temporal do regime pluviométrico do oeste do Estado de São Paulo por meios de modelos que descrevem o fenômeno no espaço, tempo e espaço-tempo. Inicialmente foi criado um banco de dados referente à precipitação pluviométrica no período de janeiro de 1970 a dezembro de 2000. Com este conjunto, fez-se uma análise de agrupamento, utilizando-se a técnica proposta por Ward para formação de grupos de estações similares e formação de anos similares. Com as estações formou-se 7 grupos homogêneos e para os anos, 3 grupos homogêneos. Nessa análise foi possível localizar grupos com o mesmo comportamento, como a região mais seca e região mais chuvosa, e os anos mais secos e chuvosos. Os resultados apresentados através do método de Ward foram semelhantes aos resultados publicados por outros autores, mesmo considerando somente a variável precipitação mensal. Dentre os modelos espaço-tempo estudados, o modelo geoestatístico espaço-tempo apresentou as melhores estimativas. As superfícies espaço-tempo geradas a partir das estimativas desse modelo descreveram o comportamento do regime pluviométrico ao longo do espaço e do tempo. Essas superfícies mostram as menores e maiores concentrações de chuva e tanto no eixo temporal quanto no eixo espacial, o regime pluviométrico descreveu um comportamento periódico. A utilização de modelos espaço-tempo pode ser útil para instituições que trabalham com banco de dados temporal e espacial que contem dados ausentes, pois esses modelos auxiliam no preenchimento desses dados com boa precisão.

**Palavras-Chave:** Geoestatística, Análise de Agrupamento, Modelos Espaço-Tempo.

## **Abstract**

Rainfall can be understood by using space-time models. The increasing search for space-time models is due to the large amount of data, stored by time and space, offered by many institutes and government agencies. Knowledge of this phenomenon may be important for agricultural, urban, and environmental planning. This work aimed to study the behavior of the space-time rainfall in western São Paulo state by means of models that describe the phenomenon in space, time, and space-time. Initially, a database was created of rainfall from January, 1970 to December, 2000. With this set, we carried out a cluster analysis using the technique proposed by Ward for the formation of similar groups of stations and the formation of similar years. It was possible to form seven homogeneous groups for the stations, and three homogeneous groups for the years. In this analysis, it was possible to find groups with the same behavior as the driest and wettest regions, and the dry and wet years. The results presented by Ward's method were similar to those published by other authors, even considering only variable monthly precipitation. Among the space-time models studied, the Geostatistical Space-Time model provided the best estimates. The space-time surface estimates generated from this model described the behavior of rainfall over space and time. These surfaces showed the lowest and highest concentrations of rain, and along the space and time axes, rainfall showed a periodic behavior. The use of space-time models can be useful for institutions that work with databases containing missing spatial and temporal data, because these models help in filling this data with high accuracy.

**Keywords:** Geostatistic, Cluster's Analysis, Space-Time Models.

## Lista de Figuras

1. Oeste Paulista com suas áreas administrativas.....	15
2. Unidades Pluviais do Estado de São Paulo.....	17
3. Exemplo de semivariograma experimental.....	21
4. Elementos amostrais regularmente distribuídos (a) e as direções convencionais (b)	21
5. Parâmetros para o cálculo do semivariograma a partir dos elementos amostrais.....	22
6. Comportamento característico dos modelos teóricos.....	23
7. Representação do modelo Potência segundo os coeficientes de declividade.....	25
8. Tolerância angular de 90° com direção de 90° (a) e tolerância angular de 90° com direção de 45° (b).....	26
9. Trecho do banco de dados.....	45
10. Banco de dados referente à região oeste do Estado de São Paulo.....	46
11. Distribuição espacial das estações de coleta do oeste paulista.....	47
12. Fluxograma da metodologia proposta.....	48
13: Gráfico dos coeficientes de fusão para as localidades.....	50
14: Dendrograma para as localidades.....	51
15. Distribuição espacial das estações pluviométricas separadas por grupos homogêneos.....	51
16. Dendrogramas individuais dos 7 grupos homogêneos para as localidades.....	52
17. Comparação dos mapas de chuva gerados por diferentes autores: Jacomo (2011) – acima, Sant’Anna Neto (1995) – esquerda e Boin (2000) – direita.....	53
18. Gráfico de linha para as médias totais anuais.....	54
19. Gráfico dos coeficientes de fusão.....	56
20. Dendrograma para os anos.....	56
21. Dendrogramas individuais dos 3 grupos homogêneos.....	57
22. Histograma dos 3 grupos homogêneos.....	57
23. Box-plot dos 3 grupos homogêneos.....	58
24. Gráfico de pontos para os 3 grupos.....	59
25. Superfície variográfica das precipitações máxima e média mensal respectivamente.....	59
26. Modelo esférico ajustado para a variável precipitação máxima mensal.....	60

27. Modelo efeito pepita puro ajustado para a variável precipitação máxima mensal na direção de 120°.....	61
28. Modelo efeito pepita puro ajustado para a variável precipitação máxima mensal na direção de 30°.....	62
29. Modelo esférico ajustado para a variável precipitação média mensal.....	62
30. Modelo potência ajustado para a variável precipitação média mensal na direção de 130°.....	63
31. Modelo esférico ajustado para a variável precipitação média mensal na direção de 40°.....	64
32. Semivariograma temporal para a variável precipitação máxima.....	65
33. Função ajustada para a variável precipitação máxima.....	65
34. Semivariograma temporal para a variável precipitação média.....	66
35. Função ajustada para a variável precipitação média.....	66
36. Superfície gerada através do modelo produto-soma generalizado da precipitação máxima mensal.....	68
37. Superfície gerada através do modelo produto-soma generalizado da precipitação média mensal.....	69
38. Semivariograma experimental e modelo esférico da componente $M_I(u)$ para o item a).....	71
39. Gráfico de dispersão dos valores observados e estimados da precipitação mensal – item a).....	72
40. Histograma do Erro da precipitação mensal – item a).....	73
41. Semivariograma experimental e modelo esférico da componente $M_I(u)$ para o item b).....	74
42. Gráfico de dispersão dos valores observados e estimados da precipitação anual – item b). .....	75
43. Histograma do Erro da precipitação anual – item b).....	75
44. Séries temporais referentes às precipitações totais.....	76
45. Mapa de superfície da variável precipitação total anual mínima e estações retiradas .....	78

46. Mapa de superfície da variável precipitação total anual máxima e estações retiradas .....	78
47: Gráficos de dispersão dos valores observados e estimados da precipitação total anual máxima – item d).....	79
48. Histogramas do Erro da precipitação total anual máxima – item d).....	79
49. Gráficos de dispersão dos valores observados e estimados da precipitação total anual média – item d).....	79
50. Histogramas do Erro da precipitação total anual média – item d).....	80
51. Gráficos de dispersão dos valores observados e estimados da precipitação total anual mínima – tem d).....	81
52. Histogramas do Erro da precipitação total anual mínima – item d).....	81
53. Distribuição espacial das estações de coleta do oeste paulista e estação retirada....	82
54. Semivariograma experimental e modelo esférico da componente $M_I(u)$ para o item e).....	83
55. Gráfico de dispersão da precipitação mensal retirando a estação C8-055 – item e). .....	84
56. Histograma do erro da estação C8-055 – item e).....	84
57. Estações de coleta próximas as estimativas.....	85
58. Semivariograma experimental e modelo esférico da componente $M_I(u)$ para o item g).....	87
59. Mapa de superfície das médias temporais do oeste paulista.....	88
60. Trecho e eixo escolhido para calcular as estimativas espaço-temporal.....	88
61. Mapa de superfície espaço-tempo do regime pluviométrico do trecho escolhido – item g).....	90
62. Região selecionada para o experimento – item i).....	91
63. Distribuição espacial da grade 9x5 para a região selecionada.....	91
64. Faixas selecionadas para as superfícies espaço-tempo – item i).....	92
65: Mapa de superfície espaço-tempo da coordenada X = 500000 (Faixa 1).....	93
66. Mapa de superfície espaço-tempo da coordenada X = 540000 (Faixa 2).....	93
67. Mapa de superfície espaço-tempo da coordenada X = 580000 (Faixa 3).....	94
68. Mapa de superfície espaço-tempo da coordenada Y = 7505000 (Faixa 4).....	95

69. Mapa de superfície espaço-tempo da coordenada Y=7545000 (Faixa 5).....	95
70: Mapa de superfície espaço-tempo da coordenada Y=7585000 (Faixa 6).....	96
71. Mapa de superfície espaço-tempo da coordenada Y=7625000 (Faixa 7).....	96
72. Mapa de superfície espaço-tempo da coordenada Y=7665000 (Faixa 8).....	97
73. Mapas de superfície do período de 1970 a 1977 .....	98
74. Mapas de superfície do período de 1978 a 1985.....	99
75. Mapas de superfície do período de 1986 a 1993 .....	100
76: Mapas de superfície do período de 1994 a 2000.....	101
77. Retorno dos maiores índices pluviométricos na mesma localidade.....	102
78. Box-plot dos anos de 1970 a 1984.....	103
79. Box-plot dos anos de 1985 a 1984.....	103

## **Lista de Tabelas**

1. Estatísticas descritivas referentes aos 7 grupos homogêneos.....	55
2. Estatísticas descritivas referentes aos 3 grupos homogêneos.....	58
3: Correlações e EAM's entre valores estimados, observados e médios.....	86
4. Correlações entre pontos observados e estimados.....	89

## Sumário

<b>1. INTRODUÇÃO.....</b>	<b>15</b>
1.1. Contextualização e Justificativa.....	15
1.2. Objetivo.....	18
1.3. Estrutura do Trabalho.....	18
<b>2. FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>19</b>
2.1. Geoestatística Clássica .....	19
2.1.1. Semivariograma.....	19
2.1.2 Modelos Teóricos.....	22
2.1.3. Anisotropia.....	25
2.1.4. Modelagem Anisotrópica.....	26
2.1.5. Hipóteses Geoestatística.....	27
2.1.6. Krigagem.....	28
2.1.6.1. Krigagem Simples.....	29
2.1.6.2. Krigagem Ordinária.....	32
2.1.6.3. Krigagem Universal .....	34
2.1.7. Cokrigagem.....	34
2.1.7.1. Semivariograma Cruzado.....	36
2.2. Modelagem Espaço-temporal .....	36
- modelo métrico .....	38
- modelo produto .....	39
- modelo linear .....	39
- modelo não-separável .....	40
- modelo de covariância produto-soma .....	40
- modelo produto-soma generalizado .....	41
- modelo geoestatístico espaço-temporal .....	42
- Máxima Entropia Bayesiana (MEB) .....	44
<b>3. EXPERIMENTO: DESCRIÇÃO, MÉTODOS E RESULTADOS .....</b>	<b>45</b>
3.1. Banco de dados.....	45
3.2. Método proposto.....	47
3.3. Programas computacionais utilizados.....	48
3.4. Resultados .....	49
3.4.1. Análise exploratória .....	49
3.4.2. Modelo produto-soma generalizado .....	59
a) geração do modelo espacial .....	59
b) geração do modelo temporal .....	64
c) geração do modelo espaço-tempo .....	66
3.4.3. Modelo geoestatístico espaço-tempo .....	70
a) período de 31 anos mensalmente (janeiro de 1970 a dezembro de 2000) .....	70
b) período de 31 anos anualmente (1970 a 2000) .....	73

c) comparação entre precipitação mensal e precipitação anual .....	76
d) precipitações totais anuais: máxima, média e mínima .....	77
e) estação de coleta distante das demais estações .....	81
f) comparação entre as estimativas geradas através do modelo e através das médias aritméticas.....	84
g) recorte no espaço-tempo.....	86
h) região no espaço-tempo .....	90
i) geoestatística no espaço .....	97
<b>4. CONCLUSÃO: Considerações finais, recomendações, trabalhos futuros .....</b>	<b>104</b>
<b>REFERÊNCIAS .....</b>	<b>107</b>

## 1. INTRODUÇÃO

### 1.1. Contextualização e Justificativa

O sucesso ou fracasso de um empreendimento pode estar relacionado às características do regime pluviométrico de uma determinada região, fazendo com que o conhecimento do comportamento desse atributo seja de grande importância nos planejamentos urbanos, agrícolas e ambientais. (SIQUEIRA et al., 2007)

O regime pluviométrico no oeste do Estado de São Paulo pode ser compreendido através de modelos espaço-temporal que tem se tornando uma importante ferramenta para o reconhecimento de vários fenômenos. A busca, cada vez mais crescente, de modelos espaço-temporal se deve a abundância de dados disponibilizados por muitos institutos. O fenômeno a ser estudado nessa dissertação é a precipitação pluviométrica no oeste do Estado de São Paulo (oeste paulista).

A região do oeste paulista ocupa 44,3% do Estado de São Paulo e está localizada entre os paralelos 21° e 23° S e meridianos 50° e 53° W, abrange cinco regiões administrativas e 107 municípios assim distribuídos: 53 municípios na região administrativa de Presidente Prudente, 25 municípios na região de Marília, 26 municípios na região de Araçatuba, 2 municípios na região de Bauru e apenas 1 município na região de São José do Rio Preto. (Figura 1)

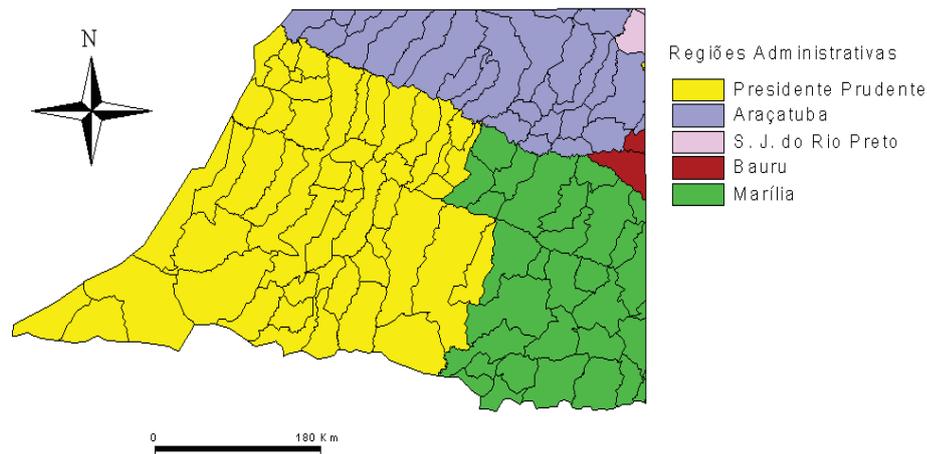


Figura 1: Oeste Paulista com suas áreas administrativas.

Monteiro (1977) descreve a respeito das características climáticas do oeste paulista como a presença do período seco (inverno) e período chuvoso, sendo que o primeiro sob influência dos sistemas polares e o segundo influenciado pelos sistemas tropicais. Já a precipitação anual média varia entre 1200 e 1500 mm, enquanto a temperatura média anual permanece acima de 22°C. (BOIN, 2000).

A respeito do comportamento do regime pluviométrico no Estado de São Paulo, Sant`Anna Neto (1995) fez uma classificação tipológica das chuvas no Estado de São Paulo conforme a Figura 2. O autor dividiu em oito unidades regionais, sendo que a região deste estudo (oeste paulista) encontra-se na unidade 7-Oeste que está localizada no Planalto Ocidental. Essa região apresenta altitudes que variam de 300 a 600 m e a precipitação anual fica entre 1100 e 1500 mm. As diferenças entre as sub-unidades ocorrem segundo fatores de ordem local, pela longa estiagem e diversidade sazonal. (FLORES, 2000) A unidade 7-Oeste, subdivide-se em seis sub-unidades, a seguir:

- Rio Grande/São José dos Dourados (7.a) - A precipitação anual totaliza, em média de 1300 a 1500 mm sofrendo a ação dos sistemas tropicais e equatoriais. A sazonalidade é a mesma da maior parte do território paulista.

- Noroeste (7.b) - A região de Araçatuba apresenta altitudes menos elevada e recebe menores quantidades de chuvas que a região anterior, com totais anuais entre 1100 a 1500 mm. O trimestre mais seco ocorre nos meses de junho a agosto e o mais chuvoso ocorre nos meses de dezembro a fevereiro.

- Alta Sorocabana (7.c) – Pertencente a região de Presidente Prudente e do Pontal do Paranapanema, onde existem as mesmas condições descritas na sub-unidade Noroeste, porém, devido a sua localização na faixa de transição dos climas controlados pelas massas tropicais e extra tropicais apresenta uma maior variabilidade interanual.

- Vale Médio do Rio Tietê (7.d) – A precipitação média anual é de 1300 a 1500 mm com altitudes entre 300 a 500 m, situado na região central do Estado. Apresenta grande variabilidade interanual e forte estiagem nos meses de inverno.

- Serra dos Agudos (7.e) – A altitude se eleva em torno de 800 m, devido à região de Marília apresentar uma pequena cuesta, provocando um efeito orográfico da chuva que aumenta a pluviosidade na área, onde os valores ultrapassam 1500 mm atingindo 1800 mm.

- Médio Vale do Rio Paranapanema (7.f) – A precipitação média anual é de 1300 a 1600 mm, apresenta baixas altitudes entre 300 a 500 m, encontra-se na faixa de transição dos climas zonais e está localizado na região de Assis/Ourinhos.

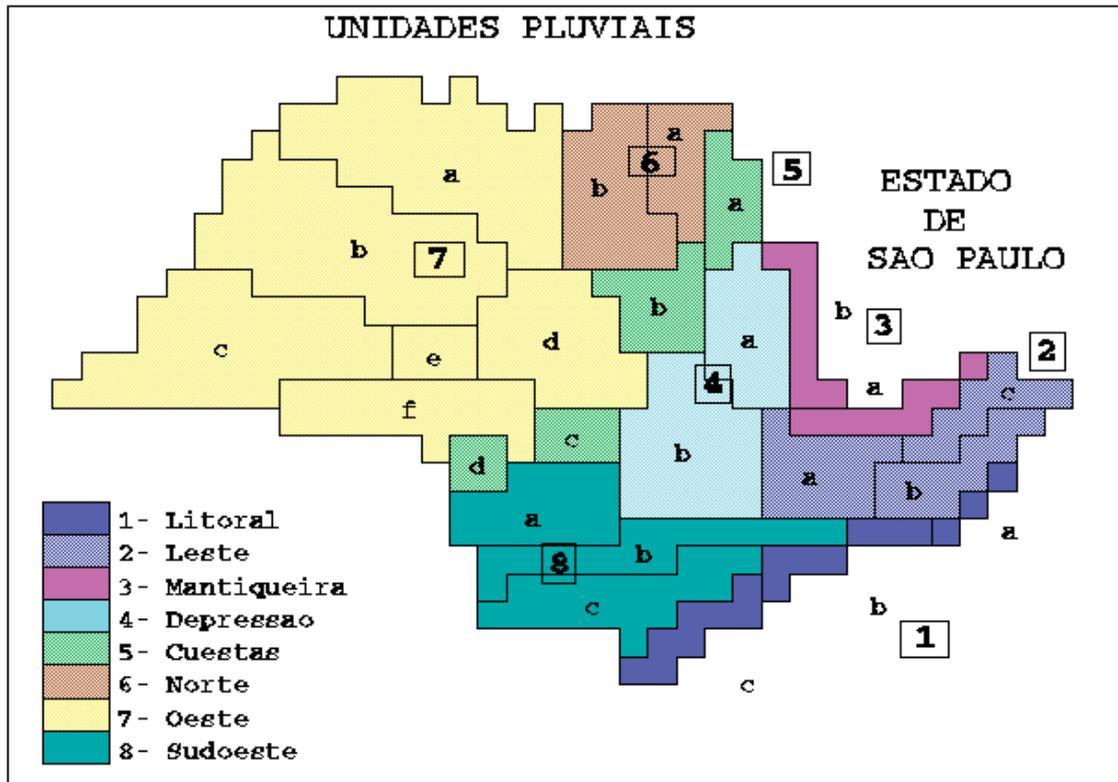


Figura 2: Unidades Pluviais do Estado de São Paulo.  
Fonte: Sant'Anna Neto (1995).

Por muitos anos, durante o século passado, a economia da maioria dos municípios do Estado de São Paulo era movida pela agricultura. Grandes fazendas ou milhares de pequenas propriedades cultivaram café, algodão, amendoim, milho, hortifruticultura, etc. que dependiam dos fatores climáticos, principalmente das chuvas. Assim, desde há muito tempo, os índices pluviométricos, temperatura, umidade, entre outras medidas vem sendo registradas por vários departamentos estaduais. E essa grande massa de dados tem incentivado os pesquisadores a buscarem modelos matemáticos que descrevam o comportamento da chuva no Estado.

Existem resultados de estudo da precipitação pluviométrica usando a geoestatística para modelagem do fenômeno em uma determinada região, como os trabalhos de Mota (2008) e Siqueira et al. (2007). Alguns trabalhos já utilizam SIG e outras ferramentas mais atuais, como no

caso o trabalho de Douaik (2009). Mas, em relação à modelagem da distribuição espaço-temporal de fenômenos dessa natureza os trabalhos se tornam raros.

## **1.2. Objetivo**

O objetivo principal deste trabalho foi estudar o regime pluviométrico no oeste do Estado de São Paulo por meio de modelos que descrevam o fenômeno no espaço, no tempo e no espaço-tempo.

Para alcançar esse objetivo, também se teve como metas:

- Aplicar as técnicas de Geoestatística clássica para analisar o comportamento da precipitação pluviométrica ao longo do tempo e no espaço, considerando essas duas informações separadamente.
- Aplicar modelos espaços-temporais que melhor represente o comportamento pluviométrico no oeste paulista.
- Produzir mapas de superfície gerados a partir das técnicas estudadas que permitam fazer previsões que sejam úteis para a agricultura, pecuária e economia.

## **1.3. Estrutura do Trabalho**

Este trabalho será apresentado em forma de capítulos, sendo o primeiro, denotado Introdução, uma apresentação, justificativa e contextualização desta pesquisa.

No segundo capítulo é apresentada a fundamentação teórica que são conceitos relacionados às análises espacial e temporal como a Geoestatística Clássica e conceitos relacionados à análise espaço-temporal como modelos relacionados a essa abordagem.

E no terceiro capítulo são apresentados os resultados do estudo, descrevendo desde a organização do banco de dados, uma análise exploratória univariada e multivariada dos dados, até a análise utilizando-se Geoestatística Clássica e à modelagem espaço-temporal.

## 2. FUNDAMENTAÇÃO TEÓRICA

### 2.1. Geoestatística Clássica

De acordo com Landim (2002), a Geoestatística teve início por volta de 1951, quando um engenheiro de minas sul-africano Daniel. G. Krige, ao estudar dados de concentração de ouro chegou à conclusão que as variâncias entre as concentrações não faziam nenhum sentido caso não fosse considerada a distância entre elementos amostrais. No entanto, o processo de formalização de Geoestatística foi impulsionado com Matheron, ao dar desenvolvimento da Teoria das Variáveis Regionalizadas. Essas variáveis diferenciam das outras pelo fato de possuir coordenadas geográficas.

A Geoestatística é definida como uma área na qual são estudados fenômenos contínuos no espaço e/ou no tempo, a qual oferece um conjunto de ferramentas estatísticas voltadas para o entendimento e modelagem da variabilidade espacial.

Esta técnica é um tópico da Estatística Aplicada cujo objeto de estudo são as Variáveis Regionalizadas, que estimam seu comportamento através de funções de densidade de probabilidade.

#### 2.1.1. Semivariograma

Uma forma de analisar o comportamento da variabilidade espacial das variáveis regionalizadas é através do coeficiente de variograma ( $\hat{\gamma}$ ). Essa análise é feita com o intuito de verificar qual a dependência espacial existente entre um elemento amostral e outro, separados por uma determinada distância. O termo “dependência espacial” se refere à relação existente entre os elementos amostrais de acordo com uma variável; quando ela existir então é possível afirmar que os elementos amostrais próximos tendem a ser mais semelhantes, caso contrário, essa afirmativa não será verdadeira.

Assim, para uma variável regionalizada  $x(i)$ , coletada em diversos pontos  $i$ , o valor de cada ponto está relacionado de algum modo com os valores obtidos a partir de pontos situados a uma certa distância, sendo razoável pensar que a influência é tanto maior quanto menor for a distância entre os pontos (LANDIM, 2002).

Através de uma amostra  $Z(x_i)$ ,  $i = 1, 2, \dots, n$  o variograma pode ser estimado pela seguinte equação:

$$2\hat{\gamma}(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (1)$$

em que:  $2\hat{\gamma}(h)$  é o variograma estimado,  $N(h)$  é o número de pares de valores medidos que estão distantes por um vetor de módulo  $h$  e  $Z(x_i)$  e  $Z(x_i + h)$  são os valores da  $i$ -ésima observação da variável regionalizada, coletados nos pontos  $x_i$  e  $x_i + h$ , separados por  $h$  (JOURNEL, 1988).

Analogamente, é possível estimar o semivariograma, que corresponde à metade de um variograma, como mostra a equação:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (2)$$

Em um semivariograma experimental espera-se que a diferença entre dois elementos amostrais diminua à medida que  $h$  também diminui, pois as observações localizadas espacialmente mais próximas tendem a ser mais semelhantes do que as separadas por uma distância maior e, conseqüentemente, mais dissimilares.

A Figura 3 ilustra um semivariograma experimental com características muito próximas do ideal. O seu padrão representa o que, intuitivamente, se espera de dados de campo, isto é, que as diferenças  $[Z(x_i) - Z(x_i + h)]$  decresçam a medida que  $h$  cresce. Logo, é esperado que  $\hat{\gamma}(h)$  aumente com a distância  $h$  (JOURNEL, 1988).

Os parâmetros do semivariograma, ilustrado na Figura 3, são:

- Alcance (a): é a distância dentro da qual os elementos amostrais apresentam-se correlacionados espacialmente;
- Patamar (C): é o valor do semivariograma correspondente ao seu alcance (a). A partir desse ponto, considera-se que não existe mais dependência espacial entre os elementos amostrais, porque a variância da diferença entre os pares de amostras torna-se invariante com a distância (CAMARGO, 1997);
- Efeito Pepita ( $C_0$ ): é a medida da variabilidade quando a distancia  $h$  tende a 0. Por definição,  $\hat{\gamma}(0) = 0$ , entretanto, na pratica, à medida que  $h$  tende para 0 (zero),  $\hat{\gamma}(h)$  se aproxima de um valor positivo chamado efeito pepita;

- Contribuição ( $C_1$ ): é a diferença entre o patamar ( $C$ ) e o efeito pepita ( $C_0$ ).

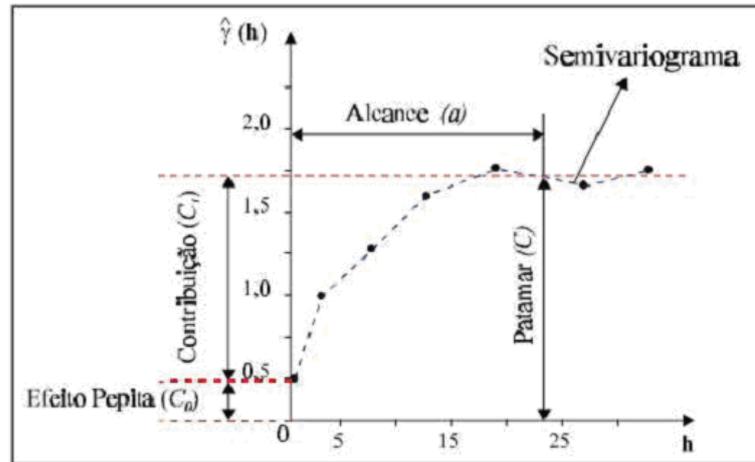


Figura 3: Exemplo de semivariograma experimental.

Fonte: Camargo (1997)

Para determinar um semivariograma experimental, no caso de todos os elementos amostrais estarem regularmente distribuídos, é necessário considerar algumas direções e realizar o cálculo de  $\hat{\gamma}(h)$  em todos os intervalos de distância  $h$ . A Geoestatística adota, convencionalmente, as direções de  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ . (Figura 4)

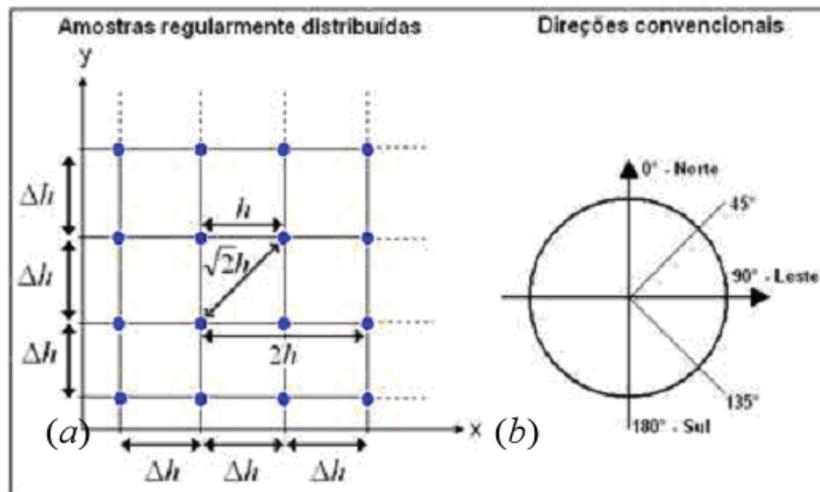


Figura 4: Elementos amostrais regularmente distribuídos (a) e as direções convencionais (b)

Fonte: Adaptada de Camargo (1997)

Dessa forma, se for definido uma distância  $h$  e uma direção  $Q^\circ$ , então todos os pares de pontos que atenderem a esse requisito serão incluídos no semivariograma, denotado por  $\hat{\gamma}(Q^\circ, h)$ .

Quando todos pares já forem analisados, os cálculos serão repetidos para  $\hat{\gamma}(Q^\circ, 2h)$ , assim, realiza-se esse processo até que a condição de parada ocorra.

Se os elementos amostrais estiverem irregularmente distribuídos, como no caso da Figura 5, será necessário introduzir novos parâmetros.

Dado um *Lag* (distância fornecida para o cálculo do semivariograma) de incremento  $I$  e tolerância  $T$ , e ainda uma direção  $D^\circ$  com tolerância angular  $A^\circ$ , fará com que qualquer par de pontos que esteja numa distância entre  $I+T$  e  $2I+T$  e numa direção qualquer compreendida entre  $(D-A)^\circ$  e  $(D+A)^\circ$  será adicionado ao cálculo do semivariograma. A largura da banda, como pode ser vista na Figura 5, é um valor que impede que o “leque” formado a partir da direção e da tolerância angular aumente muito, restringindo o número de pares, de maneira que ela não interfira no cálculo.

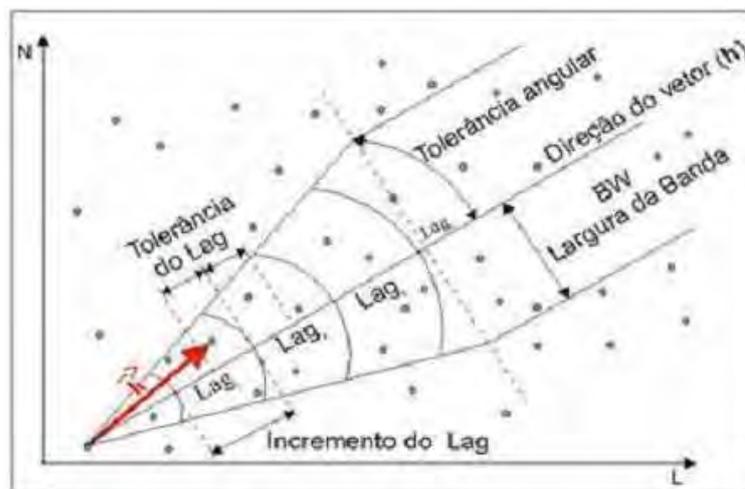


Figura 5: Parâmetros para o cálculo do semivariograma a partir dos elementos amostrais irregularmente distribuídos em duas dimensões.

Fonte: Adaptada de Camargo (1997)

### 2.1.2. Modelos Teóricos

O ajuste de um semivariograma, independente da forma como estão distribuídos os elementos amostrais, é baseado no fornecimento dos parâmetros relacionados na Figura 3, além de um modelo teórico que deve se ajustar da melhor forma possível aos valores encontrados durante o processo de análise direcional dos elementos amostrais, plotados em um gráfico bidimensional. Camargo (1997) salienta que o procedimento de ajuste não é direto e automático, mas sim interativo, no qual o pesquisador faz os ajustes necessários para obter o melhor modelo.

Dentre os modelos geralmente utilizados no processo de modelagem semivariográfica, citam-se os modelos com patamar: o esférico, o exponencial e o Gaussiano, que têm os seus formatos característicos ilustrados na Figura 6 e os modelos sem patamar como no caso do modelo potência ilustrado na Figura 7, sendo todos denominados de modelos isotrópicos conforme Isaaks e Srivastava (1989, p.374).

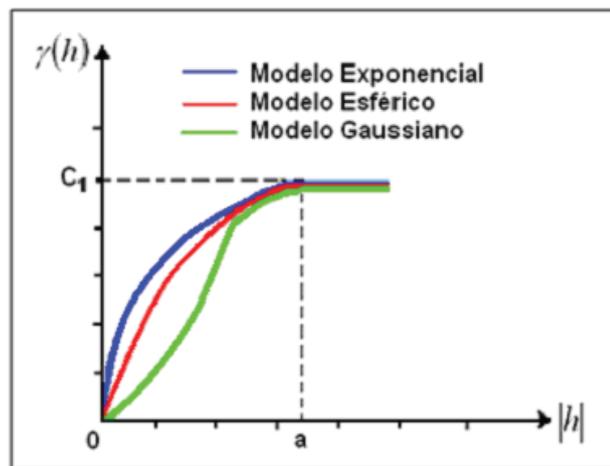


Figura 6: Comportamento característico dos modelos teóricos.  
Fonte: Camargo (1997)

O modelo esférico, um dos mais utilizados na modelagem dos fenômenos naturais, possui a equação dada por:

$$Sph(|h|) = \begin{cases} 0, & |h| = 0 \\ 1,5\left(\frac{|h|}{a}\right) - 0,5\left(\frac{|h|}{a}\right)^3, & 0 < |h| \leq a \\ 1, & |h| > a \end{cases} \quad (3)$$

A equação adaptada para o Semivariograma é dada por:

$$\gamma(h) = \begin{cases} 0 & , |h|=0 \\ C_o + C_1 \left[ \frac{3}{2} \left( \frac{|h|}{a} \right) - \frac{1}{2} \left( \frac{|h|}{a} \right)^3 \right] = C_o + C_1 [Sph(|h|)] & , 0 < |h| \leq a \\ C_o + C_1 & , |h| > a \end{cases} \quad (4)$$

O modelo exponencial possui equação normalizada dada por:

$$Exp(|h|) = \begin{cases} 0, & |h| = 0 \\ 1 - e^{-\left(\frac{|h|}{a}\right)}, & |h| \neq 0 \end{cases} \quad (5)$$

A equação adaptada para o Semivariograma é dada por:

$$\gamma(h) = \begin{cases} 0, & |h| = 0 \\ C_0 + C_1 \left[ 1 - e^{-\left(\frac{|h|}{a}\right)} \right] = C_0 + C_1 [Exp(|h|)], & |h| \neq 0 \end{cases} \quad (6)$$

O modelo Gaussiano, utilizado na modelagem de fenômenos extremamente contínuos, possui equação normalizada dada por:

$$Gau(|h|) = \begin{cases} 0, & |h| = 0 \\ 1 - e^{-\left(\frac{|h|}{a}\right)^2}, & |h| \neq 0 \end{cases} \quad (7)$$

A equação adaptada para o Semivariograma é dada por:

$$\gamma(|h|) = \begin{cases} 0, & |h| = 0 \\ C_0 + C_1 \left[ 1 - e^{-\left(\frac{|h|}{a}\right)^2} \right] = C_0 + C_1 [Gau(|h|)], & |h| \neq 0 \end{cases} \quad (8)$$

Um detalhe a favorecer a obtenção de semivariogramas experimentais bem ajustados aos modelos teóricos clássicos é a remoção da tendência (VALERIANO, 2002). Esse procedimento é necessário para atender à hipótese da Geoestatística de estacionaridade, que requer que as características estatísticas sejam constantes por toda a área analisada.

O modelo sem patamar é utilizado para modelar fenômenos que não possuem estacionaridade, isto é, apresentam tendência. O modelo potência faz parte dessa classe. Sua equação normalizada é dada por:

$$Pot(|h|) = \begin{cases} 0, & |h| = 0 \\ c|h|^e, & |h| \neq 0 \end{cases} \quad (9)$$

em que  $c$  é o coeficiente de declividade e  $e$  é o expoente.

A equação adaptada para o Semivariograma é dada por:

$$\gamma(|h|) = \begin{cases} 0, & |h| = 0 \\ C_0 + c|h|^e = C_0 + Pot(|h|), & |h| \neq 0 \end{cases} \quad (10)$$

O coeficiente de declividade ( $c$ ) influencia no grau de inclinação da curva. (Figura 7)

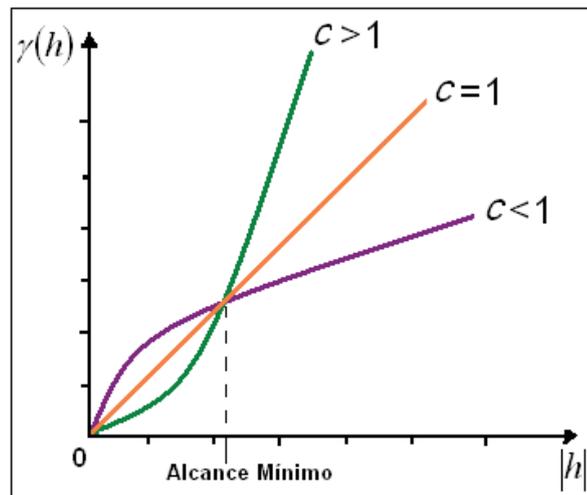


Figura 7: Representação do modelo Potência segundo os coeficientes de declividade

### 2.1.3. Anisotropia

Com a seleção do modelo teórico a ser utilizado e após o cálculo dos semivariogramas direcionais, deve-se analisar o comportamento dos dados. Quando os semivariogramas

direcionais não possuem comportamento semelhante em diferentes direções diz-se que há anisotropia. Por outro lado, se os modelos seguem comportamentos semelhantes nas diferentes direções, a distribuição então é considerada isotrópica.

A representação da distribuição de um fenômeno de caráter isotrópico geralmente é feita modelando um semivariograma omnidirecional, cujo parâmetro direcional que o caracteriza é a tolerância angular de  $90^\circ$  independente da direção tomada conforme a Figura 8. Já na existência da anisotropia, o problema toma certa complexidade, exigindo a realização da chamada modelagem anisotrópica, na qual acontece uma “combinação” entre semivariogramas que apresentam determinadas particularidades.

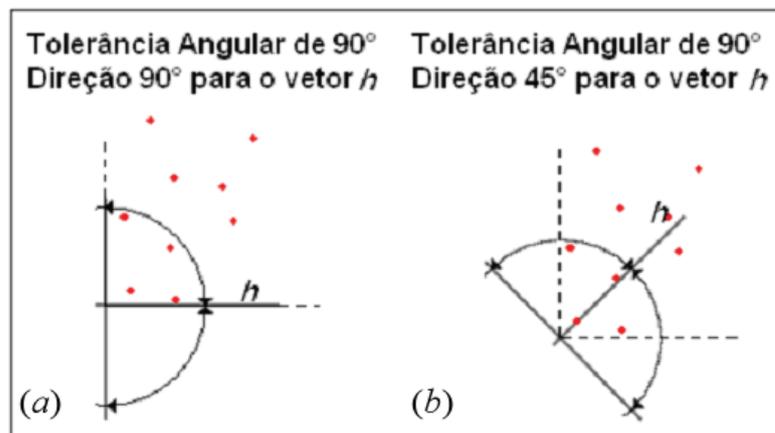


Figura 8: Tolerância angular de  $90^\circ$  com direção de  $90^\circ$  (a) e tolerância angular de  $90^\circ$  com direção de  $45^\circ$  (b).

#### 2.1.4. Modelagem Anisotrópica

Consiste em ajustar os modelos nas direções de maior e menor alcance (variabilidade) e combiná-los num único modelo. Essa modelagem é realizada nos casos de Anisotropia Geométrica, Zonal e Combinada. A Anisotropia Geométrica ocorre quando as direções de menor e maior variabilidade possuem mesmo Patamar ( $C$ ) e distintos parâmetros de alcance ( $a$ ), ambos com o mesmo tipo de modelos.

No caso da Anisotropia Zonal os modelos direcionais de maior e menor variabilidade possuem diferentes Patamares ( $C$ ) e mesmo alcance ( $a$ ), com a possibilidade de ocorrência de Efeito Pepita ( $C_0$ ) diferentes para cada modelo.

O caso mais comum dentre as anisotropia é a ocorrência da Anisotropia Combinada que é o resultado da combinação entre as Anisotropias Geométrica e Zonal, pois apresentam diferentes Patamares (C) e diferentes alcances (a) podendo apresentar diferentes Efeitos Pepita (C<sub>0</sub>).

### 2.1.5. Hipóteses Geoestatística

Para dar continuidade ao estudo geoestatístico, a variável regionalizada necessariamente tem que satisfazer pelo menos uma das hipóteses a seguir: (LANDIM, 2002)

- Hipótese de estacionariedade de 2ª ordem.
- Hipótese intrínseca.
- Hipótese de não existência de tendências regionais.

A hipótese de estacionariedade de 2ª ordem significa que não há tendências na região pelo fato de a componente determinística  $m(x)$  ser constante.

Portanto, se  $m(x)$  é igual ao valor esperado de  $Z$  na posição  $x$ , e a diferença entre  $Z(x)$  e  $Z(x+h)$  é nula e a covariância entre  $Z(x)$  e  $Z(x+h)$  existe e depende apenas do vetor distância  $h$ , pois:

$$C(h) = Cov [Z(x), Z(x+h)] = E[(Z(x) - m) \cdot (Z(x+h) - m)] = E[Z(x) \cdot Z(x+h)] - m^2, \forall x \quad (11)$$

O que implica na estacionariedade da variância, pois:

$$\begin{aligned} Var[Z(x)] &= E[Z(x) - m]^2 = E[Z^2(x)] - 2E[Z(x)] \cdot m + m^2 = \\ &= E[Z(x) \cdot Z(x+0)] - 2m^2 + m^2 = \\ &= E[Z(x) \cdot Z(x+0)] - m^2 = C(0), \forall x \end{aligned} \quad (12)$$

Que por sua vez implica na estacionariedade do variograma:

$$\begin{aligned} 2\gamma(h) &= E\{[Z(x) - Z(x+h)]^2\} = E\{Z^2(x) - 2Z(x)Z(x+h) + Z^2(x+h)\} = \\ &= E[Z^2(x)] - 2E[Z(x)Z(x+h)] + E[Z^2(x+h)] \end{aligned}$$

$$\text{Como: } E[Z^2(x)] = E[Z(x)Z(x+0)] = C(0) + m^2, \quad E[Z(x) \cdot Z(x+h)] = C(h) + m^2$$

$$\text{Então: } 2\gamma(h) = C(0) + m^2 - 2(C(h) + m^2) + C(0) + m^2 = 2C(0) - 2C(h)$$

$$\text{Simplificando: } \gamma(h) = C(0) - C(h) \quad (13)$$

A covariância e o semivariograma, sob a hipótese de estacionariedade de 2ª ordem, são ferramentas equivalentes para auxiliar na medição da dependência espacial.

A hipótese intrínseca baseia-se no fato da hipótese de estacionariedade de 2ª ordem não ser satisfeita em alguns casos.

Segundo a hipótese de estacionariedade de 2ª ordem, se existe a covariância, então existem a variância e o semivariograma da variável estudada, mas por outro lado, existem casos nas quais a capacidade de dispersão da variável é infinita, o que faz com que a covariância e a variância não existam, porém o semivariograma continua existindo. (LANDIM, 2002)

Nessa hipótese é admitido que a componente determinística que compõe a variável regionalizada,  $m(x)$ , é constante e a variância das diferenças depende somente da distância, ou seja:

$$\text{Var} [Z(x) - Z(x + h)] = E \left\{ [Z(x) - Z(x + h)]^2 \right\} = 2\gamma(h) \quad (14)$$

Esta hipótese é a mais freqüente na geoestatística, principalmente por ser menos restritiva, pois requer apenas a existência de estacionariedade do semivariograma sem restrições quanto a existência de variância finita.

A hipótese de tendência assume que a componente  $m(x)$  não é mais constante e sim representada por uma função determinística, conhecida por *Drift* ou Tendência Regional.

Neste caso é suposto que  $C(h)$  e  $\gamma(h)$  possuam estacionariedade dentro de uma vizinhança de tamanho restrito e  $m(x)$  varia regularmente dentro desta.

### 2.1.6. Krigagem

A krigagem é um processo de estimativa de valores de variáveis distribuídas no espaço e/ou no tempo, a partir de valores adjacentes enquanto considerados como interdependentes pelo semivariograma. Trata-se de um conjunto de técnicas de estimativa de superfície baseada na modelagem da estrutura de correlação espacial (CAMARGO, 1997).

A krigagem pode ser utilizada para:

- prever o valor de uma variável regionalizada em um determinado local dentro de um campo geométrico. Trata-se de um procedimento exato de interpolação que leva em consideração todos os valores observados e que pode ser a base da cartografia automatizada por meio de computadores quando se dispõe de valores de uma variável em uma determinada área;
- calcular a média de uma variável regionalizada para um volume maior que o suporte geométrico, como, por exemplo, no cálculo do teor médio de uma jazida a partir de informações obtidas de testemunhas de sondagem;
- estimativa do *Drift*, de modo análogo à análise de superfícies de tendência.

Conforme Flores (2000), a krigagem é o método que possibilita a melhor representação da continuidade dos fenômenos geográficos e, mais especialmente, do fenômeno pluvial, térmico, etc. O método ainda é capaz de fornecer, além dos valores estimados, o erro associado à estimativa, o que o faz distinto dos outros métodos de interpolação como por exemplo, o Inverso da Distância.

A krigagem utiliza informações do semivariograma definido e modelado de acordo com o comportamento e a distribuição espacial da variável em estudo para encontrar os pesos adequados a serem associados às amostras que irão estimar um ponto, área ou volume. Busca fornecer estimativas não-tendenciosas e com variância mínima, ou seja, a diferença entre os valores estimados e verdadeiros para o mesmo local deve ser nula (LANDIM, 2002).

O processo de krigagem estima um valor de um atributo, em uma posição  $u$  não amostrada, a partir de um conjunto de amostras vizinhas  $z(u_a)$ ,  $a = 1, \dots, n$ . O método de krigagem é semelhante ao de interpolação por média móvel ponderada, porém, na krigagem, os pesos dados a cada observação são determinados a partir de uma pré-análise espacial utilizando semivariogramas experimentais.

No total, a krigagem reúne mais de 30 tipos de métodos de estimação. Particularmente, serão discutidas as krigagens: simples, ordinária, universal e cokrigagem.

#### **2.1.6.1. Krigagem Simples**

Considere uma superfície composta de uma Variável Regionalizada,  $Z$ , amostrada em  $n$  locais distintos. Cada elemento amostral é localizado espacialmente por um vetor  $u_i$  que

representa uma posição em duas dimensões dada pelo par  $(x_i, y_i)$ . Assim, o valor de cada elemento amostral é denotado por  $z(u_i)$ .

Segundo Journel (1988), o objetivo é estimar o valor da variável  $Z$  num ponto desconhecido  $u_0$  com base numa combinação linear dos  $n$  valores observados com seus respectivos pesos  $\lambda_i$  e adicionando um parâmetro  $\lambda_0$ , assim a estimativa é dada por:

$$\hat{Z}_{u_0} = \lambda_0 + \sum_{i=1}^n \lambda_i \cdot Z(u_i) \quad (15)$$

É desejado que o estimador seja não tendencioso, ou seja, a diferença entre o valor amostrado e seu valor estimado seja nula. Isto é,

$$E\left[Z_{u_0} - \hat{Z}_{u_0}\right] = 0 \Rightarrow E\left[Z_{u_0}\right] = E\left[\hat{Z}_{u_0}\right] \quad (16)$$

Como a esperança dos valores estimado e amostrado são iguais, é possível afirmar que

$$E\left[\hat{Z}_{u_0}\right] = E\left[\lambda_0 + \sum_{i=1}^n \lambda_i \cdot Z(u_i)\right] = \lambda_0 + \sum_{i=1}^n \lambda_i \cdot E\left[Z(u_i)\right] \quad (17)$$

Logo, o parâmetro  $\lambda_0$  é dado por

$$\lambda_0 = E\left[\hat{Z}_{u_0}\right] - \sum_{i=1}^n \lambda_i \cdot E\left[Z(u_i)\right] \quad (18)$$

Assim, o estimador de Krigagem Simples é

$$\hat{Z}_{u_0} = E\left[Z_{u_0}\right] - \sum_{i=1}^n \lambda_i E\left[Z(u_i)\right] + \sum_{i=1}^n \lambda_i Z(u_i) \quad (19)$$

Neste caso, o método da Krigagem Simples supõe que a média  $m$  é constante, portanto:

$$E[Z_{u_0}] = E[Z_{u_i}] = m \quad (20)$$

O estimador de Krigagem Simples se transforma numa forma mais simples:

$$\hat{Z}_{u_0} = m + \sum_{i=1}^n \lambda_i [Z(u_i) - m] \quad (21)$$

Neste momento, Journel (1988), mostra que minimizando a variância do erro ( $\text{Var}[Z_{u_0} - Z_{u_0}^*]$ ), os pesos  $\lambda_i$  são obtidos a partir do seguinte sistema de equações, nomeado de sistema de Krigagem Simples.

$$\sum_{j=1}^n \lambda_j C(u_i, u_j) = C(u_i, u_0) \quad i = 1, \dots, n \quad (n \text{ equações}) \quad (22)$$

em que  $C(u_i, u_j)$  refere-se à função covariância correspondente a um vetor,  $h$ , com origem em  $u_i$  e extremidade em  $u_j$ , e  $C(u_i, u_0)$  refere-se à função semivariância correspondente a um vetor  $h$  com origem em  $u_i$  e extremidades no ponto a ser estimado  $u_0$ .

Um exemplo, retirado de Camargo (1997), com duas equações e duas incógnitas, seria:

$$\begin{cases} \lambda_1 C_{11} + \lambda_2 C_{12} = C_{10} \\ \lambda_1 C_{21} + \lambda_2 C_{22} = C_{20} \end{cases} \quad (23)$$

A correspondente variância minimizada, chamada de variância de Krigagem Simples é expressa da seguinte forma:

$$\sigma_{ks}^2 = \text{Var}[Z_{u_0} - \hat{Z}_{u_0}] = C(0) - \sum_{i=1}^n \lambda_i C(u_i, u_0) \quad (24)$$

Na notação matricial, o sistema é escrito da seguinte forma:

$\mathbf{K}\boldsymbol{\lambda} = \mathbf{k} \Rightarrow \boldsymbol{\lambda} = \mathbf{K}^{-1}\mathbf{k}$ , em que:

$$\mathbf{K} = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{bmatrix}, \quad \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} \quad \text{e} \quad \mathbf{k} = \begin{bmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{n0} \end{bmatrix}$$

$\mathbf{K}$  e  $\mathbf{k}$  são matrizes de covariâncias e  $\boldsymbol{\lambda}$  o vetor dos pesos.

### 2.1.6.2. Krigagem Ordinária

A krigagem ordinária é uma técnica de estimativa linear para uma variável regionalizada que tem sua média desconhecida, ou seja, considera-se a média flutuante ou móvel por toda a área, ao contrário da Krigagem Simples que no caso da estacionaridade de segunda ordem exige-se que a média seja conhecida.

Trata-se de uma abordagem de interpolação (geoestatística) na qual é assumida a hipótese de quase-estacionaridade, também conhecida como estacionaridade regional (LANDIM, 2002).

A formulação matemática do método da interpolação por Krigagem Ordinária é semelhante ao método da Krigagem Simples. A diferença entre esses métodos está no fato da Simples considerar a média conhecida e a Ordinária desconhecida.

Com base na relação (16) é possível concluir que as esperanças são iguais, logo

$$E[Z_{u_0}] = E\left[\lambda_0 + \sum_{i=1}^n \lambda_i \cdot Z(u_i)\right] \Rightarrow m = \lambda_0 + \sum_{i=1}^n \lambda_i \cdot m \quad (25)$$

Como já mencionado, a Krigagem Ordinária não necessita do conhecimento da média, logo, para validade da Equação (25) necessariamente o parâmetro  $\lambda_0$  deve ser nulo e a somatória dos pesos deve ser igual a 1.

Dessa forma, o estimador da Krigagem Ordinária é definido por:

$$\hat{Z}_{u_0} = \sum_{i=1}^n \lambda_i \cdot Z(u_i), \text{ com } \sum_{i=1}^n \lambda_i = 1 \quad (26)$$

Mínimizar a variância do erro,  $Var(Z_{u_0} - \hat{Z}_{u_0})$ , desde que a somatória dos pesos seja igual a 1, os pesos  $\lambda_i$  são obtidos do seguinte sistema de equação:

$$\begin{cases} \sum_{j=1}^n \lambda_j C(u_i, u_j) - \alpha = C(u_i, u_0) & i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j = 1 \end{cases} \quad (27)$$

na qual  $C(u_i, u_j)$  e  $C(u_j, u_0)$  são as semivariâncias entre os pontos  $u_i$  e  $u_j$ ,  $u_j$  e  $u_0$  e  $\alpha$  é o multiplicador de Lagrange.

Segundo Journel (1988), a variância minimizada do erro é dada por:

$$\sigma_{ko}^2 = Var\left[Z_{u_0} - \hat{Z}_{u_0}\right] = C(0) - \sum_{i=1}^n \lambda_i C(u_i, u_0) - \alpha \quad (28)$$

A Equação (28) é capaz de fornecer importantes informações sobre a confiabilidade dos valores interpolados.

Na notação matricial, o sistema é escrito da seguinte forma:

$\mathbf{K}\boldsymbol{\lambda} = \mathbf{k} \Rightarrow \boldsymbol{\lambda} = \mathbf{K}^{-1}\mathbf{k}$ , em que:

$$\mathbf{K} = \begin{vmatrix} C_{11} & C_{12} & \dots & C_{1n} & 1 \\ C_{21} & C_{22} & \dots & C_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{vmatrix}, \quad \boldsymbol{\lambda} = \begin{vmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \alpha \end{vmatrix} \text{ e } \mathbf{k} = \begin{vmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{n0} \\ 1 \end{vmatrix}$$

$\mathbf{K}$  e  $\mathbf{k}$  são matrizes de covariâncias e  $\boldsymbol{\lambda}$  o vetor dos pesos.

### 2.1.6.3. Krigagem Universal

Se os dados amostrados apresentarem tendências, existem duas metodologias que trabalham com tendência. A primeira proposta por Matheron que trabalha sem a decomposição da variável regionalizada e a segunda proposta por Journel que decompõe essa variável resultando em resíduos e superfície de tendência denominada de Krigagem Universal que será descrita a seguir.

A Variável Regionalizada que apresenta tendência pode ser decomposta em dois componentes, um *drift* e um resíduo. O *drift* é o valor médio dessa variável em certa vizinhança que varia sistematicamente e os resíduos são os valores resultantes da diferença dos valores reais e o *drift*.

Logo uma variável com tendência pode ser expressa matematicamente como:

$$Z(u) = m(u) + R(u) \quad (29)$$

Em que:

$m(u)$  representa o *drift*.

$R(u)$  representa os resíduos.

A estimação da Variável Regionalizada é a soma da superfície de resíduos estimada com a superfície que representa o *drift*.

O método de interpolação por Krigagem Universal pode ser entendido como a combinação entre Superfície de Tendência e Krigagem Ordinária.

### 2.1.7. Cokrigagem

Segundo Deutsch e Journel (1998), o termo Cokrigagem é destinado à regressão linear que utiliza dados de diferentes atributos.

É uma extensão multivariada da Krigagem, na qual cada elemento amostral é obtido de um vetor de valores, ao invés de um único valor.

Neste procedimento geoestatístico são utilizadas duas ou mais variáveis amostradas num mesmo espaço geográfico apresentando certo grau de correlação.

A utilização da cokrigagem é voltada para os casos que uma das variáveis em estudo apresente à quantidade de elementos amostral inferior a outra variável com número de elementos amostral mais denso, desde que haja correlação entre essas variáveis.

A estimativa por Cokrigagem é uma combinação linear de dados primários e secundários. (Equação (30))

$$\hat{U}_{cok}(u) = \sum_{i=1}^{n_1} a_i u_i + \sum_{j=1}^{n_2} b_j v_j \quad (30)$$

em que:

$\hat{U}_{cok}(u)$  é a estimativa no ponto  $u$ ,

$u_i$  são os dados primários próximos de  $u$ ,

$v_j$  são os dados secundários próximos de  $u$ ,

$a_i$  e  $b_j$  são os pesos determinados para as amostras.

Inicialmente, define-se a estimativa dos erros pela seguinte equação:

$$R = \hat{U}_{cok} - U_{cok} = \sum_{i=1}^{n_1} a_i U_i + \sum_{j=1}^{n_2} b_j V_j - U_{cok} \quad (31)$$

As variáveis  $U_i$  e  $V_i$  representam os fenômenos  $U$  e  $V$  nas  $n$  localizações, respectivamente.

A Equação (31) gera a matriz escrita por:

$$\mathbf{R} = \mathbf{k}^t \mathbf{U} \quad (32)$$

em que:

$$\mathbf{k}^t = (a_1, \dots, a_n; b_1, \dots, b_n; -1),$$

$$\mathbf{U}^t = (U_1, \dots, U_n; V_1, \dots, V_n; U_0).$$

A expressão da variância de  $R$  é calculada a partir da Equação (33), obtida com algumas simplificações. Utilizando os multiplicadores de Lagrange é possível obter:

$$Var(R) = Cov\{U_0 U_0\} + \mu_1 - \sum_{i=1}^n a_i Cov\{U_i U_0\} - \sum_{j=1}^n b_j Cov\{V_j U_0\} \quad (33)$$

### 2.1.7.1. Semivariograma Cruzado

A cokrigagem possui uma ferramenta semelhante à krigagem que calcula a dependência espacial denominada semivariograma cruzado. (LANDIM, 2002)

O semivariograma cruzado calcula a variabilidade entre duas variáveis separadas por uma distância  $h$  comum, assim, representado na Equação (34).

$$\gamma_{ZY}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (z_i - z_i') \cdot (y_i - y_i') \quad (34)$$

em que:

$z_i$  é o valor da variável primária  $Z$  atribuída à origem do vetor “distância”  $h$ .

$z_i'$  é o valor da variável primária  $Z$  atribuída ao fim do vetor “distância”  $h$ .

$y_i$  é o valor da variável secundária  $Y$  atribuída à origem do vetor “distância”  $h$ .

$y_i'$  é o valor da variável secundária  $Y$  atribuída ao fim do vetor “distância”  $h$ .

$N(h)$  é o número de pares.

## 2.2. Modelagem Espaço-temporal

Modelos espaços-temporais se tornam um importante instrumento de compreensão de um determinado fenômeno pelo fato de estarem associados em conjunto o espaço e o tempo. Nesse capítulo serão apresentados conceitos da modelagem espaço-temporal e alguns modelos mais utilizados na literatura.

A respeito da utilização de modelos geoestatísticos espaço-tempo, existem alguns trabalhos envolvendo diversas ciências como agronomia (HOOSBEEK, 1998), hidrologia (ROUHANI & WACKERNAGEL, 1990; GOOVAERTS & SONNET, 1993), meio ambiente

(CHRISTAKOS & VYAS, 1998), meteorologia (SOARES et al., 1996) e ciência do solo (SNEPVANGERS et al., 2003).

A maioria dos modelos espaços-temporais trabalha com estruturas de covariância das funções aleatórias espaços-temporais  $Z(s,t)$ , em que  $(s,t) \in \mathfrak{R}^d \times \mathfrak{R}$ , generalizando uma função aleatória gaussiana espaço-temporal. A covariância entre  $z(s_1,t_1)$  e  $z(s_2,t_2)$  depende das coordenadas espaço-temporais  $(s_1,t_1)$  e  $(s_2,t_2)$ . (GNEITING et al., 2005 *apud* LIMA, 2006) Na prática são exigidas três propriedades como estacionaridade, separabilidade e simetria total que serão apresentadas a seguir.

No caso da estacionaridade, a função aleatória espaço-temporal (FAET) tem covariância espacial estacionária se  $Cov\{Z(s_1,t_1), Z(s_2,t_2)\}$  depende somente das localidades  $s_1$  e  $s_2$  pelo vetor de separação espacial  $s_1 - s_2$ . No caso temporal, a covariância temporal estacionária depende somente do vetor de separação temporal  $t_1 - t_2$  como apresenta a Equação (35).

$$Cov\{Z(s_1,t_1), Z(s_2,t_2)\} = C(s_1 - s_2, t_1 - t_2) \quad (35)$$

em que:  $C$  é denotado a função de covariância espaço-tempo do processo  $(s_1,t_1)$  e  $(s_2,t_2) \in \mathfrak{R}^d \times \mathfrak{R}$  e  $C(\cdot,0)$  e  $C(0, \cdot)$  são funções de covariância puramente espacial e puramente temporal.

O campo aleatório  $Z$  possui covariância separável se existir componentes de funções de covariância puramente espacial ( $C_S$ ) e puramente temporal ( $C_T$ ) assim dado pela Equação (36).

$$Cov\{Z(s_1,t_1), Z(s_2,t_2)\} = C_S(s_1 - s_2) \cdot C_T(t_1 - t_2) \quad (36)$$

em que: as coordenadas espaços-temporais  $(s_1,t_1)$  e  $(s_2,t_2) \in \mathfrak{R}^d \times \mathfrak{R}$ .

Gneiting et al. (2005) comentam que essas estruturas permitem uma inferência e estimação computacionalmente eficiente pelo fato de considerar separadamente as estruturas espacial e temporal.

O processo espaço-tempo  $Z$  tem covariância totalmente simétrica se:

$$Cov\{Z(s_1,t_1), Z(s_2,t_2)\} = Cov\{Z(s_1,t_2), Z(s_2,t_1)\} \quad (37)$$

em que: as coordenadas espaços-temporais  $(s_1, t_1)$  e  $(s_2, t_2) \in \mathfrak{R}^d \times \mathfrak{R}$ .

Gneiting, Genton e Guttorp (2006) aplicaram modelos geoestatísticos espaço-tempo aos dados de velocidade de vento. Chandler *et al.* (2002) utilizaram os dados pluviométricos aplicando modelos generalizados (GLM).

De Iaco, Myers e Posa (2002) utilizaram modelos espaço-tempo na estimação de poluição total do ar no município de Milão na Itália e os denominaram de: modelos métricos, produtos espaço-tempo, produto-soma espaço-tempo, integrado produto-soma, não-separável e geral que serão apresentados posteriormente.

Algumas classes de modelos de covariância espaço-temporal têm se apresentado na literatura como o caso do modelo de covariância produto-soma e o modelo de covariância produto-soma generalizado, que foi descrito por De Iaco, Myers e Posa (2002).

Considere o campo aleatório espaço-tempo de segunda ordem estacionário:

$$Z = \{Z(s, t), (s, t) \in D \times T\} \quad (38)$$

em que:  $D \subset \mathfrak{R}^d$  e  $T \subset \mathfrak{R}_+$ , com covariância e semivariograma:

$$C_{ST}(h) = Cov(Z(s + h_s, t + h_t), Z(s, t)) \quad (39)$$

$$\gamma_{ST}(h) = Var[Z(s + h_s, t + h_t) - Z(s, t)] / 2 \quad (40)$$

em que:  $h = (h_s, h_t)$ ,  $(s, s + h_s) \in D^2$  e  $(t, t + h_t) \in T^2$ .

No **modelo métrico**, proposto por Dimitrakopoulos e Luo (1994)<sup>1</sup> *apud* De Iaco, Myers e Posa (2002), é assumido que:

$$C_{ST}(h_s, h_t) = C(a^2 |h_s|^2 + b^2 h_t^2) \quad (41)$$

---

<sup>1</sup> DIMITRAPOULOS, R.; LUO, X. Spatiotemporal modeling: covariances and ordinary kriging systems.

em que: os coeficientes  $a, b \in \mathfrak{R}$ . Nota-se que pela Equação (41), o mesmo tipo de modelo é assumido para as covariâncias espacial e temporal, com possíveis mudanças no alcance.

No **modelo produto**, proposto por De Cesare *et al.* (1997)<sup>2</sup> *apud* De Iaco, Myers e Posa (2002), a covariância espaço-temporal é dado por:

$$C_{ST}(h_s, h_t) = C_s(h_s) \cdot C_t(h_t) \quad (42)$$

em que: a dependência espacial é separada da temporal,  $C_s$  é uma função definida positiva em  $\mathfrak{R}^d$  e  $C_t$  é uma função definida positiva em  $\mathfrak{R}$ .

Os trabalhos de Cressie (1993)<sup>3</sup> *apud* De Iaco, Myers e Posa (2002) encontram modelos admissíveis de covariância espacial e temporal que podem ser a combinação de modelos produtos determinando modelos espaços-temporais de covariância.

O **modelo linear**, um outro tipo de separabilidade implica na adição das covariâncias espacial e temporal (ROUHANI AND HALL, 1989<sup>4</sup>, *apud* DE IACO, MYERS E POSA, 2002) dado por:

$$C_{ST}(h_s, h_t) = C_s(h_s) + C_t(h_t) \quad (43)$$

Para este modelo, Myers e Journel (1990)<sup>5</sup> *apud* De Iaco, Myers e Posa (2002) a função de covariância é somente positiva semidefinida, sendo insatisfatórias para predições ótimas.

---

<sup>2</sup> De CESARE, L. MYERS, D. E. POSA, D., Spatial temporal modeling of SO<sub>2</sub> in the Milan district. In: Baafi, E. Y., Schofield, N.A. (Eds.), *Geostatistics Wollongong '96*, Vol.2. Kluwer Academic Publishers, Dordrecht, pp. 1031-1042, 1997.

<sup>3</sup> CRESSIE, N. *Statistics for spatial data*. New York: Willey, 1993.

<sup>4</sup> ROUHANI, S., HALL, T.J., *Space-time kriging of groundwater data*. *Geostatistics*. Kluwer Academic Publishers, Dordrecht, Vol. 2, pp. 639-651. 1989.

<sup>5</sup> MYERS, D. E., JOURNEL, A. G. Variograms with sazonal anisotropies and non-invertible kriging systems. *Math Geol.* 22, 779-785, 1990.

O **modelo não-separável** é uma nova abordagem que permite obter classes de não-separabilidade; a função de covariância espaço-temporal estacionária tem sido derivada de Cressie e Hung (1999)<sup>6</sup> *apud* De Iaco, Myers e Posa (2002).

$$C_{ST}(h_s, h_t) = \int e^{ih_s^T \omega} \rho(\omega, h_t) k(\omega) d\omega \quad (44)$$

em que as duas condições seguintes são satisfeitas:

- para cada  $\omega \in \mathfrak{R}^d$ ,  $\rho(\omega, \cdot)$  é a função contínua de autocorrelação e  $k(\omega) > 0$ ;
- a função positiva  $k(\omega)$  satisfaz:

$$\int k(\omega) d\omega < \infty$$

O **modelo de covariância produto-soma** pode ser obtido da seguinte forma:

$$C_{ST}(h_s, h_t) = k_1 C_s(h_s) C_t(h_t) + k_2 C_s(h_s) + k_3 C_t(h_t) \quad (45)$$

Ou, equivalente, em termos de semivariograma:

O **variograma produto-soma** pode ser obtido da seguinte forma:

$$\gamma_{ST}(h_s, h_t) = [k_2 + k_1 C_t(0)] \gamma_s(h_s) + [k_3 + k_1 C_s(0)] \gamma_t(h_t) - k_1 \gamma_s(h_s) \gamma_t(h_t) \quad (46)$$

em que:  $C_s$  e  $C_t$  são funções de covariância,  $\gamma_s$  e  $\gamma_t$  são os correspondentes semivariogramas espacial e temporal e  $k_1 > 0$ ,  $k_2 \geq 0$ ,  $k_3 \geq 0$  para assegurar admissibilidade. Note que  $C_{st}(0,0)$  é a patamar de  $\gamma_{st}$  (patamar “global”),  $C_s(0)$  é o patamar de  $\gamma_s$  e  $C_t(0)$  é o patamar de  $\gamma_t$ .  $C_s(0)$  e  $C_t(0)$  são denominados patamares “parciais”.

---

<sup>6</sup> CRESSIE, N., HUANG, H., Classes of nonseparable, spatio-temporal stationary covariance functions. JASA 94, 1330-1340, 1999.

O **modelo produto-soma generalizado** é uma generalização do modelo de covariância produto-soma introduzido por De Cesare et al. (2001) *apud* De Iaco, Myers e Posa (2002) dada pela função semivariográfica:

$$\gamma_{ST}(h_s, h_t) = \gamma_{st}(h_s, 0) + \gamma_{st}(0, h_t) - k\gamma_{st}(h_s, 0)\gamma_{st}(0, h_t) \quad (47)$$

em que:  $\gamma_{st}(h_s, 0)$  e  $\gamma_{st}(0, h_t)$  são semivariogramas espacial e temporal válidos, respectivamente, e

$$k = \frac{(sill\gamma_{st}(h_s, 0) + sill\gamma_{st}(0, h_t) - sill\gamma_{st}(h_s, h_t))}{(sill\gamma_{st}(h_s, 0)sill\gamma_{st}(0, h_t))} \quad (48)$$

em que: *sill* indica o patamar do semivariograma.

O patamar “global”  $sill\gamma_{st}(h_s, h_t)$  pode ser estimado graficamente através da superfície do semivariograma espaço-temporal. Para isto, o seguinte condição de admissibilidade para  $\gamma_{st}(h_s, h_t)$  deve ser satisfeita:

$$0 < k \leq \frac{1}{\max\{sill(\gamma_{st}(h_s, 0)), sill(\gamma_{st}(0, h_t))\}} \quad (49)$$

em que:  $k$  é selecionado para garantir que  $\gamma_{st}(h_s, h_t)$  seja uma função positiva definida.

De Iaco, Myers e Posa (2002) fazem algumas considerações sobre os modelos estudados em seu artigo (na Seção 4.2):

- se a função de autocorrelação  $\rho$  (Equação (47)) for puramente uma função de  $h_t$ , então o modelo produto é obtido;
- os modelos produto e linear são facilmente obtidos a partir do modelo produto-soma adotando, respectivamente,  $k_2 = k_3 = 0$  e  $k_1 = 0$ ;
- os modelos de covariância produto-soma e de covariância produto-soma generalizado são não-separáveis e, em geral, são não-integráveis;

- os modelos de covariância produto-soma e o generalizado são mais flexíveis que os modelos de covariância não-separáveis para a estimar e modelar estruturas de correlação espaço-temporais.

Lima (2006) aplicou o modelo produto-soma generalizado aos dados de índice de vegetação para compreender o comportamento das plantas aquáticas no reservatório de Salto Grande localizado no município de Americana/SP.

O **modelo geostatístico espaço-temporal** estudado e aplicado nesse trabalho é o desenvolvido por Høst, Omre and Switzer (1995) que está descrito a seguir.

Sejam dadas  $n$  localizações no espaço  $u$  ( $u_1, u_2, \dots, u_n$ ) e em cada localização há uma série temporal de  $m$  tempos distintos ( $t_1, t_2, \dots, t_m$ ) com a variável aleatória  $Z(u_i, t_j)$ ,  $i = 1, 2, \dots, n$ ;  $j = 1, 2, \dots, m$ , de interesse. O modelo proposto para cada localização  $u$  em um tempo  $t$  específico é dado na Equação (50).

$$Z(u, t) = M_1(u) + R_2(t) + S(u)S(t)R(u, t) \quad (50)$$

em que:  $M_1(u)$  é o componente que representa a variação da média espacial,  $R_2(t)$  é o componente com média zero que funciona como um fator de correção temporal de  $M_1(u)$ ,  $S(u)$  e  $S(t)$  são os componentes espaço-tempo que modela o desvio padrão da variável aleatória  $Z$  e  $R(u, t)$  é o resíduo espaço-temporal com média igual a 0 e desvio padrão igual a 1.

A estimação sempre é feita numa localização fixa ( $u_0$ ) para toda a série temporal ( $t$ ) conforme a Equação (51)

$$\hat{Z}(u_0, t) = \hat{M}_1(u_0) + \hat{R}_2(t) + \hat{S}(u_0)\hat{S}(t)\hat{R}(u, t) \quad (51)$$

As equações das estimativas para calcular (51) são dadas a seguir: (KYRIAKIDIS e JOURNAL, 1999; RIBEIRO e MINGOTI, 2007)

$$M_1(u_i) = \frac{1}{T} \sum_{j=1}^T Z(u_i, t_j) \quad (52)$$

em que:  $M_I$  representa a média da série temporal para a localização  $u_i$ .

$$\hat{M}_1(u_0) = \sum_{\{u_i \in V_{u_0}\}} \lambda(u_i) M_1(u_i) \quad (53)$$

em que:  $\lambda$  são os pesos atribuídos por krigagem e  $\hat{M}_1(u_0)$  estimado por krigagem ordinária.

$$\hat{R}_2(t) = \sum_{i=1}^n \lambda(u_i) (Z(u_i, t) - \hat{M}_1(u_i)) \quad (54)$$

$$\hat{S}(u_0) = \sum_{\{u_i \in V_{u_0}\}} \lambda(u_i) \hat{S}_1(u_i) \quad (55)$$

$$\hat{R}(u, t) = \frac{[Z(u, t) - \hat{M}_1(u_i) - \hat{R}_2(t)]}{[\hat{S}_1(u_i)] \times [\hat{S}_2(t)]} \quad (56)$$

em que:

$$\hat{S}_1^2(u_i) = \frac{1}{T} \sum_{j=1}^T (Z(u, t_j) - \hat{M}_1(u_i) - \hat{R}_2(t_j))^2 \quad (57)$$

$$\hat{S}_2^2(t) = \frac{1}{w^2} \sum_{i=1}^n \lambda(u_i) (Z(u_i, t) - \hat{M}_1(u_i) - \hat{R}_2(t))^2 \quad (58)$$

$$w^2 = \sum_{i=1}^n \lambda(u_i) \hat{S}_1^2(u_i) \quad (59)$$

Os autores Kyriakidis e Journel (1999) e Ribeiro e Mingoti (2007) estudaram e aplicaram esse modelo, sendo que o primeiro trabalho apresentou uma revisão da modelagem espaço-temporal e o segundo apresentou uma metodologia computacional para este modelo espaço-temporal posteriormente adotada nessa dissertação.

A **Máxima Entropia Bayesiana** (MEB ou BME, do inglês, *Bayesian Maximum Entropy*) ou Moderna Geoestatística Espaço-Temporal é outro método de análise desenvolvido pelo George Christakos e colaboradores desde a década de 1980. Essa metodologia permite realizar o mapeamento espaço-temporal e sua extrapolação temporal da variável em estudo e difere dos demais métodos de interpolação pelo fato de adquirir os dados de diferentes fontes e tipos de conhecimento. Outro destaque dessa metodologia é a incorporação sistemática de dados *soft*, dados que são incertos, expressos por curvas ou intervalos de probabilidade. As medidas feitas com precisão (observações disponíveis) são denominadas de dados *hard*. (CHRISTAKOS, 2000)

A MEB, ao contrário de outras técnicas de interpolação, não necessita assumir nenhuma distribuição a priori, como é necessário assumir uma distribuição gaussiana para diversas técnicas de krigagem, podendo incorporar momentos multivariados não-gaussianos de diversas formas. O processo de mapeamento é dado através da escolha do preditor de acordo com as suas necessidades. Suas principais formas são: MEB moda, MEB mediana, MEB média condicional (equivalente a Krigagem Ordinária) e MEB percentil (equivalente a Krigagem Indicativa). Seus resultados gerados fornecem estimativas mais acuradas pelo fato de incorporar sistematicamente de diversas fontes de informação. Este método é a união de diversas metodologias, como a Análise Estocástica, Teoria da Informação, Geoestatística e Estatística Bayesiana.

Os autores Douaik et al. (2005) fizeram um mapeamento da salinidade do solo aplicando o método MEB. Money et al. (2009) avaliaram a contaminação fecal ao longo da Bacia do Rio Raritam ao norte de Nova Jersey nos Estados Unidos. Bogaert & D'Or (2001) reconstruíram mapas de valores contínuos com o método da MEB e comparar a eficiência deste método com o método mais utilizado em ciências do solo o Método de Quantificação das Legendas (MQL).

### 3. EXPERIMENTO: DESCRIÇÃO, MÉTODOS E RESULTADOS

#### 3.1. Banco de dados

Os dados de precipitação pluviométrica do Estado de São Paulo estão disponibilizados pelo Departamento de Águas e Energia Elétrica (DAEE) no endereço eletrônico <<http://www.dae.sp.gov.br/cgi-in/Carrega.exe?arq=/hidrometeorologia/bancodados.htm>>. A Figura 9 ilustra o banco de dados nesse endereço eletrônico.

Banco de Dados Pluviométricos do Estado de São Paulo												
Pesquisar por:												
Município	Municípios	Prefixo	Tipo de Dados:									
Município	Presidente Prudente	D8-003	Chuva mensal									
OK												
Município	Prefixo	Nome	Altitude	Latitude	Longitude	Bacia						
Presidente Prudente	D8-003	Presidente Prudente	460 m	22°06'	51°23'	Mandaguari						
<a href="#">Download da Série Histórica (1936 até 2004)</a>												
CHUVA MENSAL (mm)												
Ano	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1936	---	---	---	---	---	---	---	---	---	---	---	286,0
1937	268,5	189,0	265,8	72,5	15,5	70,8	4,3	86,0	7,0	191,3	191,9	101,3
1938	---	---	---	---	---	---	---	---	---	---	---	---
1939	---	---	---	---	---	---	---	0,0	---	---	---	---
1940	---	---	---	---	---	---	---	---	---	---	---	---
1941	---	---	---	---	---	---	---	---	---	---	---	---
1942	---	---	---	---	---	---	---	---	---	---	63,5	169,0
1943	79,7	155,0	201,0	50,0	3,0	26,0	0,0	0,0	120,0	15,4	154,0	108,0
1944	97,0	170,0	200,0	53,0	23,0	38,0	2,0	0,0	30,0	150,0	205,0	20,0
1945	214,1	175,5	131,6	74,6	21,8	221,2	51,7	9,5	43,5	94,3	55,0	---
1946	---	---	---	---	---	---	---	---	---	---	---	---
1947	230,5	376,9	103,7	21,2	150,6	1,2	6,5	17,5	145,2	95,0	4,7	---
1948	234,0	207,0	92,8	21,0	31,0	0,0	56,0	0,7	55,5	25,0	104,3	83,1
1949	64,2	127,9	147,2	24,2	46,7	51,5	0,0	11,0	0,0	67,6	86,0	250,0
1950	172,3	266,9	178,2	67,7	22,8	25,9	18,0	0,0	23,3	176,3	112,1	79,7
1951	440,9	302,0	378,7	12,2	39,4	73,9	---	42,0	20,0	119,3	261,5	81,2

Figura 9: Trecho do banco de dados.

Os dados do DAEE referem-se a um período que varia aproximadamente entre 1936 e 2004, coletados por 1116 estações distribuídas em todo o Estado de São Paulo. Cada estação de coleta possui uma localização diferente e um banco de dados referente à precipitação pluviométrica, com informações sobre: chuva diária, chuva mensal, máximas mensais, pluviosidade máxima diária, pluviosidade acumulada mensal e pluviosidade máxima média mensal.

Neste trabalho, inicialmente coletou-se um total de 109 dados das estações da região oeste do Estado de São Paulo. Como pode-se observar na Figura 9, o banco de dados do DAEE

apresenta todas as informações separadamente para cada estação, nos anos em que houve a coleta.

Os dados do DAEE foram reorganizados de modo que cada arquivo tinha as informações pluviométricas de todas as estações referentes a um período de tempo específico. Por exemplo, para a análise espacial do estudo geostatístico correspondente ao ano de 1995, necessita-se das informações espaciais referentes às estações neste ano, e se caso o estudo for temporal em uma localização específica, serão necessários os dados de todo o período referente a essa localização.

Geralmente, o arquivo de dados do DAEE contém campos com dados ausentes como ilustrado na Figura 9 e a quantidade de dados ausentes depende de cada ano de coleta, pois em um determinado ano contém somente um dado coletado ou em outro ano somente um dado ausente. O preenchimento desses dados foi feito através de uma média aritmética dos valores de precipitação pluviométrica de 3 ou 4 estações mais próxima a estação com essa falta de informação, ressaltando que fez-se esse preenchimento de acordo com o comportamento dos dados nessas estações. Essa etapa demandou tempo devido às quantidades de dados para organização e de campos sem informação.

O período coletado foi de 31 anos, de janeiro/1970 a dezembro/2000. A Figura 10 apresenta um trecho do banco de dados utilizado em estudo.

	A	B	C	D	E	F	G
1	CODEST	X_COORD	Y_COORD	JAN_90	FEV_90	MAR_90	ABR_90
2	c7-001	506884.00	7572674.00	301.60	34.60	143.00	153.10
3	c7-003	555413.00	7672196.00	262.30	53.30	258.30	147.50
4	c7-004	556787.00	7570729.00	213.40	6.60	109.90	91.00
5	c7-006	591343.00	7594552.00	263.50	68.50	164.80	85.00
6	c7-008	565678.00	7640798.00	176.90	85.10	187.30	74.50
7	c7-009	557087.00	7655588.00	204.70	51.60	261.00	97.90
8	c7-010	596724.00	7629574.00	188.20	138.10	193.10	113.60
9	c7-011	586272.00	7613029.00	271.80	74.90	146.80	89.70
10	c7-012	531127.00	7650124.00	158.70	96.80	214.10	105.20
11	c7-016	569024.00	7614957.00	301.60	34.60	143.00	153.10
12	c7-023	579595.00	7659184.00	191.50	70.80	223.20	121.00
13	c7-024	574311.00	7638916.00	196.40	46.30	189.40	105.60
14	c7-028	569158.00	7646318.00	186.65	38.10	221.10	87.40
15	c7-033	538040.00	7648264.00	198.50	84.10	234.50	146.20
16	c7-034	513834.00	7650147.00	301.10	67.80	190.70	112.90

Figura 10: Banco de dados referente à região oeste do Estado de São Paulo.

O banco de dados da região oeste do Estado de São Paulo contém os campos: código da estação, coordenadas (UTM), 360 colunas correspondentes aos meses do período de 1970 a 2000, totais anuais, máximo, mínimo e médio (mensal, anual e correspondente a cada estação).

A distribuição espacial das estações de coleta referente à precipitação pluviométrica está representada na Figura 11.

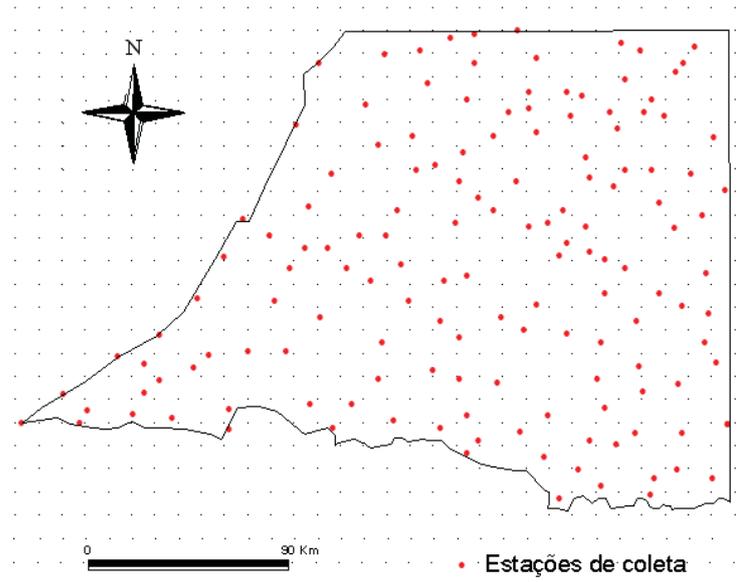


Figura 11: Distribuição espacial das estações de coleta do oeste paulista.

### 3.2. Método proposto

Para alcançar objetivo principal deste trabalho que é estudar o regime pluviométrico no Estado de São Paulo por meio de modelo que melhor represente o fenômeno no espaço, no tempo e no espaço-tempo, foram aplicadas técnicas estatísticas derivadas da Geoestatística: Geoestatística Clássica e Modelagem Espaço-temporal com os modelos produto-soma generalizado proposto por De Iaco, Myers e Posa (2002) e geoestatístico espaço-tempo proposto por Høst, Omre and Switzer (1995). O modelo produto-soma generalizado é um modelo separável, pois trata de maneira independente a estrutura espacial e temporal e foi utilizado por Lima (2006), o modelo geoestatístico espaço-tempo já foi utilizado por Kyriakidis e Journel (2009) e Ribeiro e Mingoti (2007). Optou-se por essas técnicas pelo fato de conter uma grande quantidade de dados no espaço e no tempo e principalmente por descrever continuamente o fenômeno através da modelagem.

Essas técnicas se diferenciam de outros métodos de interpolação pela maneira como os pesos são atribuídos a diferentes elementos amostrais. Inicialmente, o método de interpolação por krigagem foi desenvolvido para solucionar problemas de mapeamentos geológicos, mas seu uso

expandiu-se com sucesso no mapeamento de solos, mapeamento hidrológico, mapeamento atmosférico e outros campos correlatos. A Figura 12 representa o fluxograma que sintetiza a metodologia proposta.

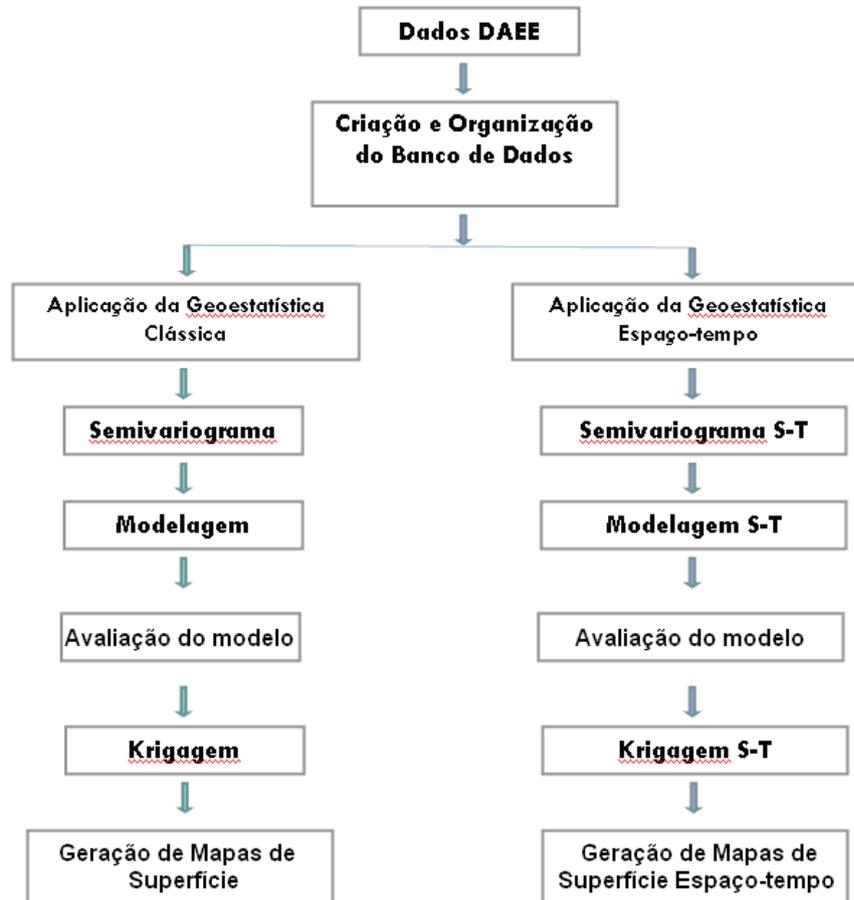


Figura 12: Fluxograma da metodologia proposta.

### 3.3. Programas computacionais utilizados

Na modelagem no espaço e no tempo utilizou-se SIG/SPRING e o software geoestatísticos VARIOWIN. Na modelagem espaço-temporal, os programas R/GeoR e MATLAB. Uma descrição sucinta desses programas será apresentada a seguir.

O SIG/SPRING é um Sistema de Informação Geográfica (SIG) com funções de processamento de imagens, análise espacial, modelagem numérica de terreno e consulta a bancos de dados espaciais, gratuito e disponibilizado em <[www.dpi.inpe.br/spring](http://www.dpi.inpe.br/spring)>. Tem como objetivo a análise, em duas dimensões, de dados espacialmente distribuídos. A estrutura utiliza a biblioteca pública de funções geoestatísticas GSLIB (DEUTSCH & JOURNAL, 1998).

O VARIOWIN é composto por três programas (arquivos \*.exe) executados isoladamente. São eles: PREVAR2D, utilitário que constrói uma matriz de distâncias para todos os possíveis pares de dados existentes no arquivo \*.dat; VARIO2D, utilitário que usa o arquivo \*.pcf, originado pelo PREVAR2D, para fazer a variografia exploratória em 2D; MODEL, utilitário que executa, de modo interativo, o ajuste, a um modelo teórico, do semivariograma experimental obtido pelo VARIO2D e faz a modelagem interativa para anisotropia geométrica e zonal. É também um software gratuito disponibilizado em <[www-sst.unil.ch/research/variowin/](http://www-sst.unil.ch/research/variowin/)>.

R é uma linguagem de programação muito difundido no meio estatístico por realizar várias análises estatísticas. Ele tem um módulo exclusivo denominado geoR para o estudo geoestatístico e foi usado para modelagem espaço-tempo da precipitação pluviométrica. O software de domínio público pode ser encontrado em <[www.r-project.org](http://www.r-project.org)>.

### **3.4. Resultados**

#### **3.4.1. Análise exploratória**

Essa etapa tem como objetivo reunir em grupos homogêneos as estações de coleta do oeste paulista através da característica comum entre elas que são as medidas de precipitação pluviométrica.

O método ou algoritmo de agrupamento ou cluster aplicado aos dados de precipitação pluviométrica foi um método hierárquico que utiliza a distância euclidiana para medir o grau de similaridade ou dissimilaridade entre os indivíduos. Com o objetivo de formar grupos mais heterogêneos possíveis e de modo que os elementos dentro de cada grupo fossem homogêneos, adotou-se o procedimento proposto por Ward, fundamentado na “mudança de variação” entre os grupos e dentro dos grupos que estão sendo formados em cada passo de agrupamento. De acordo com Johnson e Wichern (2007), esse procedimento também é denominado de “mínima variância”, pois em cada passo do algoritmo, os dois conglomerados que apresentam a menor distância entre elas são combinados formando um único grupo.

Inicialmente, a análise de agrupamento foi realizada considerando-se como observações todas as estações de coleta e em seguida somente os anos de coleta, sendo as variáveis a precipitação mensal para as estações pluviométricas e a precipitação anual para os anos em estudo. O programa computacional utilizado para aplicar esta análise foi o MINITAB, software estatístico desenvolvido nos Estados Unidos pelo State College, Pensilvânia, cuja licença é paga.

- classificação das estações pluviométricas:

Foram consideradas 108 estações de coleta no período de 31 anos separados mensalmente (janeiro/1970 a dezembro de 2000).

Como não havia a priori, número de agrupamentos (clusters) desejado, a escolha deste número baseou-se na comparação gráfica dos números de grupos com os coeficientes de fusão, isto é, o valor numérico (distâncias ou semelhança) para os quais vários elementos se unem para formar um grupo. Conforme a Figura 13, pode-se observar que a partir do grupo 7 não há saltos significativos nos valores das distâncias entre grupos (coeficientes de fusão), embora a partir do terceiro grupo as diferenças entre os grupos também não sejam tão grandes.

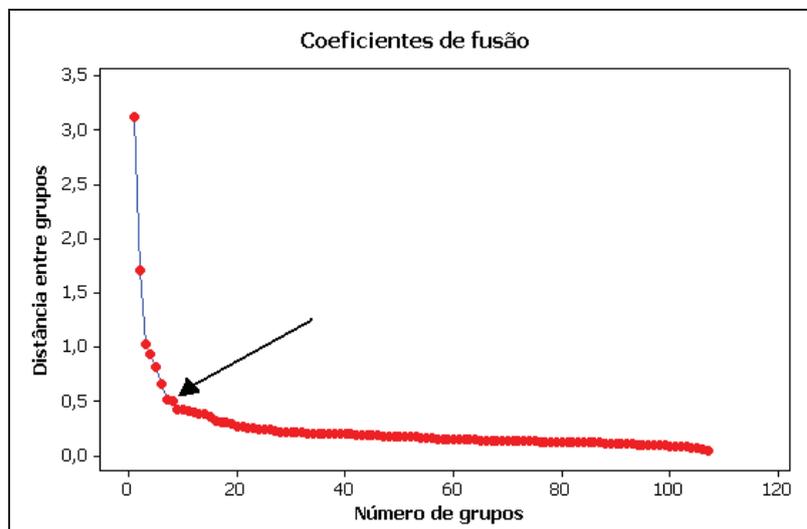


Figura 13: Gráfico dos coeficientes de fusão para as localidades.

Após definir o número de grupos adequado, o próximo passo foi a geração do dendrograma com a divisão de 7 grupos que está ilustrado na Figura 14.

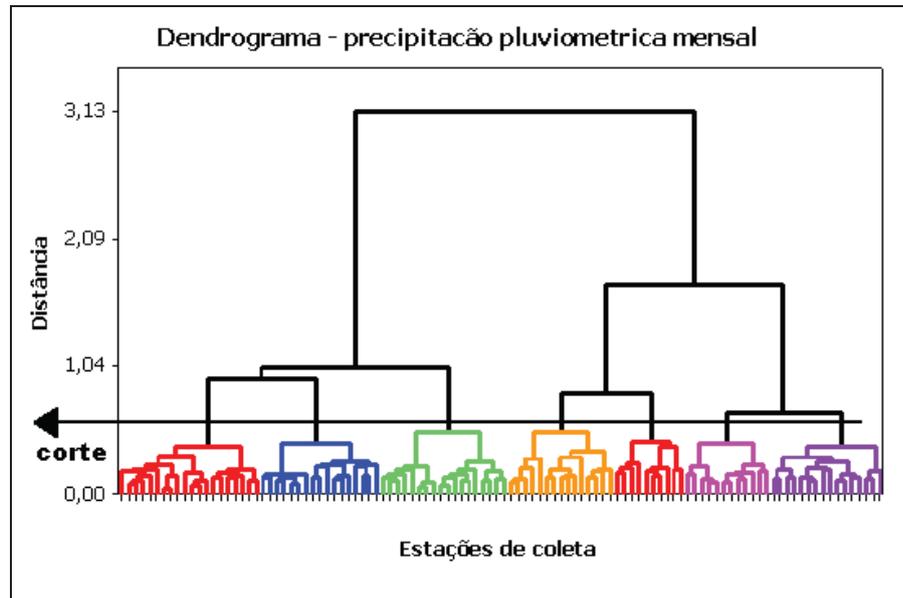


Figura 14: Dendrograma para as localidades.

Para facilitar a visualização e interpretação, foram gerados os dendrogramas individuais para cada grupo homogêneo que estão ilustrados na Figura 16.

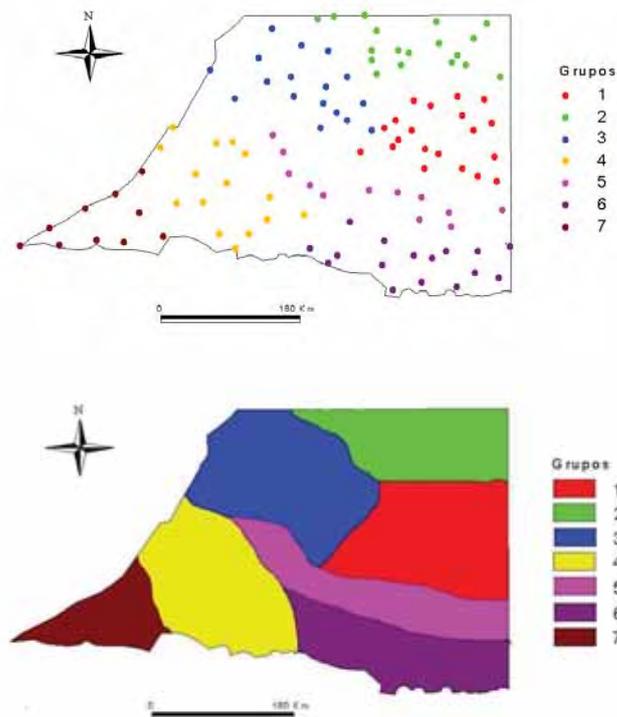


Figura 15: Distribuição espacial das estações pluviométricas separadas por grupos homogêneos

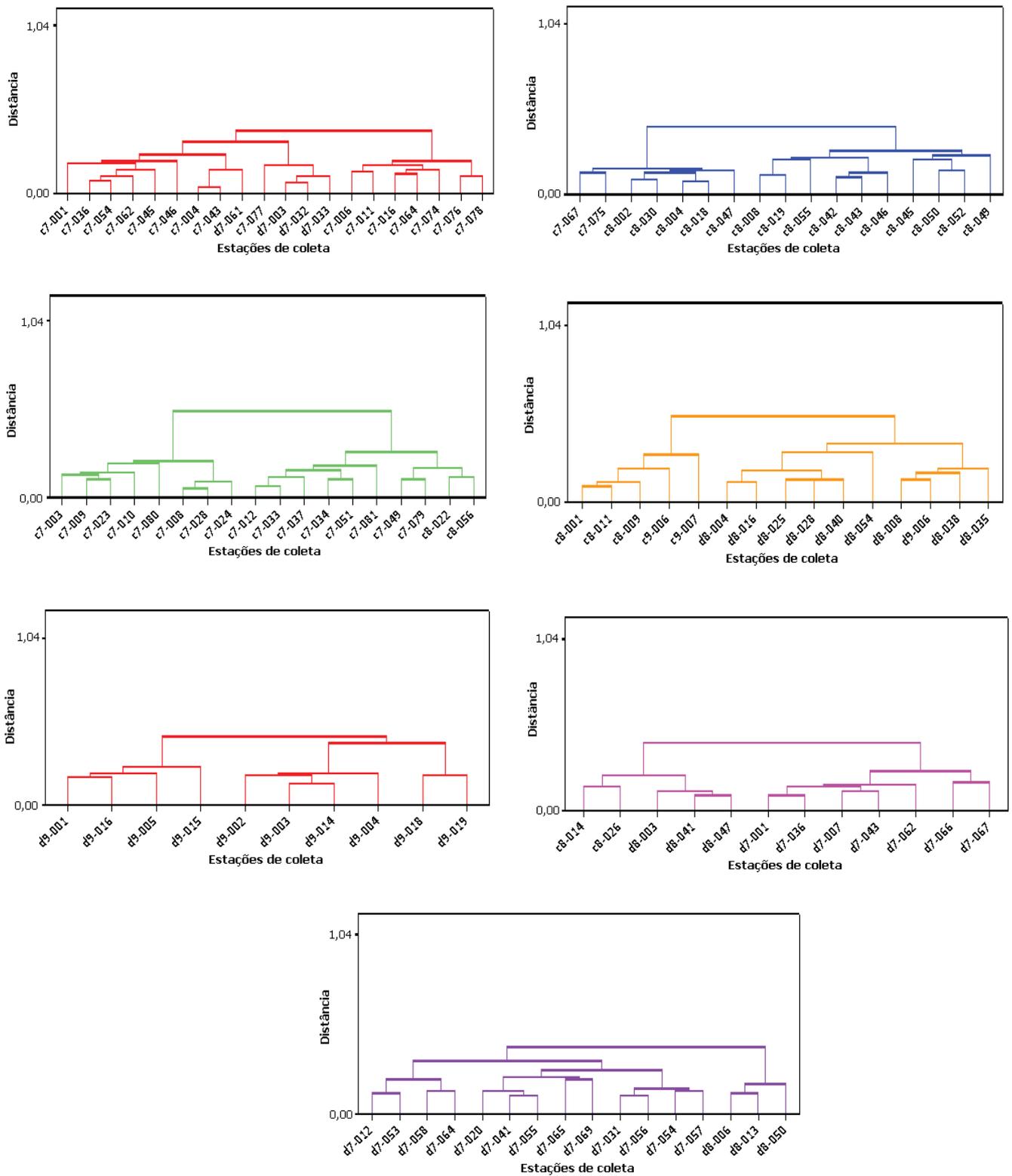


Figura 16: Dendrogramas individuais dos 7 grupos homogêneos para as localidades.

Conforme as Figuras 1 e Figura 15, a região administrativa de Presidente Prudente é formada pelos grupos 4 e 7, aproximadamente 82 % do grupo 3 e 25% do grupo 1. A região de Marília contem 70% das estações do grupo 1, 50% do grupo 5 e 75% do grupo 6. Já a região de Araçatuba contém as estações do grupo 2 e outras em quantidades bem menores (25% do grupo 1 e aproximadamente 18% do grupo 3). Em Bauru há somente a presença de 5% do Grupo 1.

Esses agrupamentos de estações foram colocados no mapa, para verificar se estão distribuídos aleatoriamente na região de estudo ou se houve agrupamento por vizinhança conforme a Figura 15.

Neste caso, foi considerada somente a variável precipitação mensal, os autores Sant'Anna Neto (1995) e Boin (2000) consideraram outras variáveis altamente correlacionadas como a altitude, clima e correntes atmosféricas para chegarem a conclusões semelhantes a respeito da classificação da precipitação no oeste paulista. Sant'Anna Neto (1995) classificou a chuva em 6 regiões homogêneas e Boin (2000) em 9 regiões homogêneas conforme a Figura 17.

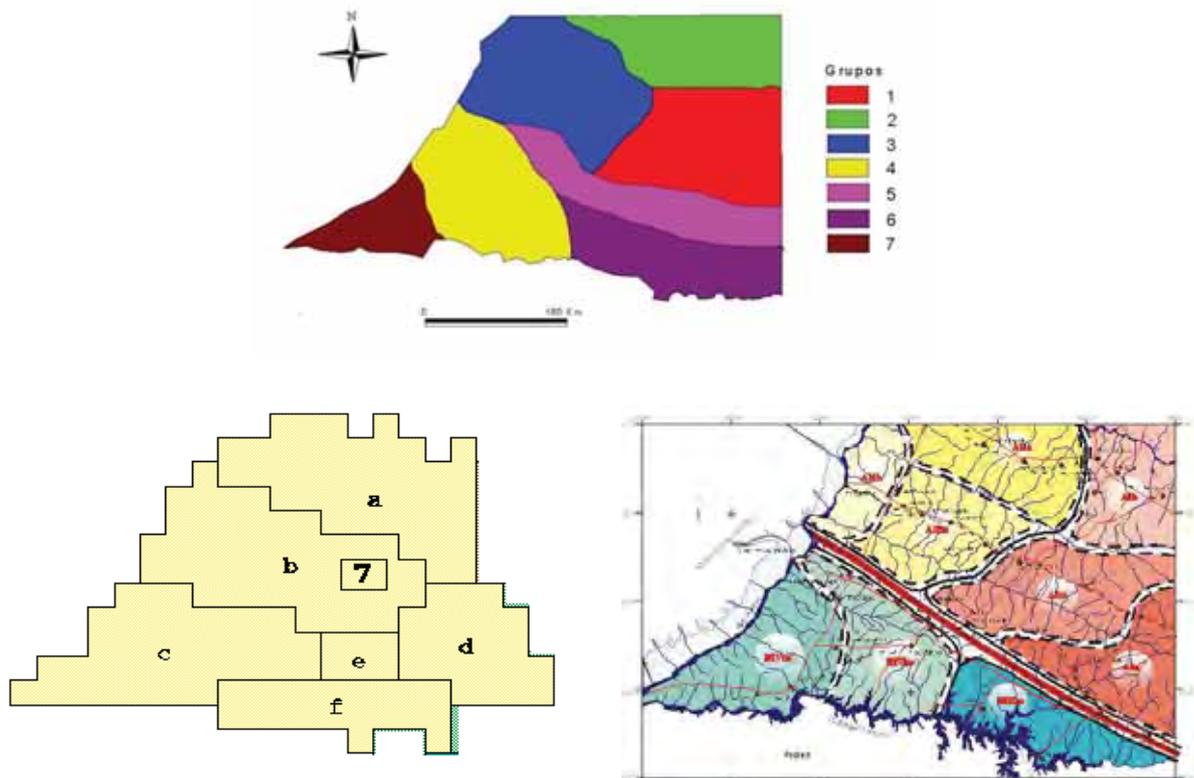


Figura 17: Comparação dos mapas de chuva gerados por diferentes autores: Jacomo (2011) – acima, Sant'Anna Neto (1995) – esquerda e Boin (2000) – direita.

As comparações entre os grupos foram feitas através das médias totais anuais para todas as estações referentes a cada grupo para os 31 anos. Os gráficos das médias anuais dos 7 grupos estão ilustrados na Figuras 18.

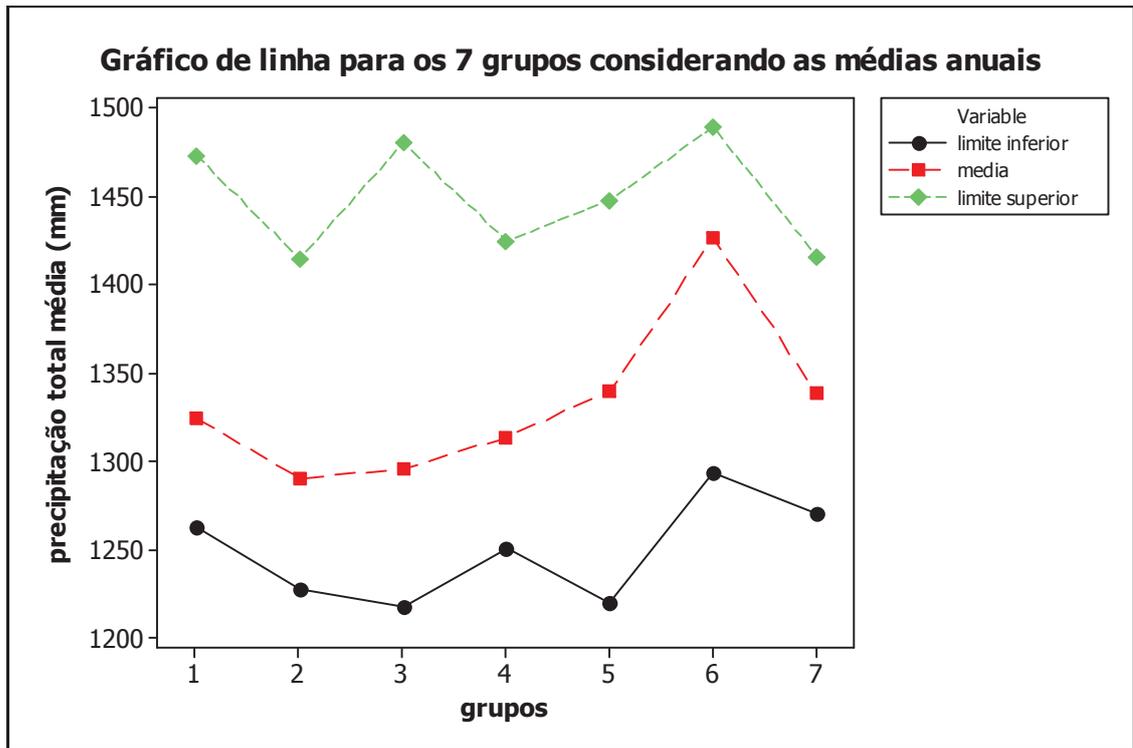


Figura: 18: Gráfico de linha para as médias totais anuais.

De acordo com os gráficos apresentados na Figura 18, o grupo 6 apresenta os maiores valores médios anuais de precipitação pluviométrica e os grupos 2 e 3 apresentam os menores valores de precipitação. As estatísticas descritivas estão representadas na Tabela 1.

O grupo 6 representa a região chuvosa e está localizado 75% na região de Marília e 25% na região de Presidente Prudente, com precipitação anual variando em torno de 1400 mm.

Os grupos 2 e 3 representam a região seca e estão localizados ao norte da região administrativa de Presidente Prudente e na região de Araçatuba, com precipitação anual variando em torno de 1300 mm.

Tabela 1: Estatísticas descritivas referentes aos 7 grupos homogêneos.

	<b>Total Mínimo</b>	<b>Total Máximo</b>	<b>Média Anual Limite Inferior</b>	<b>Média Anual</b>	<b>Média Anual Limite Superior</b>
<b>Grupo 1</b>	39136,46	45660,3	1262,47	1324,56	1472,91
<b>Grupo 2</b>	38053,2	43858,7	1227,52	1290,12	1414,80
<b>Grupo 3</b>	37739,9	45878,4	1217,42	1295,95	1479,95
<b>Grupo 4</b>	38780,79	44166,97	1250,99	1313,49	1424,74
<b>Grupo 5</b>	37822,15	44874,64	1220,07	1339,83	1447,57
<b>Grupo 6</b>	40087,35	46167,75	1293,14	1426,52	1489,28
<b>Grupo 7</b>	39392,11	43886,4	1270,71	1338,24	1415,69

- classificação dos anos:

Foram consideradas 108 estações de coleta no período de 31 anos separados anualmente (1970 a 2000).

Como não havia a priori, número de agrupamentos (clusters) desejado, a escolha deste número baseou-se na comparação gráfica dos números de grupos com os coeficientes de fusão sendo que foi escolhido um total de 3 grupos. (Figura 19)

Após definir o número de grupos adequado, o passo seguinte foi a geração do dendrograma com a divisão de 3 grupos que está ilustrado na Figura 20 e os dendrogramas de cada grupo na Figura 21.

As comparações entre os grupos foram feitas através das médias aritméticas anuais de precipitações pluviométricas de todas as estações de cada grupo. Por exemplo, o grupo foi formado por 12 anos semelhantes e o grupo 2 por 11 anos e foram geradas médias aritméticas de todos os 12 anos do grupo 1 para cada estação de coleta para poderem ser comparadas com as médias geradas nos 11 anos semelhantes do grupo 2 em todas as 108 estações. Os histogramas e os box-plots dos 3 grupos estão ilustrados nas Figuras 22 e 23, respectivamente.

Observa-se que os menores valores de precipitação pluviométrica estão concentrados no grupo 1 e os maiores no grupo 3, que receberam denominação de anos mais secos e mais chuvosos, respectivamente.

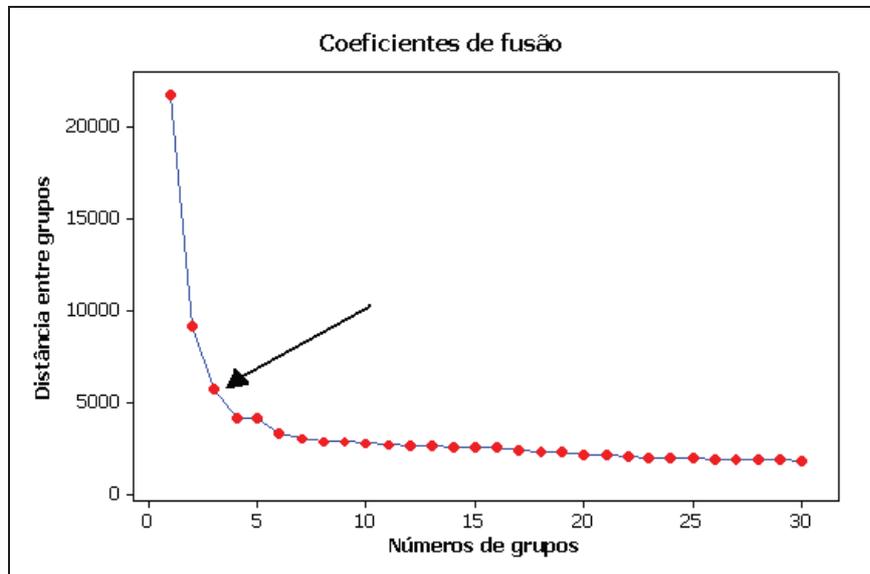


Figura 19: Gráfico dos coeficientes de fusão.

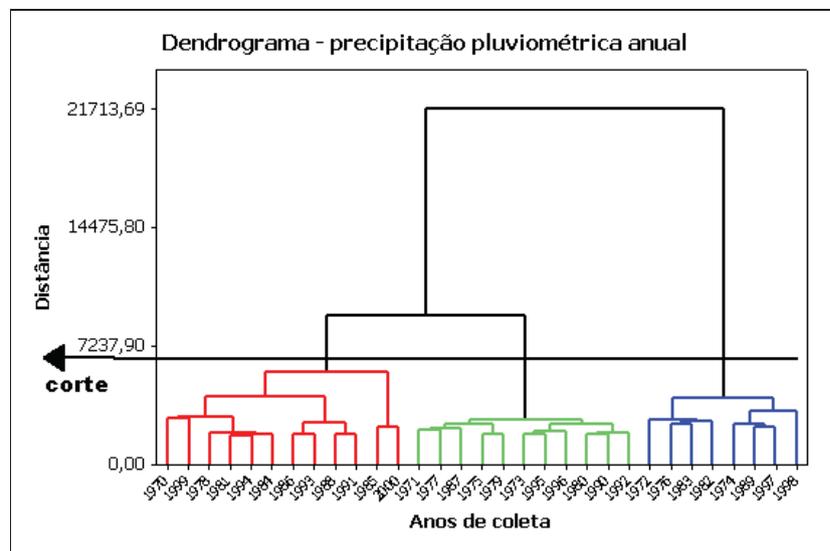


Figura 20: Dendrograma para os anos.

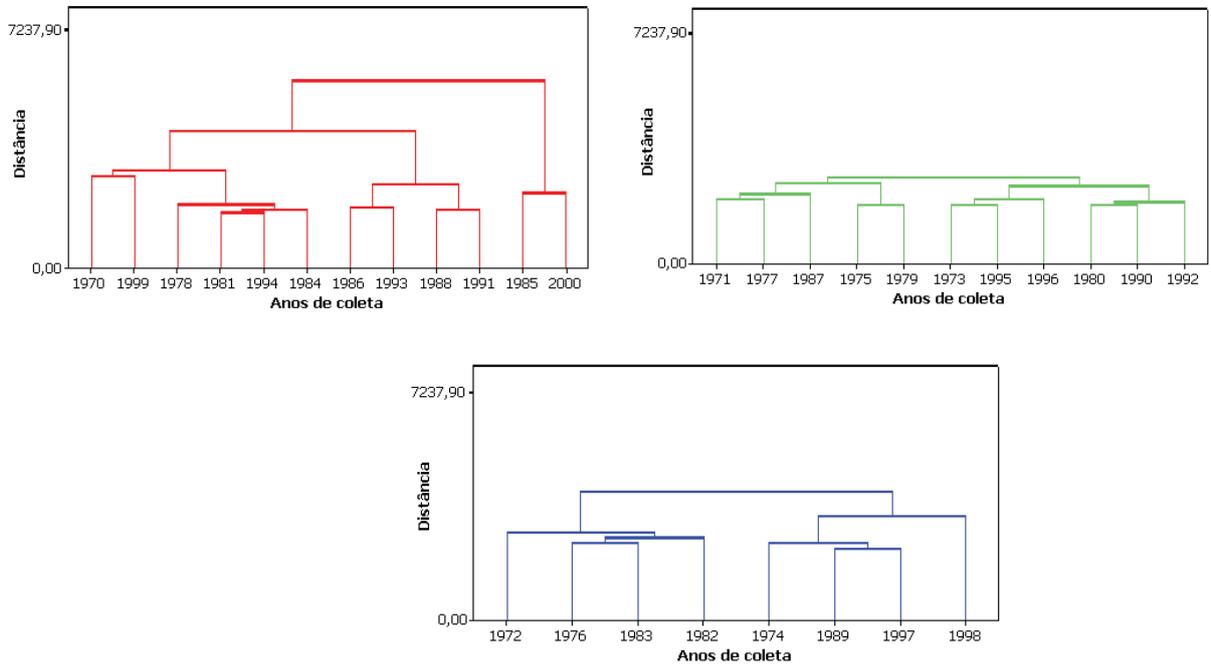


Figura 21: Dendrogramas individuais dos 3 grupos homogêneos

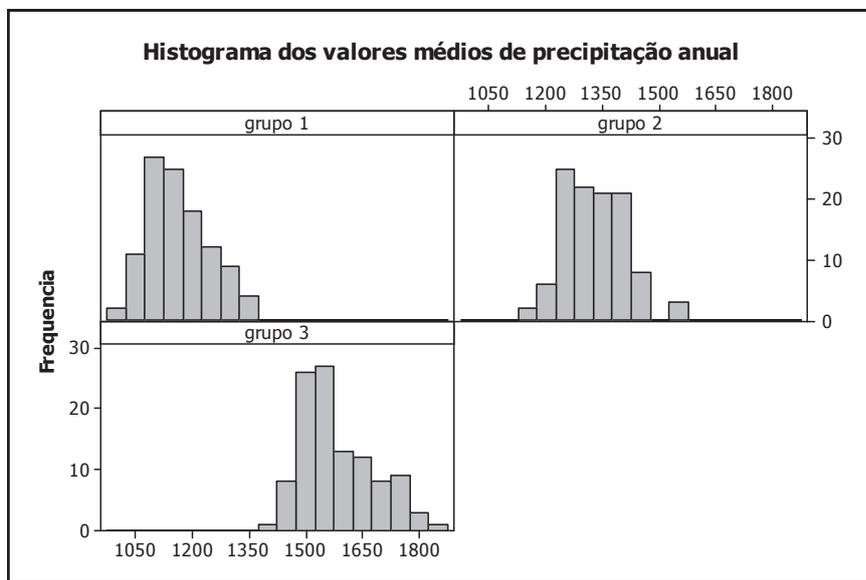


Figura 22: Histograma da distribuição das precipitações médias anuais para os 3 grupos de anos semelhantes.

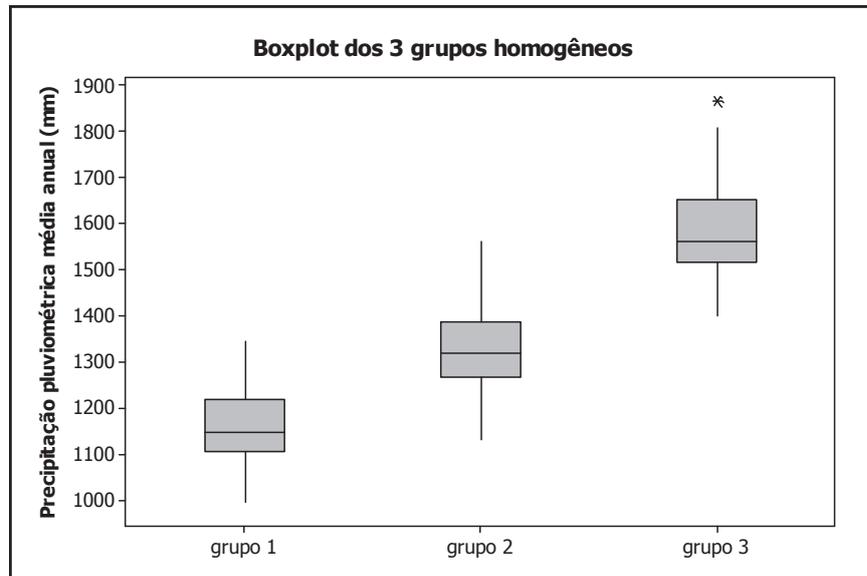


Figura 23: Box-plot das precipitações médias anuais para os 3 grupos de anos semelhantes.

Os anos mais secos são os correspondentes aos de 1970, 1978, 1981, 1984, 1985, 1986, 1988, 1991, 1993, 1994, 1999 e 2000 classificados no grupo 1. Os mais chuvosos são os anos de 1972, 1974, 1976, 1982, 1983, 1989, 1997 e 1998 classificados no grupo 3. Os demais anos foram denotados habituais e são do grupo 2. As estatísticas descritivas estão apresentadas na Tabela 2 e o gráfico de pontos com as respectivas classificações na Figura 24.

Tabela 2: Estatísticas descritivas referentes aos 3 grupos homogêneos.

	<b>N</b>	<b>Média</b>	<b>Variância</b>	<b>D.Padrão</b>	<b>Mínimo</b>	<b>Mediana</b>	<b>Máximo</b>
<b>Grupo 1</b>	108	1163,3	6751,4	82,2	998,1	1149,2	1346,5
<b>Grupo 2</b>	108	1328,1	6500,1	80,6	1134,5	1320,4	1563,1
<b>Grupo 3</b>	108	1585,1	9161,3	95,7	1400,6	1563,6	1865,9

A ocorrência do período seco mais extenso foi nos anos de 1984, 1985 e 1986, o período chuvoso mais extenso foi no máximo de 2 anos como apresenta os anos de 1979-1980, 1995-1996. Os períodos mais semelhantes foram o período de 1982 a 1985 e o período de 1997 a 2000.

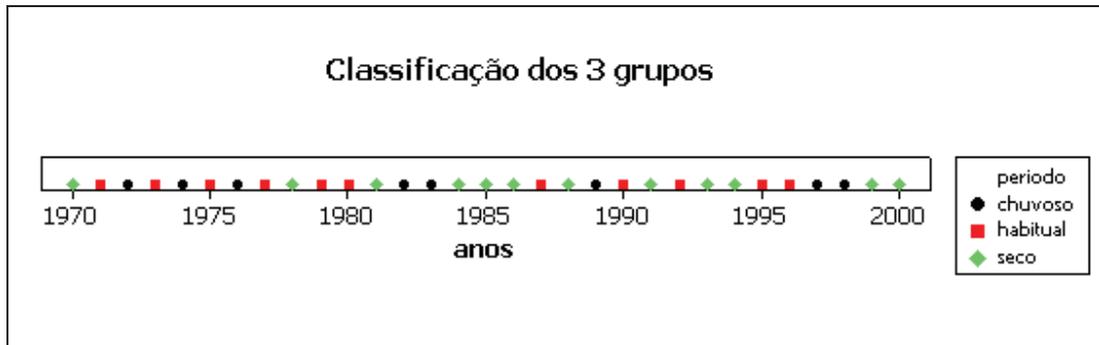


Figura 24: Gráfico de pontos para os 3 grupos.

### 3.4.2. Modelo produto-soma generalizado

A geração desse modelo espaço-tempo consiste na soma e produto dos modelos espacial e temporal que são derivados da geoestatística.

#### a) geração do modelo espacial

As variáveis selecionadas foram: precipitação máxima mensal e precipitação média mensal. A precipitação máxima mensal corresponde aos valores máximos de chuva no período de 11 anos (jan/1990 a dez/2000) e a média mensal aos valores médios (média aritmética) desse período em cada estação. Inicialmente foi gerada a superfície variográfica para cada variável de interesse, com intuito de descrever o comportamento espacial da variável em estudo. (Figura 25)

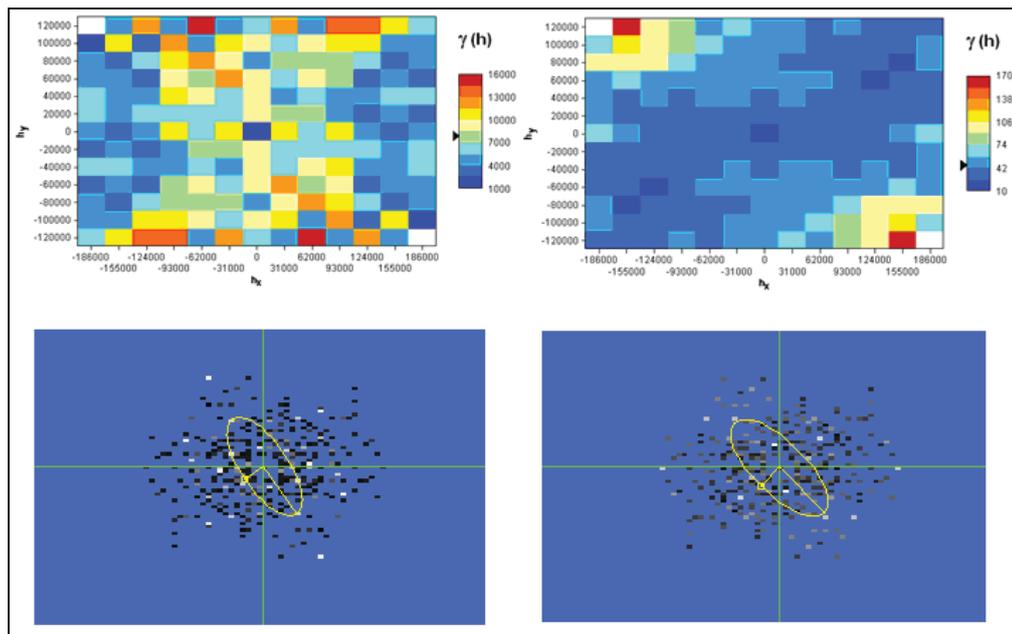


Figura 25: Superfície variográfica das precipitações máxima e média mensal respectivamente.

Nesse caso todas as variáveis em estudo apresentam comportamento anisotrópico, pois possuem uma direção com maior variabilidade espacial. No caso da precipitação máxima mensal, as direções de maior e menor variabilidade foram de  $120^\circ$  e  $30^\circ$ , respectivamente. Já na precipitação média mensal as direções de maior e menor variabilidade foram de  $130^\circ$  e  $40^\circ$ , respectivamente.

A seguir foram gerados os modelos para as variáveis em estudo considerando os casos isotrópico e anisotrópico.

- precipitação máxima mensal: *caso isotrópico*

Neste caso, foi ajustado um modelo esférico a partir do semivariograma omnidirecional para a variável precipitação máxima mensal com os parâmetros: Efeito Pepita ( $C_0$ ) = 2409,33, alcance ( $a$ ) = 92 km e contribuição ( $C_1$ ) = 5658,6. Nesse caso a precipitação máxima mensal no oeste do Estado de São Paulo possui uma dependência espacial de 92 km. (Figura 26)

O modelo matemático ajustado para a variável precipitação máxima mensal a partir da Equação (4) é:

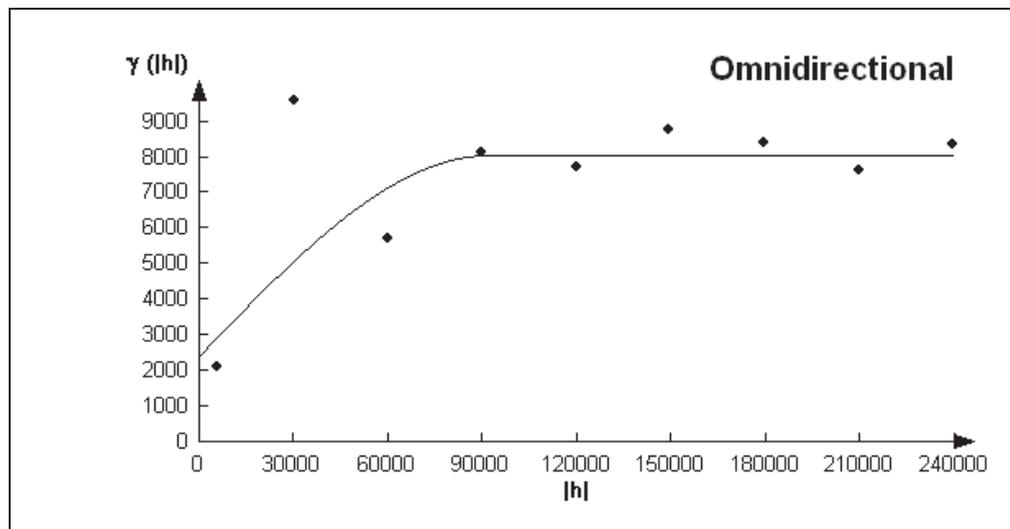


Figura 26: Modelo esférico ajustado para a variável precipitação máxima mensal

$$\gamma(h) = \begin{cases} 2409,33, & |h|=0 \\ 2409,33 + 5658,6 \cdot \left[ \frac{3}{2} \left( \frac{|h|}{92000} \right) - \frac{1}{2} \left( \frac{|h|}{92000} \right)^3 \right] = 2409,33 + 5658,6 [Sph(|h|)], & 0 < |h| \leq 92000 \\ 8067,93, & |h| > 92000 \end{cases} \quad (60)$$

- precipitação máxima mensal: *caso anisotrópico*

As direções de maior e menor variabilidade espacial são 120° e 30°, respectivamente. Essas direções não apresentaram dependência espacial (ausência de alcance) e foi ajustado o modelo efeito pepita puro com os parâmetros 8500 para a direção de 120° e 7520 para a direção 30°.

Neste caso, a variável precipitação máxima mensal pode ser aplicada outro tipo de procedimento de interpolação, devido a não existência da dependência espacial, como por exemplo, o Inverso da Distância.

Os casos de efeito pepita puro ocorreram devido à suavização dos dados, pois foram considerados somente os pontos máximos das estações do período de 11 anos. Os mesmos dados trabalhados para os anos chuvoso e seco mostraram que não acontecia os casos de efeito pepita puro conforme Flores (2000). O autor separou os dados de precipitação pluviométrica do oeste paulista em estações do ano (primavera, verão, outono e inverno) e “anos-padrão” aplicando a geoestatística para todos esses casos.

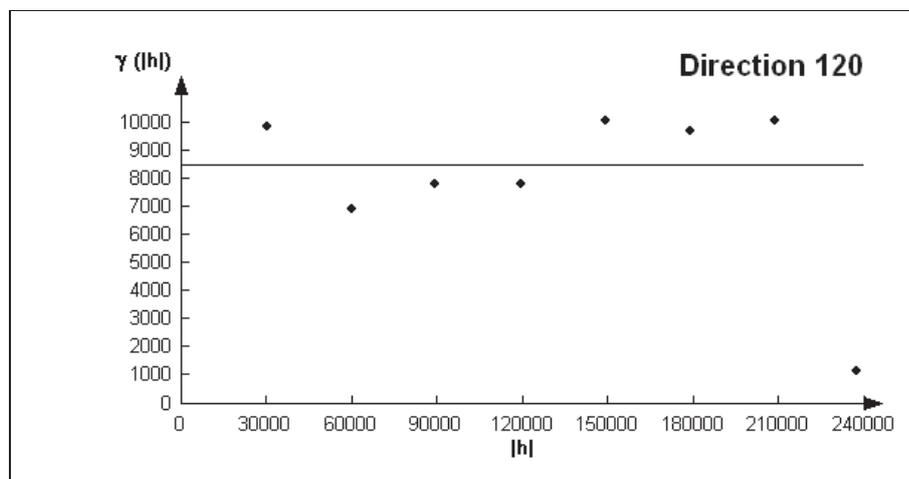


Figura 27: Modelo efeito pepita puro ajustado para a variável precipitação máxima mensal na direção de 120°

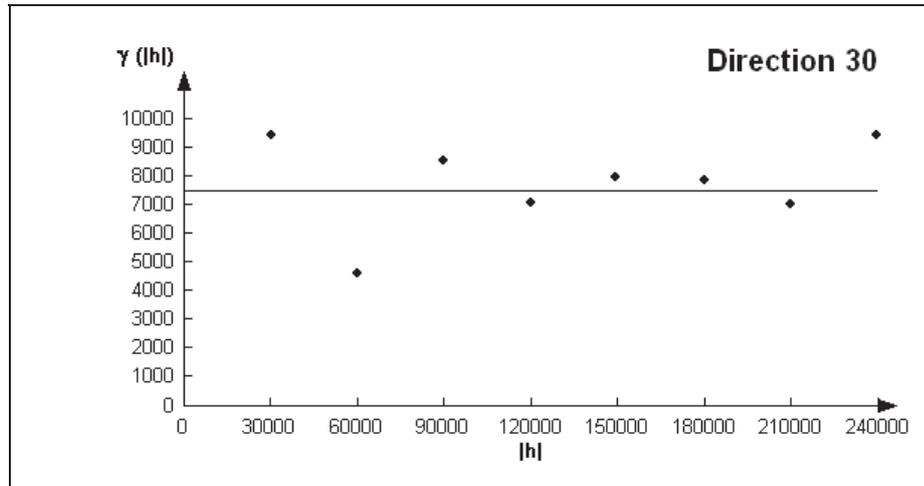


Figura 28: Modelo efeito pepita puro ajustado para a variável precipitação máxima mensal na direção de 30°

- precipitação média mensal: caso isotrópico

Neste caso, foi ajustado um modelo esférico a partir do semivariograma omnidirecional para a variável precipitação média mensal com os parâmetros: Efeito Pepita ( $C_0$ ) = 16,48, alcance ( $a$ ) = 123 km e contribuição ( $C_1$ ) = 34. Nesse caso a precipitação média mensal no oeste do Estado de São Paulo possui uma dependência espacial de 123 km. (Figura 29)

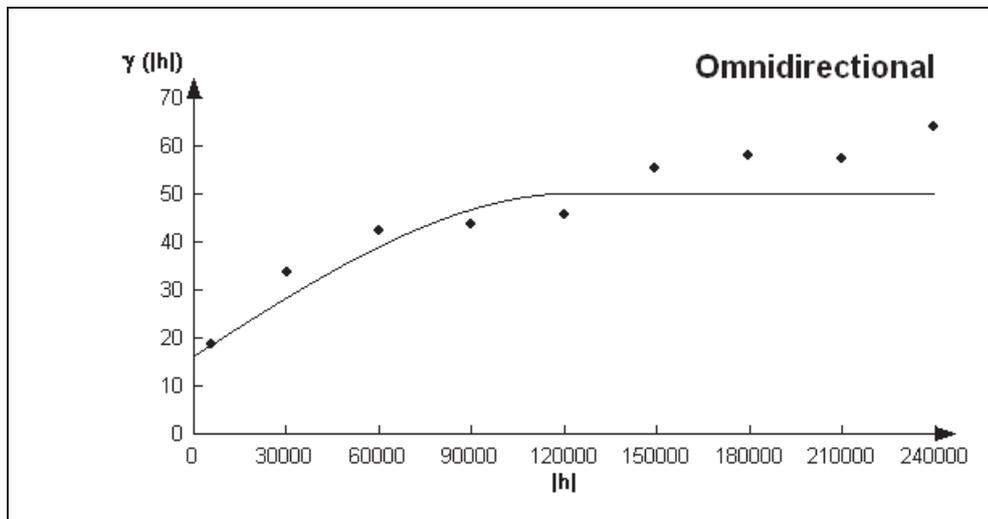


Figura 29: Modelo esférico ajustado para a variável precipitação média mensal

O modelo matemático ajustado para a variável precipitação média mensal a partir da Equação (4) é:

$$\gamma(h) = \begin{cases} 16,48, & |h|=0 \\ 16,48 + 34 \cdot \left[ \frac{3}{2} \left( \frac{|h|}{123000} \right) - \frac{1}{2} \left( \frac{|h|}{123000} \right)^3 \right] = 16,48 + 34 \cdot [Sph(|h|)], & 0 < |h| \leq 123000 \\ 50,48, & |h| > 123000 \end{cases} \quad (61)$$

- precipitação média mensal: caso anisotrópico

As direções de maior e menor variabilidade espacial são 130° e 40°, respectivamente. A direção de 130° apresentou tendência (ausência de patamar ou não-estacionaridade) e foi ajustado o modelo potência. Já direção de 40° apresentou efeito pepita puro com o parâmetro de 38,5. Estudos mais detalhados podem ocorrer na variável precipitação média mensal como a remoção da tendência da direção de 130° e aplicação da Krigagem Universal.

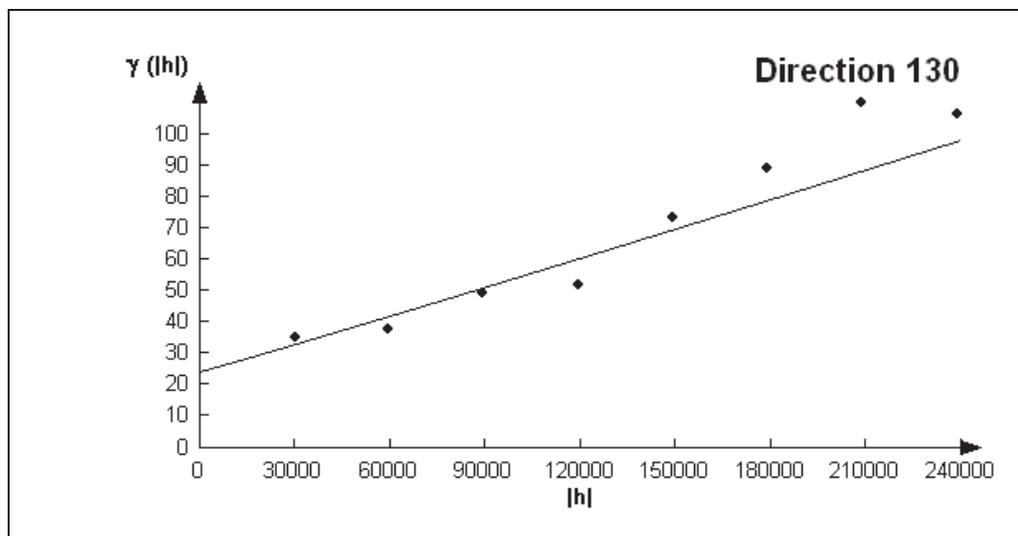


Figura 30: Modelo potência ajustado para a variável precipitação média mensal na direção de 130°.

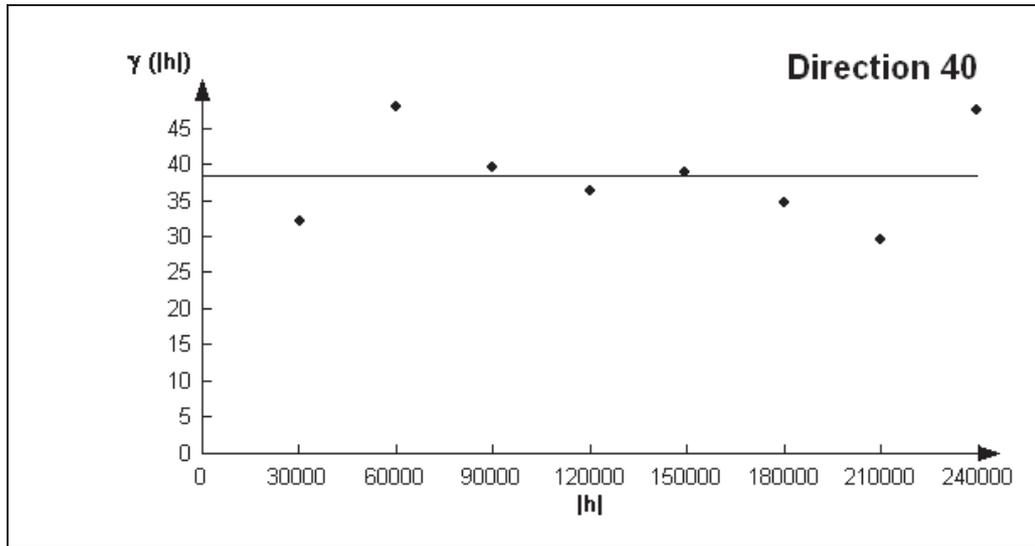


Figura 31: Modelo esférico ajustado para a variável precipitação média mensal na direção de 40°.

### b) geração do modelo temporal

As variáveis temporais selecionadas foram: precipitação máxima e precipitação média. A variável precipitação máxima é a quantidade máxima de chuva em um determinado tempo, por exemplo, a precipitação máxima no mês de janeiro de 1990 foi de 546 mm. Já a variável precipitação média é a quantidade média de chuva em um determinado tempo.

Todas as estações de coleta de dados do oeste do Estado de São Paulo (111 estações) estão organizadas em 132 meses e o período de estudo foi correspondente a 11 anos.

#### - precipitação máxima:

Para a geração do semivariograma temporal foi considerado o espaçamento de *lag* de 2 meses, tolerância do lag de 0,01 meses e número de lags igual a 20 meses.

A função gerada para os dados do semivariograma temporal da variável precipitação máxima, apresentada na Figura 32, foi à fusão de uma função periódica com a função exponencial apresentada na Figura 33.

$$\gamma(t) = 31000 \cdot (1 - \exp(-t/1,14)) - 10770,5 \cdot (1 + \cos(2 \cdot \pi \cdot t / -12,192)) \quad (62)$$

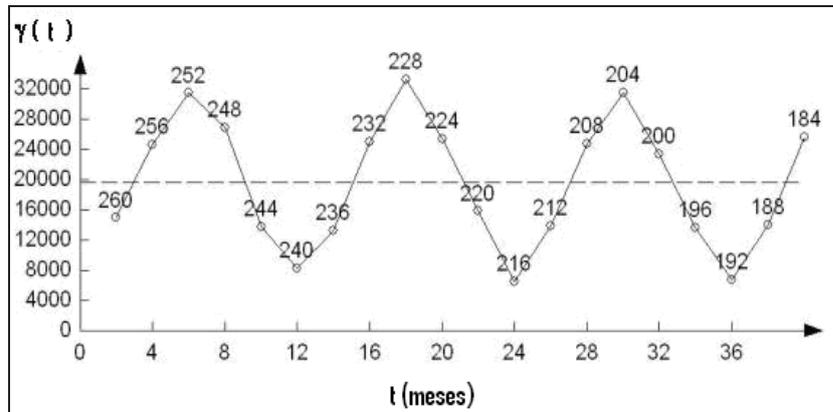


Figura 32: Semivariograma temporal para a variável precipitação máxima.

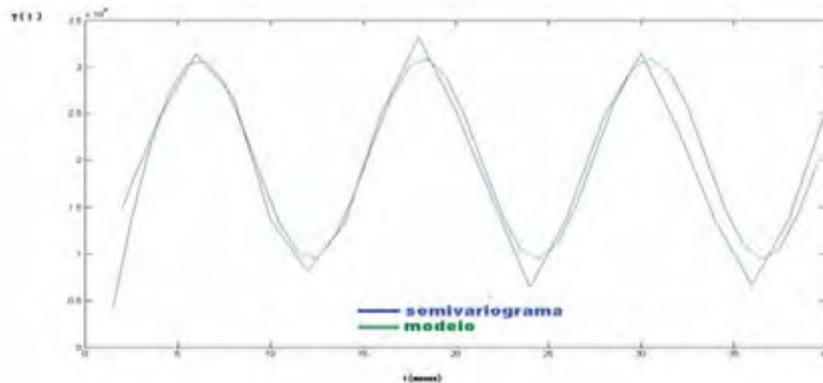


Figura 33: Função ajustada para a variável precipitação máxima.

- precipitação média:

Para a geração do semivariograma temporal foi considerado o espaçamento de *lag* de 2 meses, tolerância do lag de 0,01 meses e número de lags igual a 20 meses.

A função gerada para os dados do semivariograma temporal da variável precipitação média, apresentada na Figura 34, foi à fusão de uma função periódica com a função exponencial apresentada na Figura 35.

$$\gamma(t) = 9134 \cdot (1 - \exp(-t/1,219)) - 2703,9915 \cdot (1 + \cos(2,5 \cdot \pi \cdot t - 15,27)) \quad (63)$$

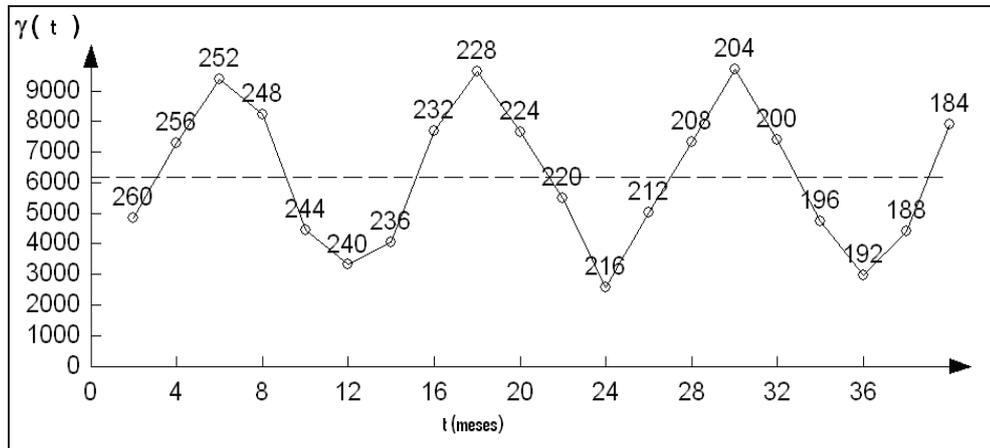


Figura 34: Semivariograma temporal para a variável precipitação média.

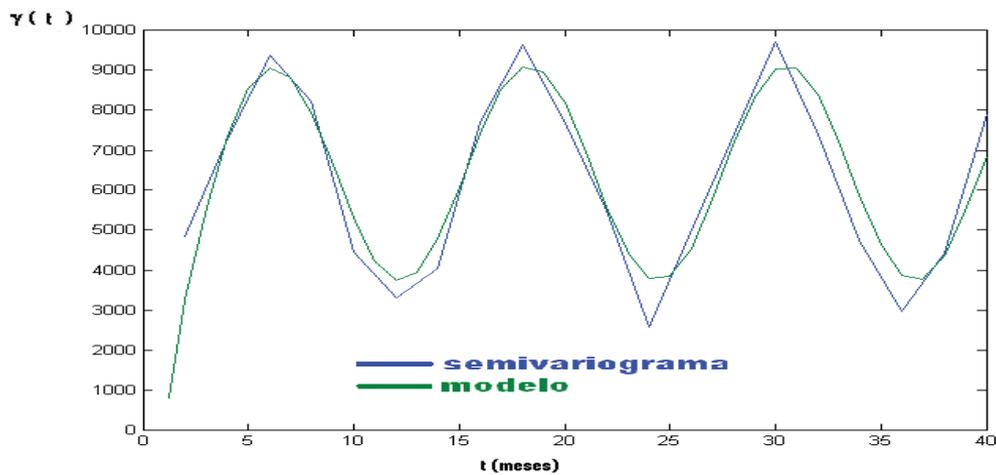


Figura 35: Função ajustada para a variável precipitação média.

### c) geração do modelo espaço-tempo

Na análise espaço-temporal, foi utilizado o modelo já discutido no Capítulo 2 da classe de modelos separáveis, denominado modelo produto-soma generalizado. Esse modelo consiste em associar as componentes espacial e temporal, para isto, é necessário determinar o valor de  $k$ , que “pondera” o produto dos modelos espacial e temporal.

Os modelos espaciais e temporais escolhidos são os correspondentes a variável precipitação mensal: média e máxima.

-precipitação máxima mensal no espaço e precipitação máxima no tempo:

No primeiro passo determinou-se os patamares espacial e temporal para atender o critério de valores admissíveis para  $k$ . Neste caso, os patamares são:  $C\gamma_{st}(h_s, 0) = 8067,93$  e  $C\gamma_{st}(0, h_t) = 31000$ :

$$0 < k \leq 1/\max\{8067,93; 31000\} \rightarrow 0 < k \leq 0,0001 \quad (64)$$

Através desse critério, pode-se obter o intervalo de valores admissíveis para o patamar global  $C\gamma_{st}(h_s, h_t)$ .

$$0 < \frac{(8067,93 + 31000 - C\gamma_{st}(h_s, h_t))}{(8067,93 \cdot 31000)} \leq 0,0001$$

$$14047,35 < C\gamma_{st}(h_s, h_t) \leq 39057,93 \quad (65)$$

O modelo produto-soma generalizado espaço-temporal da Equação (47) ajustado aos dados de precipitação máxima é dado por:

$$\begin{aligned} \gamma_{ST}(h_s, h_t) = & [2409,33 + 5658,6 \left[ \frac{3}{2} \left( \frac{|h|}{92000} \right) - \frac{1}{2} \left( \frac{|h|}{92000} \right)^3 \right] + [31000 \cdot (1 - \exp(-h_t/1,14)) - \\ & 10770,5 \cdot (1 + \cos(2 \cdot \pi \cdot h_t / -12,192))] - k \left\{ 2409,33 + 5658,6 \left[ \frac{3}{2} \left( \frac{|h|}{92000} \right) - \frac{1}{2} \left( \frac{|h|}{92000} \right)^3 \right] \right\} \cdot \\ & \cdot [31000 \cdot (1 - \exp(-h_t/1,14)) - 10770,5 \cdot (1 + \cos(2 \cdot \pi \cdot h_t / -12,192))] \end{aligned} \quad (66)$$

em que:  $0 < k \leq 0,0001$ ,  $h_s$  distância espacial em metros e  $h_t$  distância temporal em meses.

Neste caso, observa-se que o produto dos modelos espacial e temporal não tem muita significância, pois o valor  $k$  que pondera a soma e produto desses modelos está próximo de 0. O modelo produto-soma generalizado está ilustrado na Figura 36.

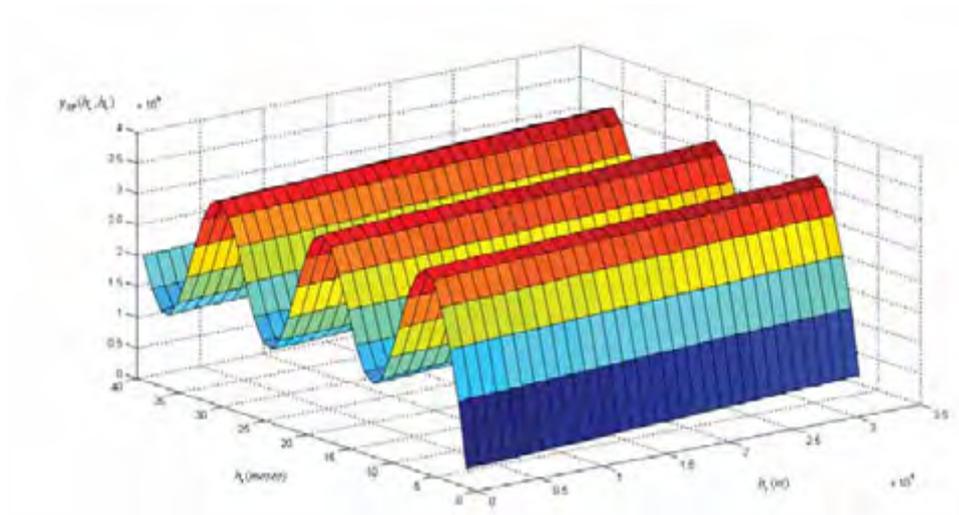


Figura 36: Superfície gerada através do modelo produto-soma generalizado da precipitação máxima mensal.

-precipitação média mensal no espaço e precipitação média no tempo:

O primeiro passo foi determinar os patamares espacial e temporal para atender o critério de valores admissíveis para  $k$ . Neste caso, os patamares são:  $C\gamma_{st}(h_s, 0) = 50,48$  e  $C\gamma_{st}(0, h_t) = 9000$  então:

$$0 < k \leq 1/\max\{50,48; 9000\} \rightarrow 0 < k \leq 1/9000 \rightarrow 0 < k \leq 0,0001 \quad (67)$$

Através desse critério, pode-se obter o intervalo de valores admissíveis para o patamar global  $C\gamma_{st}(h_s, h_t)$ .

$$0 < \frac{(50,48 + 9000 - C\gamma_{st}(h_s, h_t))}{(50,48 \cdot 9000)} \leq 0,0001$$

$$9005,49 < C\gamma_{st}(h_s, h_t) \leq 9050,48 \quad (68)$$

O modelo produto-soma generalizado espaço-temporal da Equação (47) ajustado aos dados de precipitação média é dado por:

$$\begin{aligned}
\gamma_{ST}(h_s, h_t) = & 16,48 + 34 \left[ \frac{3}{2} \left( \frac{|h_s|}{123000} \right) - \frac{1}{2} \left( \frac{|h_s|}{123000} \right)^3 \right] - [9134 \cdot (1 - \exp(-h_t/1,219)) - \\
& 2703,9915 \cdot (1 + \cos(2,5 \cdot \pi \cdot h_t / -15,27))] - k \left\{ 16,48 + 34 \left[ \frac{3}{2} \left( \frac{|h_s|}{123000} \right) - \frac{1}{2} \left( \frac{|h_s|}{123000} \right)^3 \right] \right\} \cdot \\
& \cdot [9134 \cdot (1 - \exp(-h_t/1,219)) - 2703,9915 \cdot (1 + \cos(2,5 \cdot \pi \cdot h_t / -15,27))]
\end{aligned}
\tag{69}$$

em que:  $0 < k \leq 0,0001$ ,  $h_s$  é a distância espacial em metros e  $h_t$  é a distância temporal em meses.

Neste caso, observa-se que o produto dos modelos espacial e temporal não tem muita significância, pois o valor  $k$  que pondera a soma e produto desses modelos está próximo de 0. O modelo produto-soma generalizado está ilustrado na Figura 37.

Devido à baixa variabilidade no espaço da variável precipitação média mensal, a superfície gerada através do modelo produto-soma generalizado no eixo espacial permaneceu aproximadamente constante em todos os tempos.

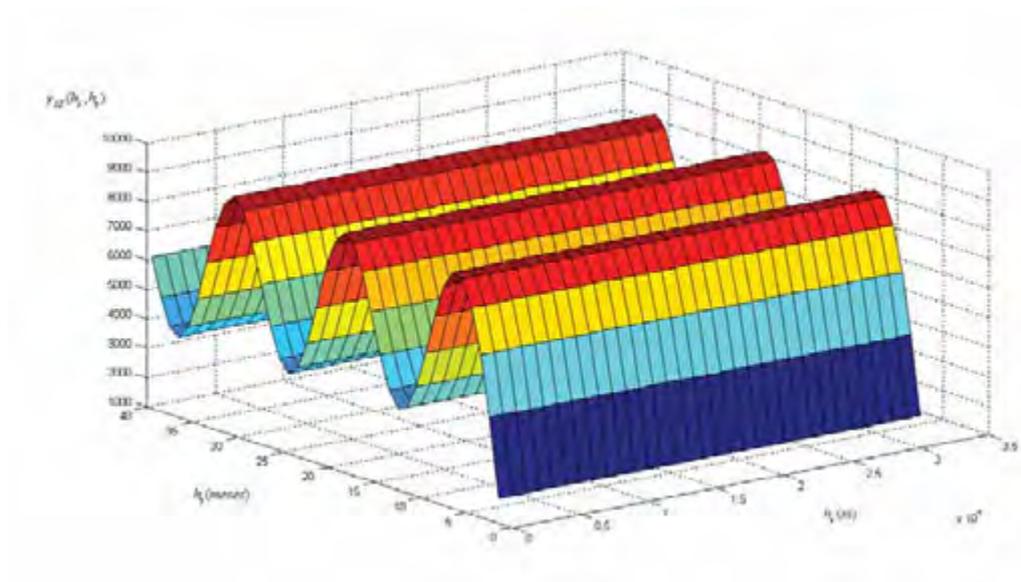


Figura 37: Superfície gerada através do modelo produto-soma generalizado da precipitação média mensal.

### 3.4.3. Modelo geoestatístico espaço-tempo

Para construção dos modelos e suas avaliações foram estudados vários cenários: retirada de uma observação, retirada de uma observação situada mais distante das demais, retirada de várias observações, análise de um transecto, análise de uma sub-área grande, etc. Foram considerados os seguintes casos:

- período de 31 anos mensalmente (total de 105 estações e 372 tempos distintos) e para avaliação retirando-se uma estação (D7-053) que são apresentados em a);
- período de 31 anos anualmente (total de 105 estações e 31 tempos distintos) e para avaliação retirando-se novamente a estação D7-053, apresentados em b);
- comparação da precipitação total anual observada com as estimativas obtidas para precipitação mensal e precipitação anual é apresentada em c);
- retirada de 2 localidades diferentes para estimação das variáveis: precipitações totais (máxima, média e mínima) anuais, totalizando omissão de 6 estações, cuja análises são apresentadas em d);
- retirada de uma estação que está mais distante das demais, apresentada em e);
- comparação entre as estimativas geradas pelo modelo e pelas médias aritméticas, omitindo as mesmas 6 estações do item d) é apresentada em f);
- um recorte no espaço-tempo, fazendo um transecto sobre 12 estações meteorológicas, cujas análises são apresentadas em g);
- uma área retangular grande com 3 faixas no eixo X e 5 faixas no eixo Y, análise apresentada em h);
- geração de mapas de superfície espacial através da krigagem ordinária em cada ano; no contexto espacial.

#### a) período de 31 anos mensalmente (janeiro de 1970 a dezembro de 2000)

O conjunto de casos é composto por 105 localizações no espaço  $(u_1, u_2, \dots, u_{105})$  e a cada uma é associada uma série temporal de 372 tempos distintos  $(t_1, t_2, \dots, t_{372})$ , representando os meses do período de janeiro de 1970 a dezembro de 2000. A variável aleatória  $Z(u_i, t_j)$ ,  $i = 1, 2, \dots, 105$ ;  $j = 1, 2, \dots, 372$ , representa os valores de precipitação mensal.

A estação de coleta D7-053 ( $x = 580409$ ,  $y = 7480219$ ) foi retirada do banco de dados para a geração do modelo espaço-temporal e posteriormente foi comparada com a sua estimativa dada pelo modelo gerado.

O semivariograma experimental e o modelo esférico foram gerados pelas médias temporais da componente  $M_I(u)$ . (Figura 38)

Os parâmetros do modelo esférico ajustado ao semivariograma experimental de  $M_I(u)$  foram gerados no software VARIOWIN e utilizados no R. Os parâmetros são: Efeito Pepita ( $C_0$ ) = 10,0665, alcance ( $a$ ) = 185307,8364 m e Contribuição ( $C_1$ ) = 27,0144, descritos na Equação (70).

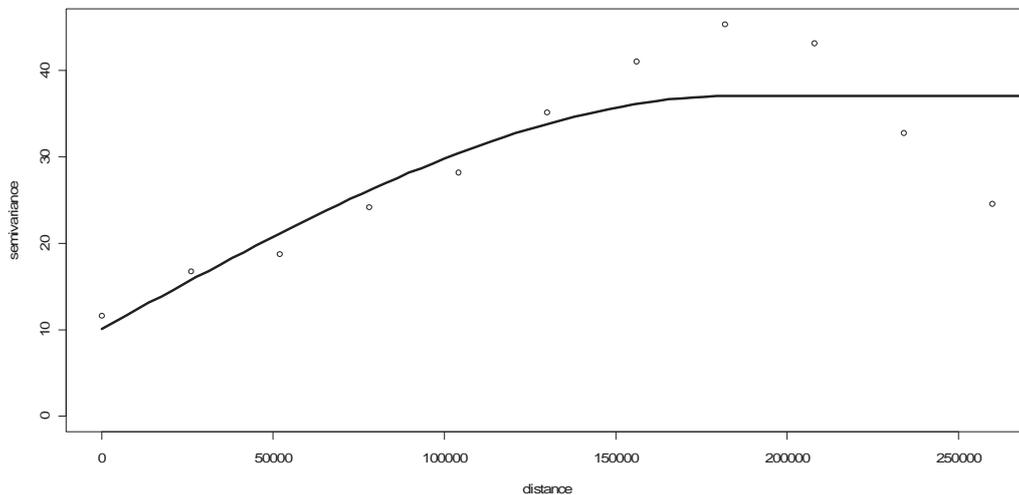


Figura 38: Semivariograma experimental e modelo esférico da componente  $M_I(u)$  para o item a), período de 31 anos mensalmente.

$$\gamma(h) = \begin{cases} 10,06, & |h|=0 \\ 10,06 + 27,01 \left[ \frac{3}{2} \left( \frac{|h|}{185307,83} \right) - \frac{1}{2} \left( \frac{|h|}{185307,83} \right)^3 \right], & 0 < |h| \leq 185307,83 \\ 37,08, & |h| > 185307,83 \end{cases} \quad (70)$$

A estimativa espaço-temporal é calculado para uma localização específica ( $u_0$ ) em todos os tempos ( $t_1, t_2, \dots, t_j$ ). Nesse caso, a estimação foi para a localidade ( $u_0$ ) referente à estação D7-053 com coordenadas  $x = 580409$  e  $y = 7480219$  em todos os 372 meses,  $t_1, t_2, \dots, t_{372}$ . (jan/1970 a dez/2000)

A componente  $\hat{M}_1(u_0)$  foi estimada aplicando-se a krigagem ordinária na vizinhança da localidade de interesse ( $u_0$ ), no caso da estação D7-053 resultou em 119,7845. Este resultado é independente do tempo  $t$ . As demais estimativas de  $\hat{Z}(u_0, t)$  foram calculadas individualmente utilizando-se os pesos  $\lambda$  obtidos da krigagem ordinária da Equação (53).

Nessa localidade de interesse  $u_0 = (580409,7480219)$  obteve-se 372 valores estimados de precipitação mensal. Esses valores estimados foram confrontados com os valores observados nessa localização.

A correlação e o Erro Absoluto Médio (EAM) entre os valores estimados e observados da precipitação mensal na localização de interesse foram de 0,923 e 22,4452, respectivamente. Isto significa que o modelo geoestatístico espaço-tempo apresentou estimativas próximas dos valores observados.

O gráfico de dispersão entre os valores estimados e observados da precipitação mensal durante os 31 anos está ilustrado na Figura 39. Neste gráfico, os índices pluviométricos foram divididos em dois períodos correspondentes aos meses chuvosos e secos. O período chuvoso corresponde aos meses de outubro, novembro, dezembro, janeiro, fevereiro e março, já no período seco foram os meses de abril, maio, junho, julho, agosto e setembro.

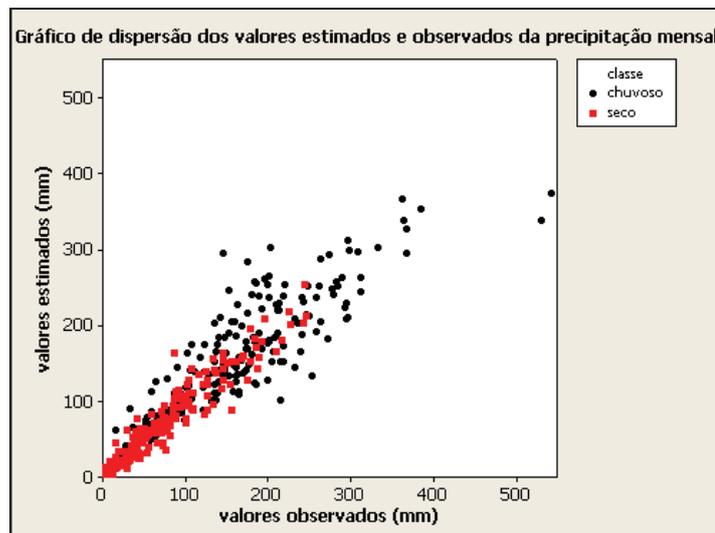


Figura 39: Gráfico de dispersão dos valores observados e estimados da precipitação mensal para o período de 31 anos– item a).

Pela Figura 39, observa-se que o modelo geoestatístico espaço-tempo forneceu melhores estimativas para os meses do período seco (pouca concentração de chuva, menores índices pluviométricos), enquanto que no período chuvoso, observam-se alguns pontos discordantes. O teste de normalidade Shapiro-Wilk foi aplicado às medidas dos erros da Figura 40 e conclui-se que os erros seguem uma distribuição normal com 0,05 de significância.

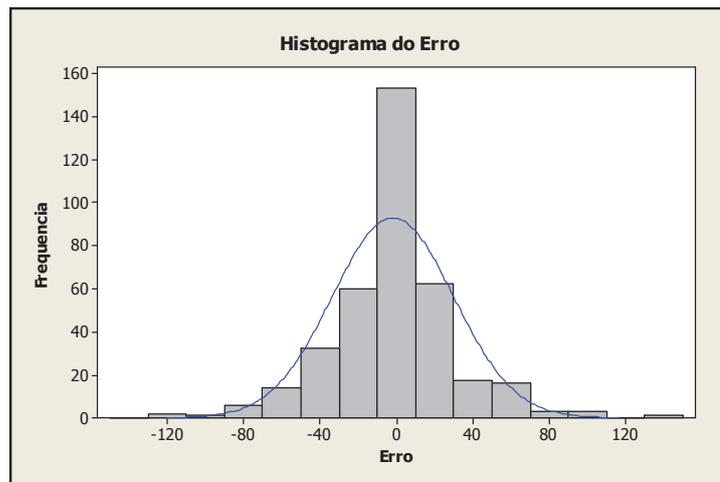


Figura 40: Histograma do Erro da precipitação mensal para o período de 31 anos – item a).

#### **b) período de 31 anos anualmente (1970 a 2000)**

Neste caso, continua-se com as 105 localizações no espaço  $(u_1, u_2, \dots, u_{105})$ , mas agora a cada uma está associada uma série temporal de 31 tempos distintos  $(t_1, t_2, \dots, t_{31})$  representando os anos de 1970 a 2000. Então, a variável aleatória  $Z(u_i, t_j)$ ,  $i = 1, 2, \dots, 105$ ;  $j = 1, 2, \dots, 31$ , representa os valores de precipitação anual, que é a soma de todas as precipitações mensais referentes a cada ano.

Como anteriormente, a estação de coleta D7-053 ( $x = 580409$ ,  $y = 7480219$ ) foi retirada do banco de dados para a geração do modelo espaço-temporal e posteriormente comparada com as estimativas.

O semivariograma experimental e o modelo esférico foram gerados com as médias temporais da componente  $M_I(u)$ . (Figura 41)

O modelo que melhor se ajustou ao semivariograma experimental de  $M_I(u)$  foi o esférico, cujos parâmetros foram gerados no software VARIOWIN e utilizados no R. Eles são: Efeito

Pepita ( $C_0$ ) = 1401,517 , alcance (a) = 186736,125 m e Contribuição ( $C_1$ ) = 3954,438, que estão descritos na Equação (71).

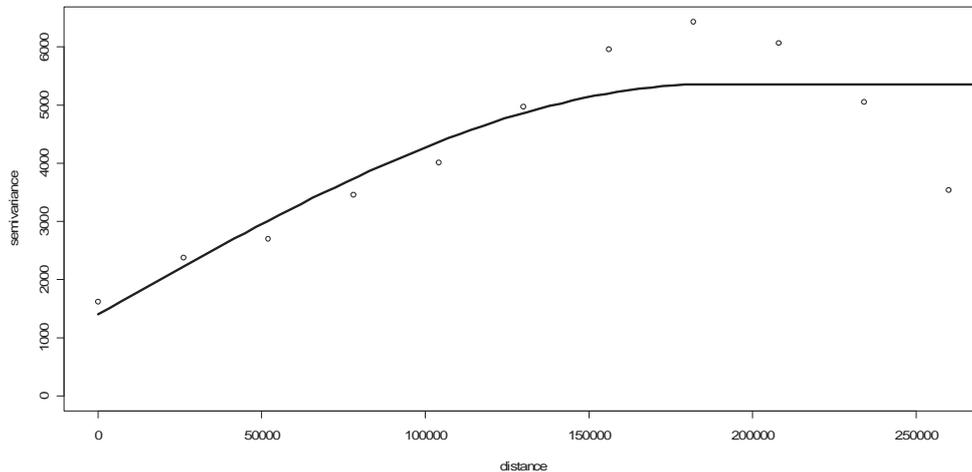


Figura 41: Semivariograma experimental e modelo esférico da componente  $M_1(u)$  para o período de 31 anos mensalmente - item b).

$$\gamma(h) = \begin{cases} 1401,51, & |h|=0 \\ 1401,51 + 3954,43 \left[ \frac{3}{2} \left( \frac{|h|}{186736,12} \right) - \frac{1}{2} \left( \frac{|h|}{186736,12} \right)^3 \right], & 0 < |h| \leq 186736,12 \\ 5355,95, & |h| > 186736,12 \end{cases} \quad (71)$$

A estimativa espaço-temporal foi calculado para a localidade ( $u_0$ ) referente à estação D7-053 com coordenadas  $x = 580409$  e  $y = 7480219$  em todos os 31 anos,  $t_1, t_2, \dots, t_{31}$ . (1970 a 2000)

A componente  $\hat{M}_1(u_0)$  foi calculada aplicando a krigagem ordinária na vizinhança da localidade de interesse ( $u_0$ ) resultando em 1431,861. Este resultado é independente do tempo  $t$ . As demais estimativas de  $\hat{Z}(u_0, t)$  foram calculadas individualmente utilizando-se os pesos  $\lambda$  obtidos dessa interpolação por krigagem ordinária.

Para a localidade de interesse  $u_0 = (580409, 7480219)$  obteve-se 31 estimativas de precipitação anual, que foram confrontados com os valores observados nessa localização.

Calculou-se, novamente, a correlação e o Erro Absoluto Médio (EAM) entre os valores estimados e observados da precipitação mensal na localização de interesse, que foram respectivamente iguais a 0,882 e 97,9738. Assim, tem-se que o modelo geoestatístico espaço-tempo apresentou estimativas próximas dos valores observados.

O gráfico de dispersão e o histograma do erro entre os valores estimados e observados da precipitação anual dos 31 anos está ilustrado nas Figuras 42 e 43, respectivamente.

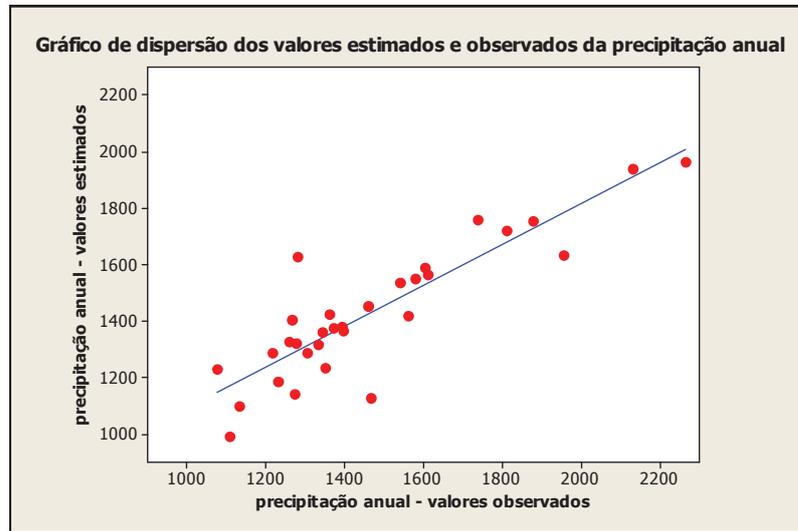


Figura 42: Gráfico de dispersão dos valores observados e estimados da precipitação anual para o período de 31 anos – item b).

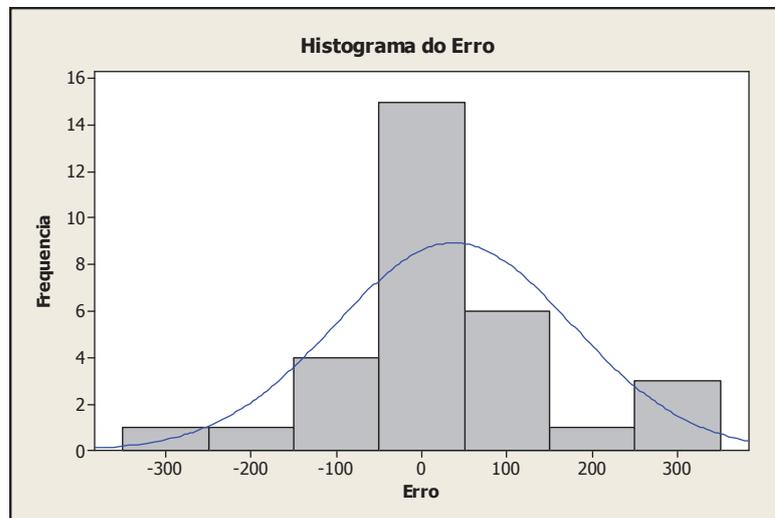


Figura 43: Histograma do Erro da precipitação anual para o período de 31 anos – item b).

O teste de normalidade Shapiro-Wilk foi aplicado às medidas dos erros da Figura 43 e conclui-se que os erros seguem uma distribuição normal no nível de significância de 0,05.

### c) comparação entre precipitação mensal e precipitação anual

As estimativas da precipitação anual podem ser obtidas somando-se todas as estimativas mensais do período de um ano. Essas estimativas foram comparadas com aquelas obtidas diretamente com os totais anuais, ou seja, com os valores estimados gerados pelos modelos obtidos em a) (período de 31 anos mensalmente) e b) (período de 31 anos anualmente).

Neste procedimento, foram consideradas três variáveis: precipitação total anual observada (PTAO), precipitação total anual estimada (PTAE) e precipitação total anual estimada mensalmente (PTAEM), sendo a última variável a soma das precipitações mensais descritas anteriormente.

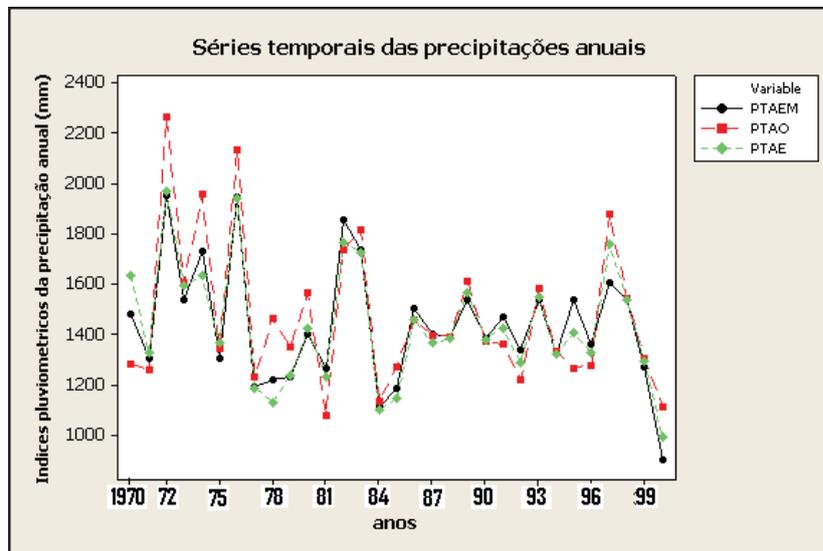


Figura 44: Séries temporais referentes às precipitações totais

As séries temporais referentes a essas variáveis estão apresentadas na Figura 44. As variáveis PTAEM e PTAE apresentam comportamentos bastante semelhantes no período de estudo e com valores iguais em anos como os de 1971 a 1977, 1979 a 1994 e 1997 a 2000. Comparando com os valores observados, os valores estimados do modelo espaço-tempo estão próximos dos valores observados, com pequeno afastamento para os anos em que os índices pluviométricos foram maiores como nos anos de 1972, 1974, 1976 e 1997. As correlações entre as variáveis (PTAO, PTAE) e (PTAO, PTAEM) foram de 0,882 e 0,878 respectivamente e o Erro Absoluto Médio (EAM) foram de 97,9738 e 110,193 respectivamente. De acordo com os

resultados das correlações e dos EAM's, a variável PTAE apresentam melhores estimativas em comparação com a PTAEM.

#### **d) precipitações totais anuais: máxima, média e mínima**

Neste caso, três variáveis de estudo foram geradas representando as precipitações totais anuais: máxima, média e mínima. A precipitação total anual máxima (PTAMa) são os valores máximos anuais referente ao período de 31 anos, a precipitação total anual média (PTAMe) são os valores médios e a precipitação total anual mínima (PTAMi) são os valores mínimos.

Para a estimação, foram retiradas do banco de dados um total de seis estações de coleta (2 para PTAMa, 2 para PTAMe e 2 para PTAMi) escolhidas através da visualização dos mapas de superfícies gerados a partir dessas três variáveis conforme as Figuras 45 e 46. Os mapas de superfícies gerados a partir dessas variáveis auxiliaram na escolha das estações, pois visualizam as regiões de maiores e menores concentrações de chuva.

As estações correspondentes a PTAMi possuem os menores índices pluviométricos, em torno de 400 mm, representados por C8-050 ( $x = 486169$ ,  $y = 7649458$ ) e C8-019 (410154,7635148). As estações correspondentes a PTAMe possuem os índices pluviométricos, em torno de 1500 mm, representados por C7-028 (569158,7646827) e D8-047 (474226,7546827). Por último, as estações referentes à PTAMa possuem os índices pluviométricos mais elevados, em torno de 2350 mm, representados por D8-013 (491438,7493344) e C8-049 (468865,7653813).

#### **- precipitação total anual máxima (PTAMa):**

As correlações entre os valores observados e estimados das estações C8-049 e D8-013 são 0,862 e 0,913 respectivamente e o Erro Absoluto Médio (EAM) foram de 30,3949 e 24,7118 respectivamente.

Mais uma vez, o modelo geoestatístico espaço-tempo forneceu melhores estimativas para os meses do período seco (pouca concentração de chuva, menores índices pluviométricos), do que no período chuvoso. A medida que o índice pluviométrico aumenta (valores observados), as estimativas ficam mais distantes dos valores observados.

O teste de normalidade Shapiro-Wilk foi aplicado às medidas dos erros da Figura 47 e conclui-se que os erros seguem uma distribuição normal com nível de significância de 0,05.

- precipitação total anual média (PTAME):

As correlações entre os valores observados e estimados das estações C7-028 e D8-047 são 0,817 e 0,935, respectivamente e o Erro Absoluto Médio (EAM) foram de 35,8813 e 20,7359 respectivamente.

Através dos resultados das altas correlações e as análises dos gráficos de dispersão da Figura 48, pode-se concluir que o modelo geoestatístico espaço-tempo apresentou estimativas próximas dos valores observados.

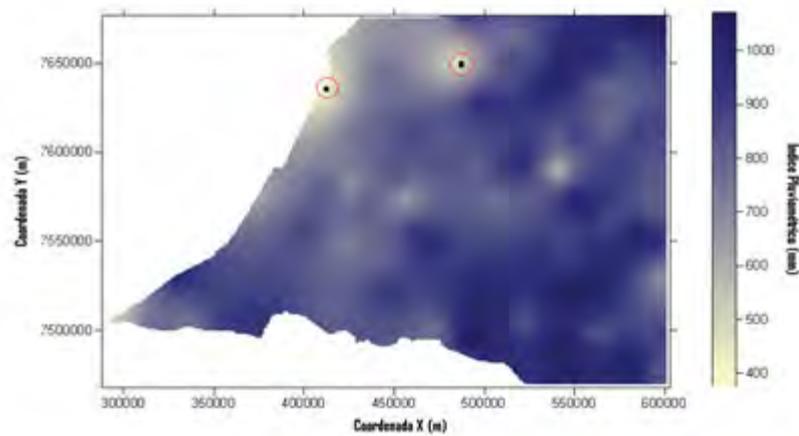


Figura 45: Mapa de superfície da variável precipitação total anual mínima e estações retiradas

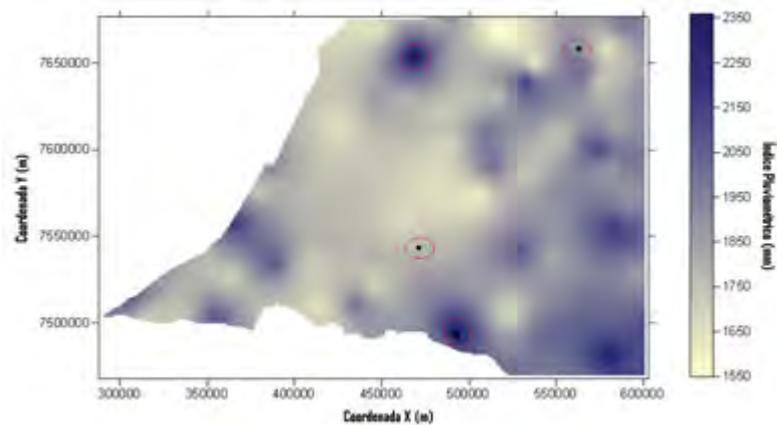


Figura 46: Mapa de superfície da variável precipitação total anual máxima e estações retiradas

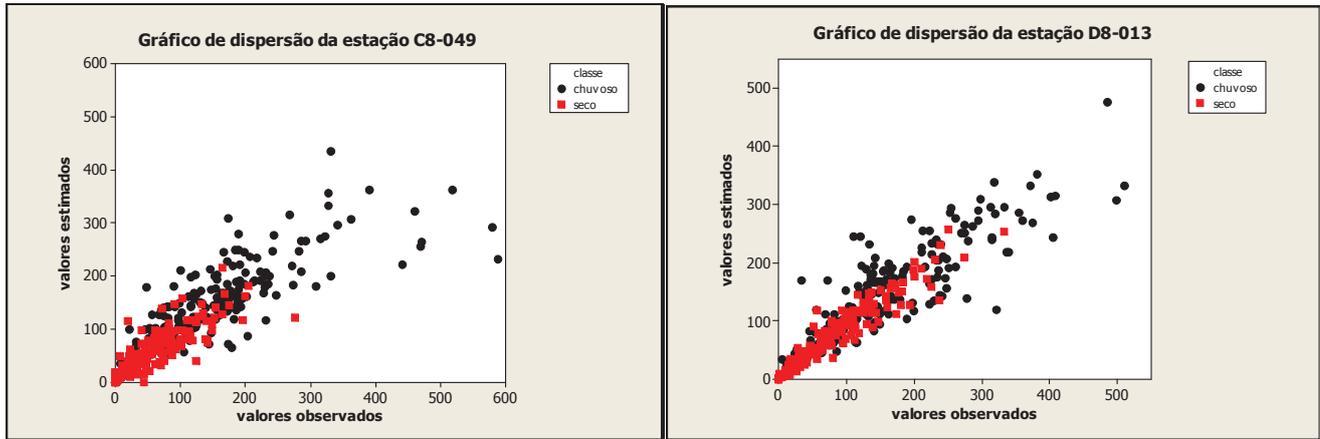


Figura 47: Gráficos de dispersão dos valores observados e estimados da precipitação total anual máxima – item d).

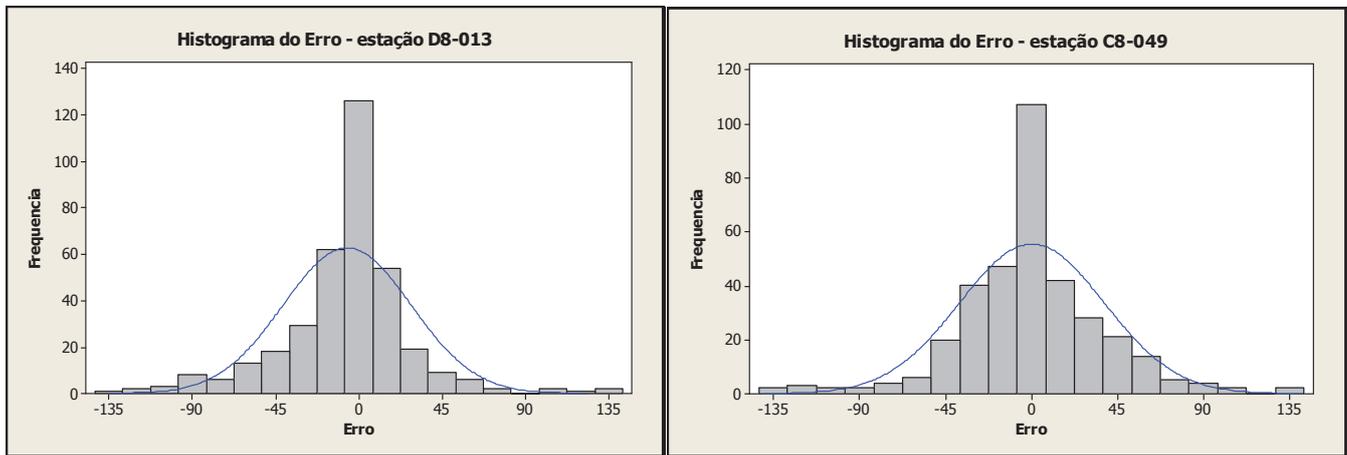


Figura 48: Histogramas do Erro da precipitação total anual máxima – item d).

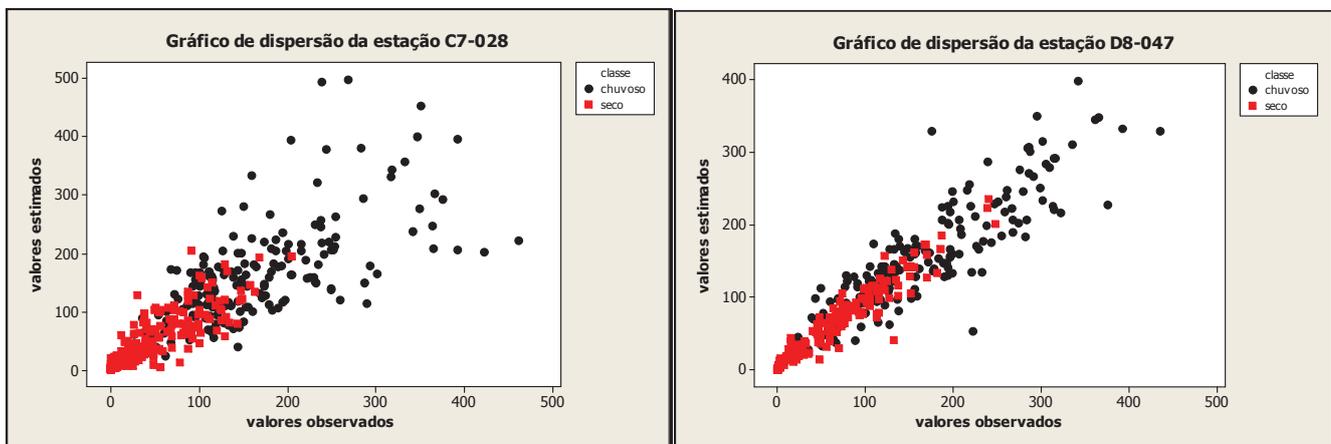


Figura 49: Gráficos de dispersão dos valores observados e estimados da precipitação total anual média – item d).

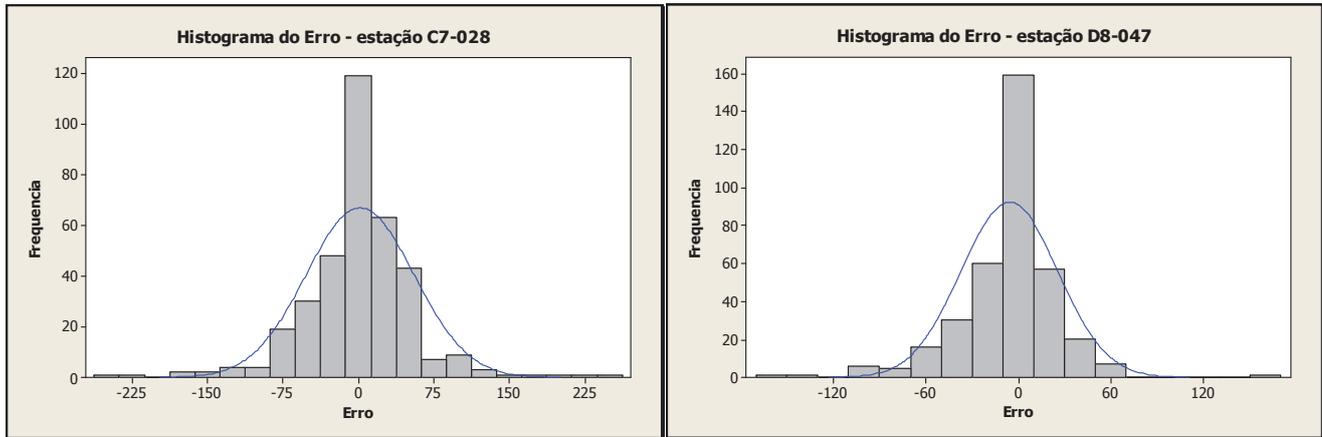


Figura 50: Histogramas do Erro da precipitação total anual média – item d).

De modo geral, o modelo geostatístico espaço-tempo conseguiu captar melhor os valores correspondentes aos meses do período seco (pouca concentração de chuva, menores índices pluviométricos). No período chuvoso, a maioria dos valores estimados da estação C7-028 não apresentou valores próximos dos observados, ao contrário dos da estação D8-047 que apresentaram valores próximos aos observados.

O teste de normalidade Shapiro-Wilk foi aplicado às medidas dos erros da Figura 50 e conclui-se que os erros seguem uma distribuição normal com 0,05 de significância.

- precipitação total anual mínima (PTAMi):

As correlações entre os valores observados e estimados das estações C8-050 e C8-019 são 0,886 e 0,902, respectivamente e o Erro Absoluto Médio (EAM) foram de 26,1724 e 24,3981 respectivamente.

Através dos resultados das altas correlações e as análises dos gráficos de dispersão apresentados na Figura 51 e histogramas do erro apresentadas na Figura 52, pode-se concluir que o modelo geostatístico espaço-tempo apresentou estimativas próximas dos valores observados, conseguindo captar melhor os valores correspondentes aos meses do período seco (pouca concentração de chuva, menores índices pluviométricos). Em relação ao período chuvoso houve muitos pontos discordantes nas duas estações de coleta.

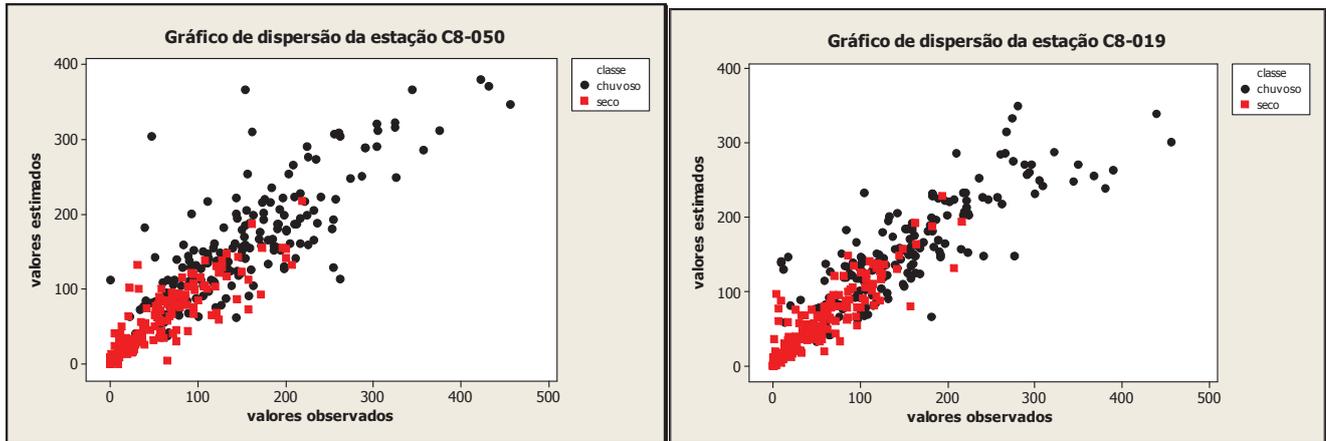


Figura 51: Gráficos de dispersão dos valores observados e estimados da precipitação total anual mínima – tem d).

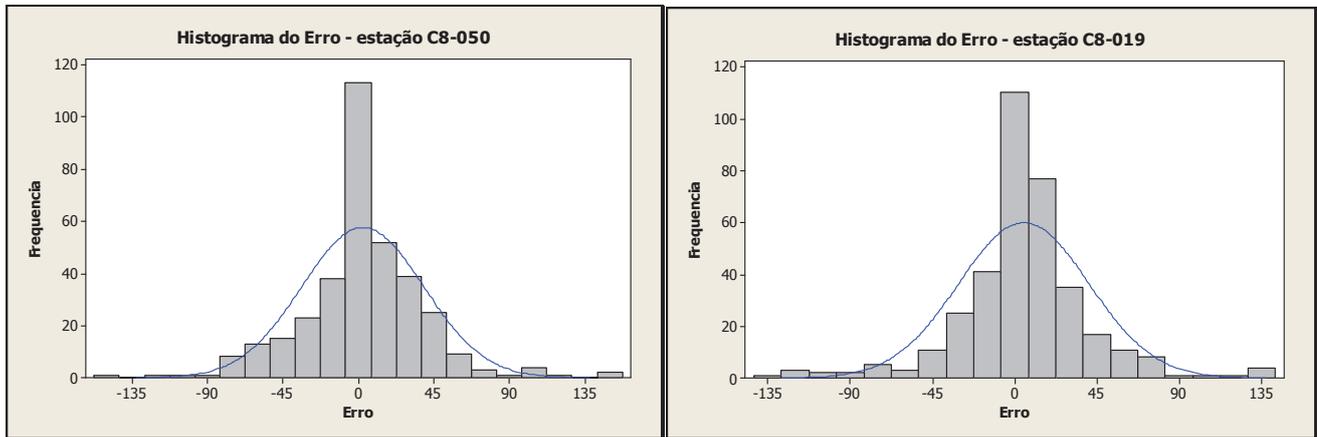


Figura 52: Histogramas do Erro da precipitação total anual mínima – item d).

Além disso, pode-se concluir que o modelo geoestatístico espaço-tempo apresenta melhores estimativas para os valores com menores índices pluviométricos (abaixo de 200 mm), conforme Figura 51.

O teste de normalidade Shapiro-Wilk foi aplicado às medidas dos erros da Figura 52 e conclui-se que os erros seguem uma distribuição normal com 0,05 de significância.

#### e) estação de coleta distante das demais estações

Uma das maneiras de comprovar a eficiência das estimativas do modelo geoestatístico espaço-tempo foi retirar a estação de coleta mais distante das demais estações e gerar o modelo

sem essa estação comparando com os seus valores reais. A estação retirada foi a C8-055, com coordenadas  $x = 425806$  e  $y = 7613089$ , que está ilustrada na Figura 53.

Neste caso foram 104 localizações no espaço  $(u_1, u_2, \dots, u_{104})$  e em cada uma com uma série temporal de 372 tempos distintos  $(t_1, t_2, \dots, t_{372})$  representando os meses do período de janeiro de 1970 a dezembro de 2000. A variável aleatória  $Z(u_i, t_j)$ ,  $i = 1, 2, \dots, 107$ ;  $j = 1, 2, \dots, 372$ , representa os valores de precipitação mensal.

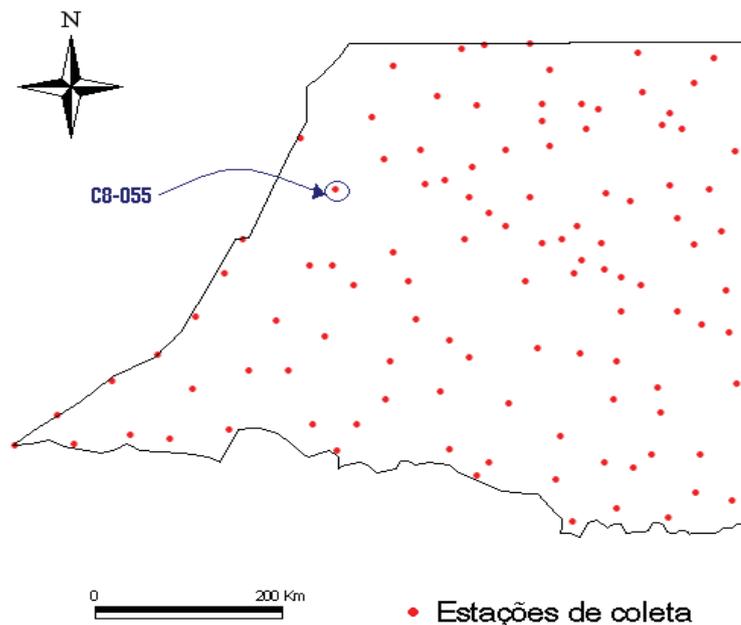


Figura 53: Distribuição espacial das estações de coleta do oeste paulista e estação retirada.

O semivariograma experimental e o modelo esférico foram gerados através das médias temporais da componente  $M_I(u)$ . (Figura 54)

O modelo que melhor se ajustou ao semivariograma experimental de  $M_I(u)$  foi o esférico, cujos parâmetros são: Efeito Pepita ( $C_0$ ) = 9,9663, alcance ( $a$ ) = 185800 m e Contribuição ( $C_1$ ) = 29,457, que estão descritos na Equação (72).

$$\gamma(h) = \begin{cases} 9,9663, & |h|=0 \\ 9,9663 + 29,457 \left[ \frac{3}{2} \left( \frac{|h|}{185800} \right) - \frac{1}{2} \left( \frac{|h|}{185800} \right)^3 \right], & 0 < |h| \leq 185800 \\ 39,4233, & |h| > 185800 \end{cases}, \quad (72)$$

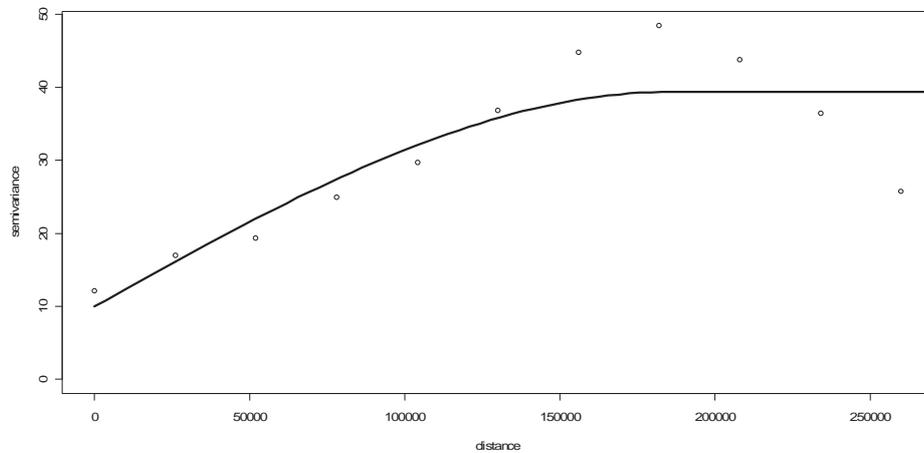


Figura 54: Semivariograma experimental e modelo esférico da componente  $M_1(u)$  para a estação de coleta distante das demais estações - item e).

A estimativa espaço-temporal foi calculada para a estação C8-055 ( $u_0$ ) em todos os tempos ( $t_1, t_2, \dots, t_{372}$ ), ou seja, tem todos os meses de 31 anos.

A componente  $\hat{M}_1(u_0)$  é igual a 105,4776 e nessa localidade de interesse  $u_0 = (425806, 7613089)$  obteve-se 372 valores estimados de precipitação mensal que foram confrontados com os valores observados.

A correlação e o Erro Absoluto Médio entre os valores estimados e observados da precipitação mensal na localização de interesse foram de 0,881 e 26,1141 respectivamente.

Isto significa que o modelo geoestatístico espaço-tempo apresentou estimativas próximas dos valores observados mesmo com a presença de dados distantes (estações de coletas distantes). O gráfico de dispersão e o histograma do erro entre os valores estimados e observados da precipitação mensal durante os 31 anos estão ilustrados na Figura 55 e Figura 56, respectivamente.

O teste de normalidade Shapiro-Wilk foi aplicado às medidas dos erros da Figura 56 e conclui-se que os erros seguem uma distribuição normal com nível de significância de 0,05.

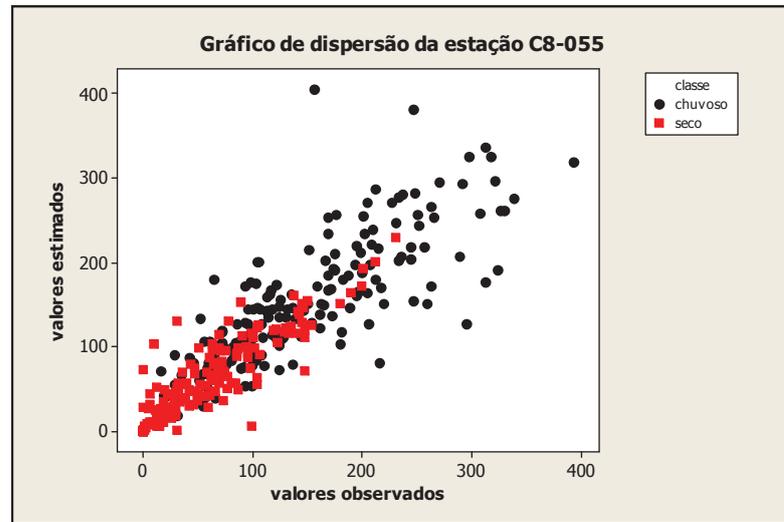


Figura 55: Gráfico de dispersão da precipitação mensal retirando a estação C8-055 – item e).

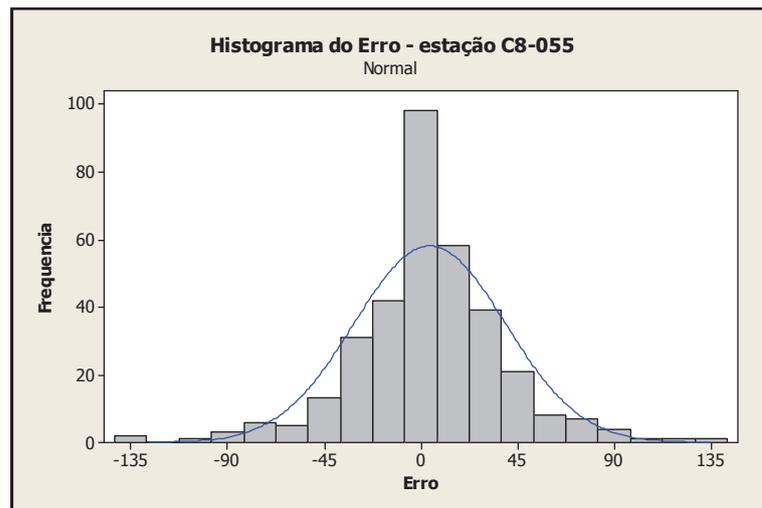


Figura 56: Histograma do erro da estação C8-055 – item e).

#### f) comparação entre as estimativas geradas através do modelo e através das médias aritméticas

As estações de coleta retiradas foram as mesmas referentes às variáveis precipitação mensal mínima e precipitação mensal máxima do item d), ou seja as estações C8-50, C8-19, D8-013 e C8-049. O semivariograma experimental e o modelo esférico das médias temporais da componente  $M_1(u)$  utilizados foram os gerados para o período de 31 anos mensalmente, item a) ilustrado na Figura 38 e Equação (70).

Supostamente, a precipitação pluviométrica apresenta o mesmo comportamento para distâncias próximas, portanto foram calculadas médias aritméticas referentes às três estações mais próximas de cada estação retirada. Essas estações estão ilustradas na Figura 57 e descritas a seguir.

- precipitação mensal mínima:

Estações retiradas: C8-050 e C8-019.

Estações próximas as retiradas: (C8-049, C8-052, C7-051) e (C8-008, C8-055, C8-043) respectivamente.

- precipitação mensal máxima:

Estações retiradas: D8-013 e C8-049.

Estações próximas as retiradas: (D8-050, D8-006, D7-065) e (C8-022, C8-050, C8-045) respectivamente.

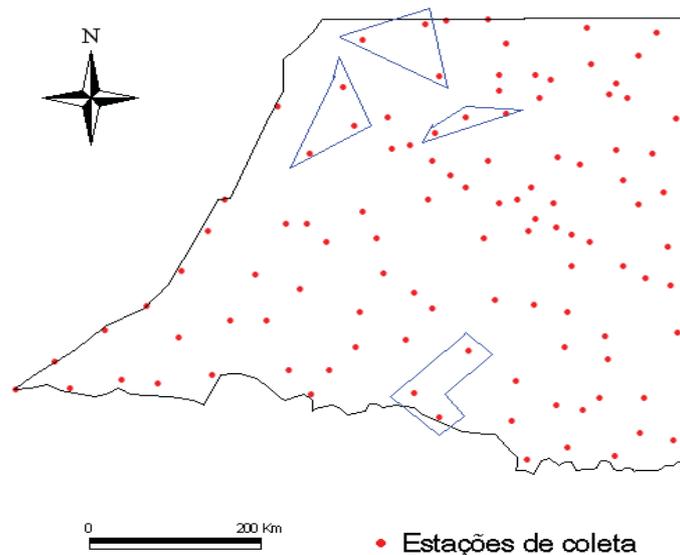


Figura 57: Estações de coleta próximas as estimativas

As correlações e os EAM's das estações retiradas entre os valores estimados, observados e gerados a partir das médias aritméticas estão apresentadas na Tabela 3.

Com exceção da estação C8-019, as maiores correlações e os menores EAM's ocorreram entre os valores observados e estimados, mas a diferença é mínima. Esse resultado é esperado pelo

fato de o modelo geoestatístico espaço-tempo considerar todas as estações tanto no espaço quanto no tempo para o cálculo de suas estimativas.

Tabela 3: Correlações e EAM's entre valores estimados, observados e médios

Estações retiradas	Correlações e EAM's entre valores observados e estimados	Correlações e EAM's entre valores observados e médios
<b>C8-050</b>	0,886 26,172	0,884 26,474
<b>C8-019</b>	0,902 24,398	0,904 24,045
<b>D8-013</b>	0,913 24,712	0,906 25,606
<b>C8-049</b>	0,862 30,395	0,843 31,236

#### g) recorte no espaço-tempo

Neste caso foram 111 localizações no espaço ( $u_1, u_2, \dots, u_{111}$ ) e em cada localização possui uma série temporal de 132 tempos distintos ( $t_1, t_2, \dots, t_{132}$ ) representando os meses do período de janeiro de 1990 a dezembro de 2000. A variável aleatória  $Z(u_i, t_j)$ ,  $i = 1, 2, \dots, 111$ ;  $j = 1, 2, \dots, 132$ , representa os valores de precipitação mensal.

O semivariograma experimental e o modelo esférico foram gerados pelas médias temporais da componente  $M_I(u)$  e está ilustrado na Figura 58. O mapa de superfície para as médias temporais está ilustrado na Figura 59.

O modelo que melhor se ajustou ao semivariograma experimental de  $M_I(u)$  foi o esférico, cujos parâmetros são: Efeito Pepita ( $C_0$ ) = 30,509, alcance ( $a$ ) = 288000 m e Contribuição ( $C_1$ ) = 29,598, que estão descritos na Equação (73).

$$\gamma(h) = \begin{cases} 30,509, & |h| = 0 \\ 30,509 + 29,598 \left[ \frac{3}{2} \left( \frac{|h|}{288000} \right) - \frac{1}{2} \left( \frac{|h|}{288000} \right)^3 \right], & 0 < |h| \leq 288000 \\ 60,107, & |h| > 288000 \end{cases}, \quad (73)$$

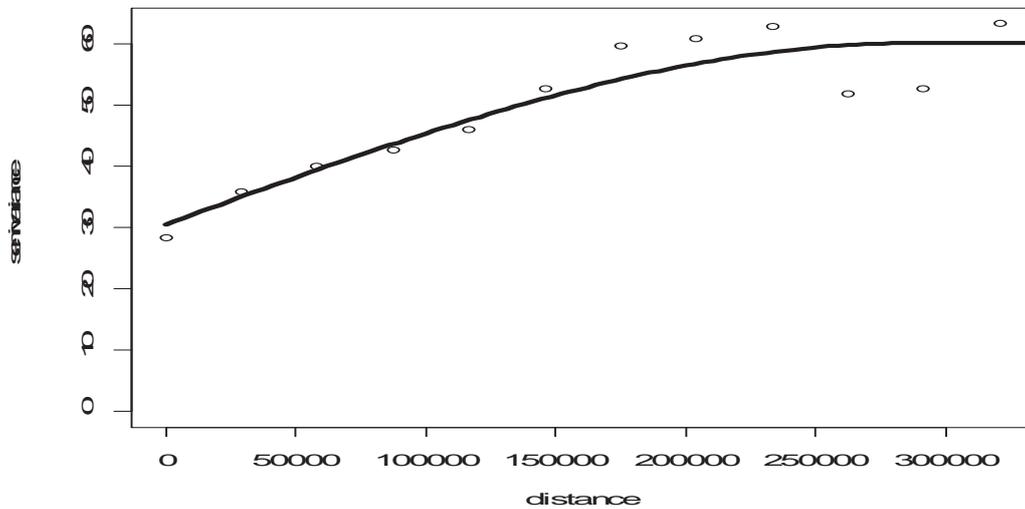


Figura 58: Semivariograma experimental e modelo esférico da componente  $M_I(u)$  para o item g).

A região mais chuvosa pertence à região com coordenadas em  $X$  mais altas, com maiores concentrações nas coordenadas  $(x, y) = (540000, 7400000)$  e  $(510000, 7500000)$ . Nas regiões com maiores altitudes ocorrem maiores índices pluviométricos, como se pode observar no oeste paulista na Serra de Marília localizado na parte inferior a direita da Figura 59.

O trecho ilustrado na Figura 60 corresponde a 12 localizações no espaço, em que cada uma possui uma serie temporal de 132 informações correspondendo aos meses de janeiro/1990 a dezembro/2000, totalizando 1584 elementos amostrais no espaço-tempo  $(Z(u_i, t_j), i = 1, \dots, 12; j = 1, \dots, 132)$ . A estimação espaço-temporal para cada localização de interesse  $(u_0)$  foi feita para cada uma das 12 localizações no espaço.

As correlações  $(\rho)$  entre os pontos observados e estimados estão apresentados na Tabela 4. Todas as correlações ficaram próximas de 1, isto significa que os valores estimados através do modelo espaço-temporal apresentaram estimativas próximas aos valores reais.

Por último, foi gerado o mapa de superfície espaço-temporal a partir da krigagem ordinária nos 1584 elementos amostrais. (Figura 61)

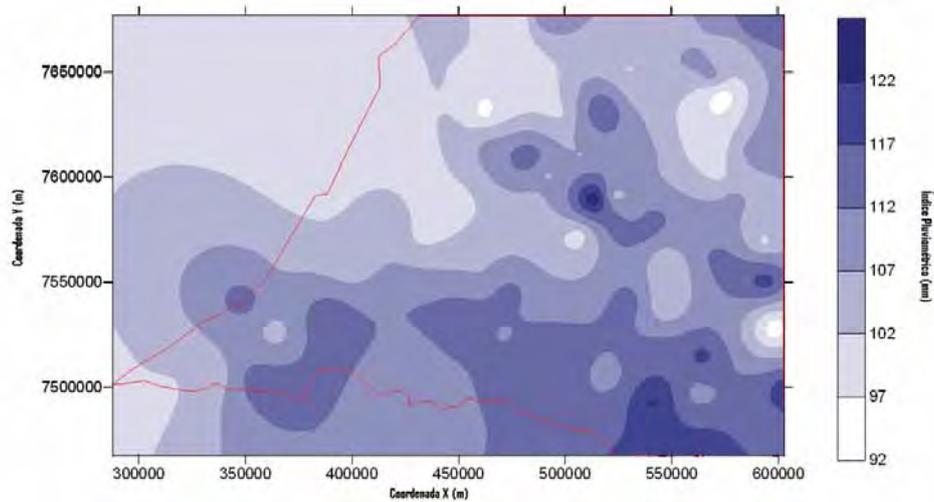


Figura 59: Mapa de superfície das médias temporais do oeste paulista.

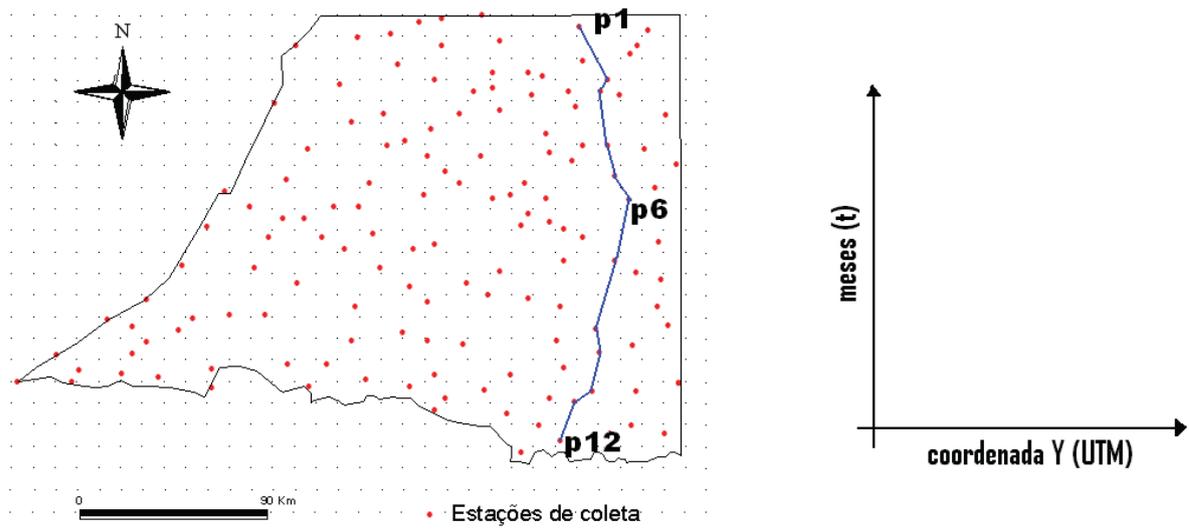


Figura 60: Trecho e eixo escolhido para calcular as estimativas espaço-temporal.

O mapa de superfície apresentado na Figura 61 descreve o comportamento do regime pluviométrico ao longo do espaço-tempo. Observa-se tanto no eixo temporal quanto no eixo espacial que o regime pluviométrico tem um comportamento periódico. O período mais seco está representado na parte superior da Figura 61 (ano de 2000) e o mais chuvoso próximo aos meses de dezembro de 1998 e janeiro de 1999.

A altitude influencia no regime pluviométrico do trecho em estudo, pois as maiores altitudes estão na região pertencente a Serra de Marília que contem coordenadas em Y mais baixas e os maiores índices pluviométricos (região chuvosa).

Tabela 4. Correlações entre pontos observados e estimados.

	<b>P1</b>	<b>P2</b>	<b>P3</b>	<b>P4</b>	<b>P5</b>	<b>P6</b>
<b>Coordenada X</b>	555413	565678	569024	569158	572409	579251
<b>Coordenada Y</b>	7672196	7640798	7614957	7646318	7600183	7589082
<b>Correlação (ρ)</b>	0,967	0,965	0,934	0,982	0,961	0,965

Tabela 4. Correlações entre pontos observados e estimados. (continuação)

	<b>P7</b>	<b>P8</b>	<b>P9</b>	<b>P10</b>	<b>P11</b>	<b>P12</b>
<b>Coordenada X</b>	563491	561657	565154	553074	565160	546169
<b>Coordenada Y</b>	7526425	7496912	7516023	7491409	7515348	7472981
<b>Correlação (ρ)</b>	0,970	0,956	0,960	0,955	0,966	0,940

O produto gerado através dessa modelagem espaço-temporal pode ser utilizado na visualização dos dados, por exemplo, na localização de coordenada Y igual a 7540 km, o mês mais seco do período em estudo é janeiro de 1995. Outra informação que se pode retirar da Figura 61 são as categorias de precipitação pluviométrica (habitual, chuvoso e seco) conforme classificados pelos autores Shimabukuro, Flores e Oliveira (2003), pois esse mapa apresenta as concentrações das variações dos índices pluviométricos.

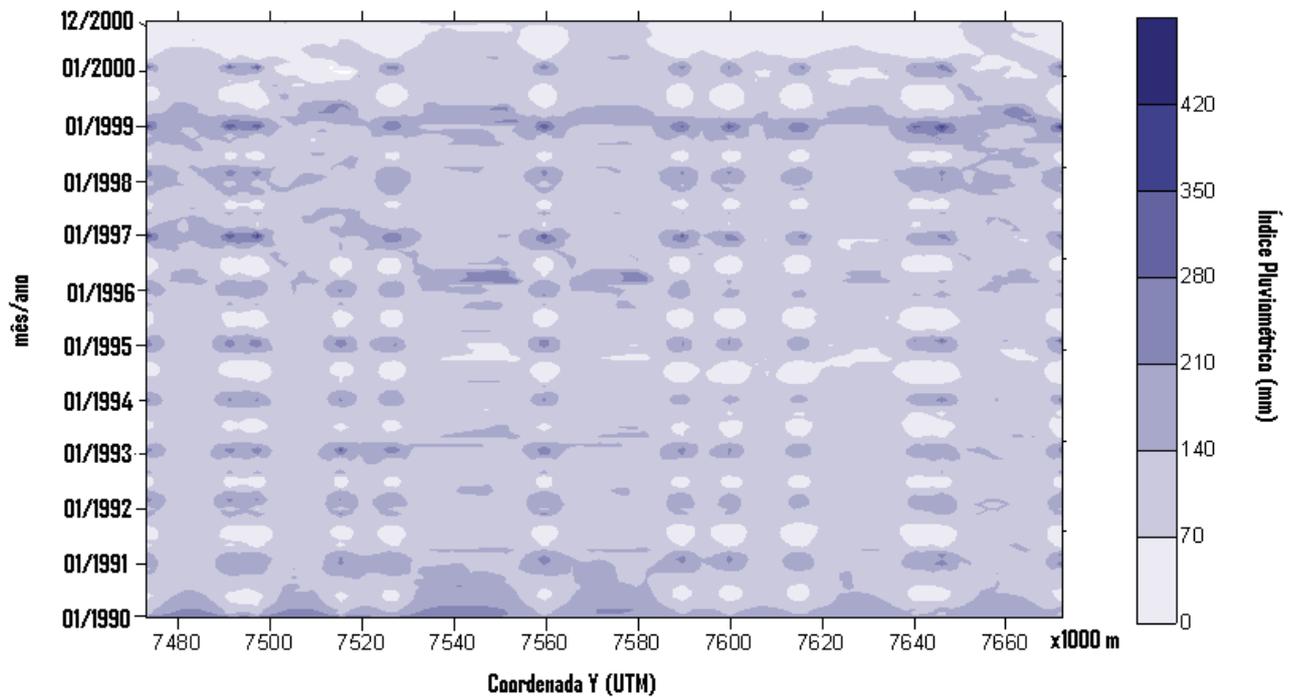


Figura 61: Mapa de superfície espaço-tempo do regime pluviométrico do trecho escolhido – item g).

#### h) região no espaço-tempo

São 105 localizações no espaço  $(u_1, u_2, \dots, u_{105})$  e em cada uma possui uma série temporal de 372 tempos distintos  $(t_1, t_2, \dots, t_{372})$  representando os meses do período de janeiro de 1970 a dezembro de 2000. A variável aleatória de interesse  $Z(u_i, t_j)$ ,  $i = 1, 2, \dots, 105$ ;  $j = 1, 2, \dots, 372$ , representa os valores de precipitação mensal.

O semivariograma experimental e o modelo esférico foram os gerados no item a) encontrados na Figura 38 e Equação (70), a região selecionada está ilustrada na Figura 62.

As estimações espaços-temporais foram geradas em uma grade espacial de 9x5 referente a 45 pontos no espaço em que a cada ponto corresponde uma série de 372 dados temporais, assim totalizando 16740 pontos no espaço-tempo. A importância dos dados estimados por modelos espaço-tempo é a maior precisão se for comparado somente pelos modelos espaciais e temporais. A distribuição da grade 9x5 está ilustrada na Figura 63.

O passo seguinte foi à escolha das faixas para a geração das superfícies espaço-tempo. As faixas escolhidas foram às referentes à coordenada X para 500000m, 540000m, 580000m e coordenada Y para 7500000m, 7540000m, 7580000m, 7620000m e 7660000m que estão representadas na Figura 64.

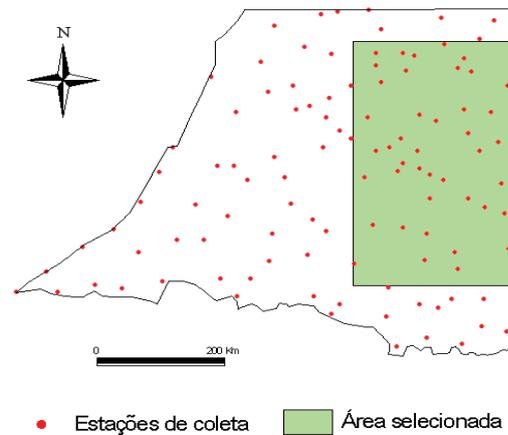


Figura 62: Região selecionada para o experimento – item i).

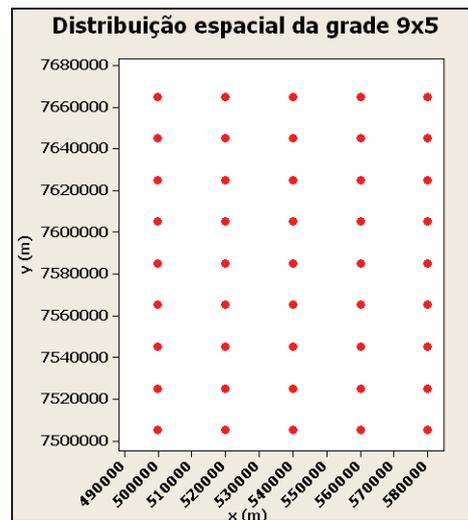


Figura 63: Distribuição espacial da grade 9x5 para a região selecionada.

As faixas em X possuem 9 pontos no espaço e 372 pontos no tempo totalizando 3348 pontos no espaço-tempo. Já em Y possuem 5 pontos no espaço e 372 pontos no tempo totalizando 1860 pontos no espaço-tempo. Os mapas de superfície espaço-tempo foram gerados para todas as 8 faixas a partir da técnica de interpolação por Krigagem Ordinária.

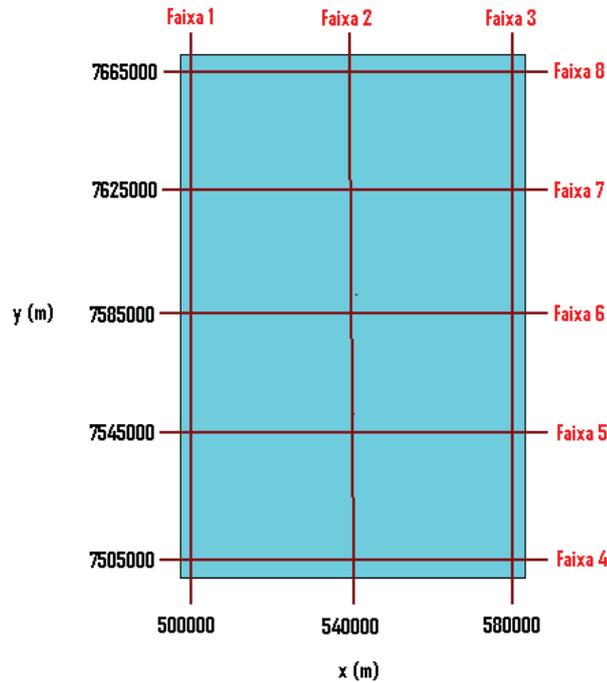


Figura 64: Faixas selecionadas para as superfícies espaço-tempo – item i).

O objetivo desses mapas de superfície espaço-tempo foi visualizar o comportamento da precipitação pluviométrica e para analisar o regime pluviométrico, como os períodos mais seco e chuvoso de cada faixa analisada.

O mapa de superfície espaço-tempo da Figura 65 apresenta alguns trechos com período mais chuvosos referentes à aproximadamente aos anos de 1983 e 1998. O período mais seco está representado no ano de 2000. No restante do mapa desta figura a precipitação pluviométrica se comporta de modo cíclico, ou seja, alternando períodos seco e chuvoso, tanto para somente o eixo temporal quanto para o eixo espacial. Conclui-se o mesmo para as Faixa 2 e Faixa 3.

À medida que se percorre o eixo X, conclui-se que as regiões correspondentes aos períodos seco e chuvoso continuam as mesmas para o casos temporal e espacial (eixo Y).

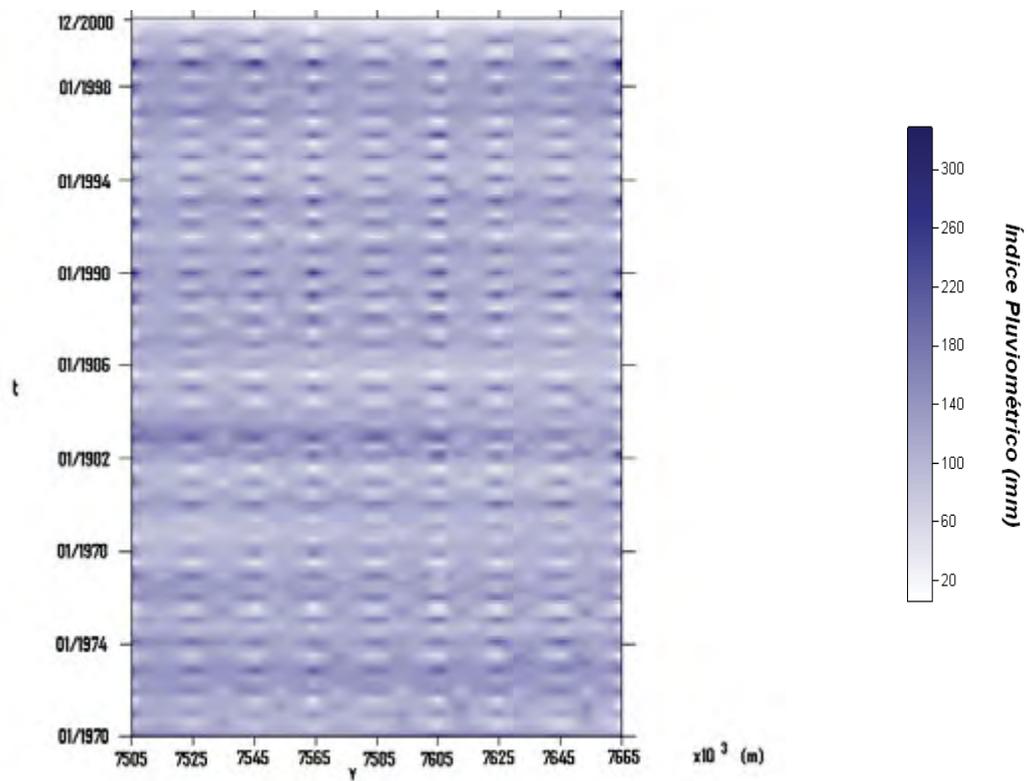


Figura 65: Mapa de superfície espaço-tempo da coordenada X = 500000 (Faixa 1).

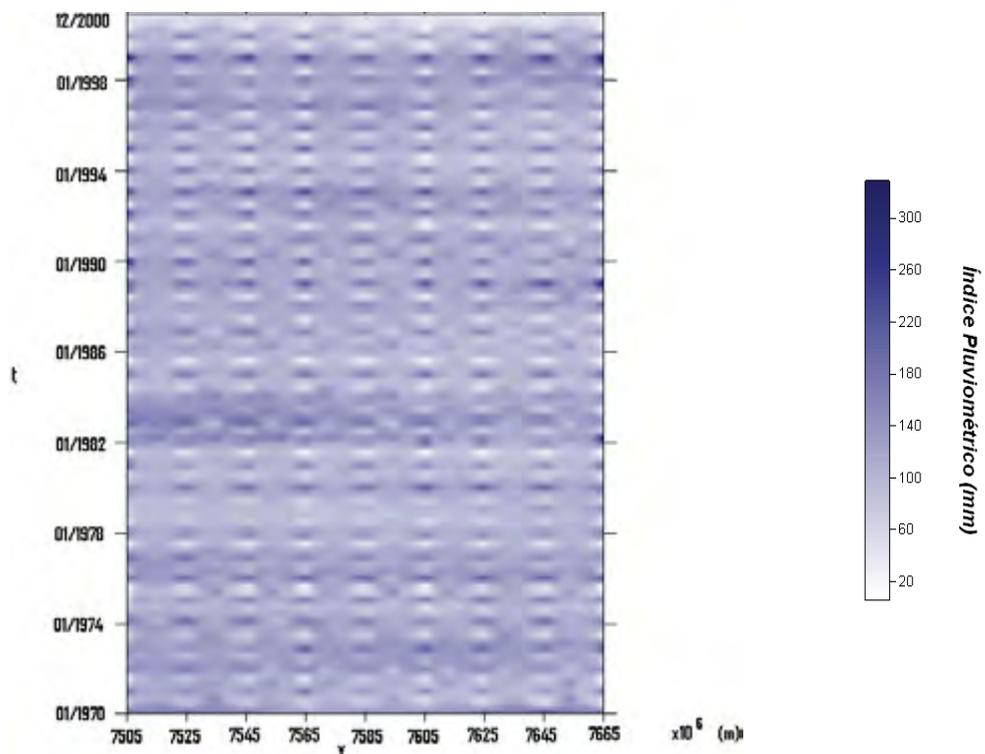


Figura 66: Mapa de superfície espaço-tempo da coordenada X = 540000 (Faixa 2).

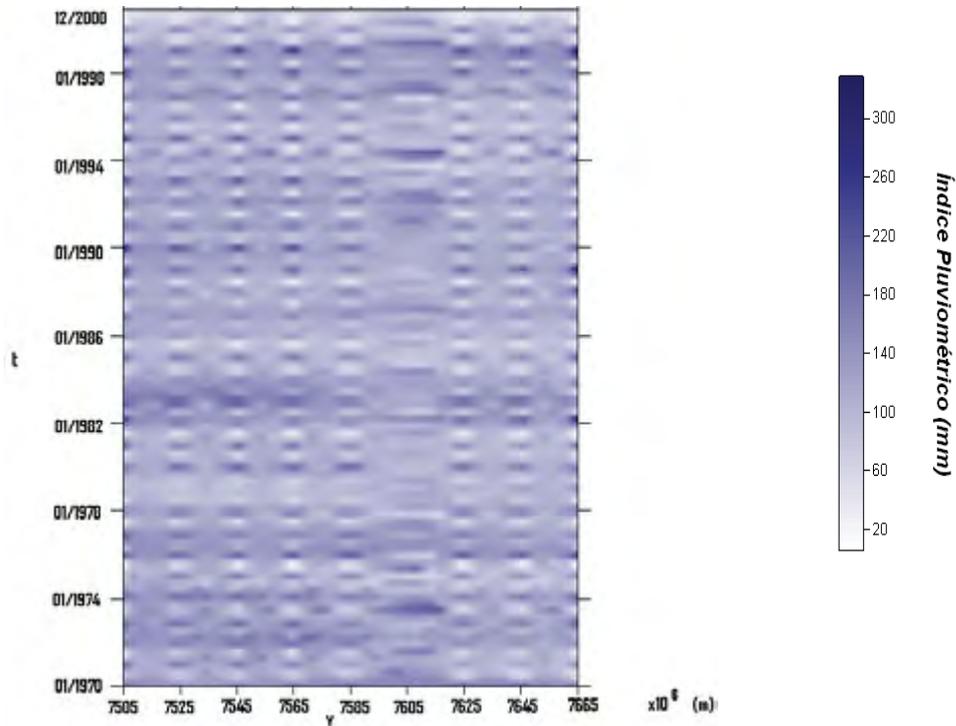


Figura 67: Mapa de superfície espaço-tempo da coordenada X = 580000 (Faixa 3).

O mapa de superfície espaço-tempo da Figura 68 apresentou regiões com maiores índices pluviométricos correspondentes ao período chuvoso que estão aproximadamente nos anos de 1983 e 1998. Já o período mais seco está no ano de 2000.

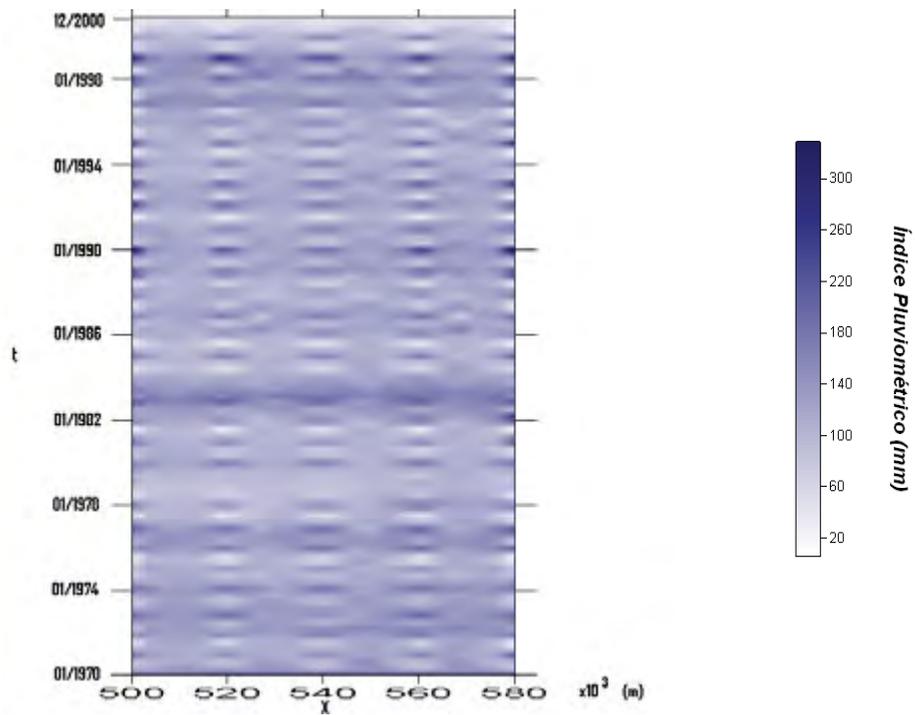


Figura 68: Mapa de superfície espaço-tempo da coordenada Y = 7505000 (Faixa 4).

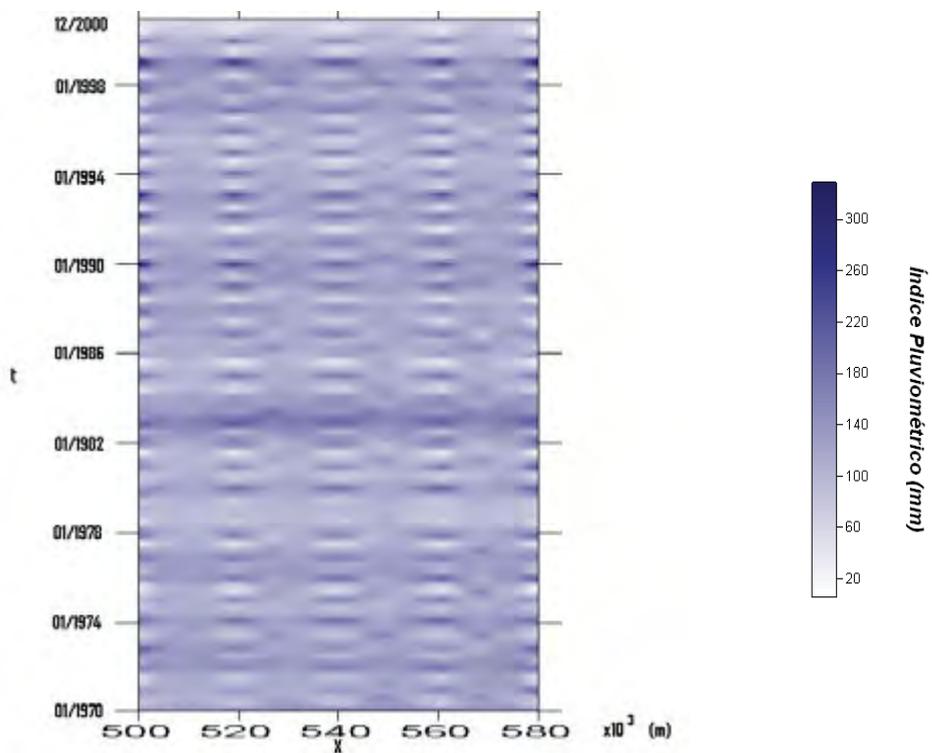


Figura 69: Mapa de superfície espaço-tempo da coordenada Y=7545000 (Faixa 5).

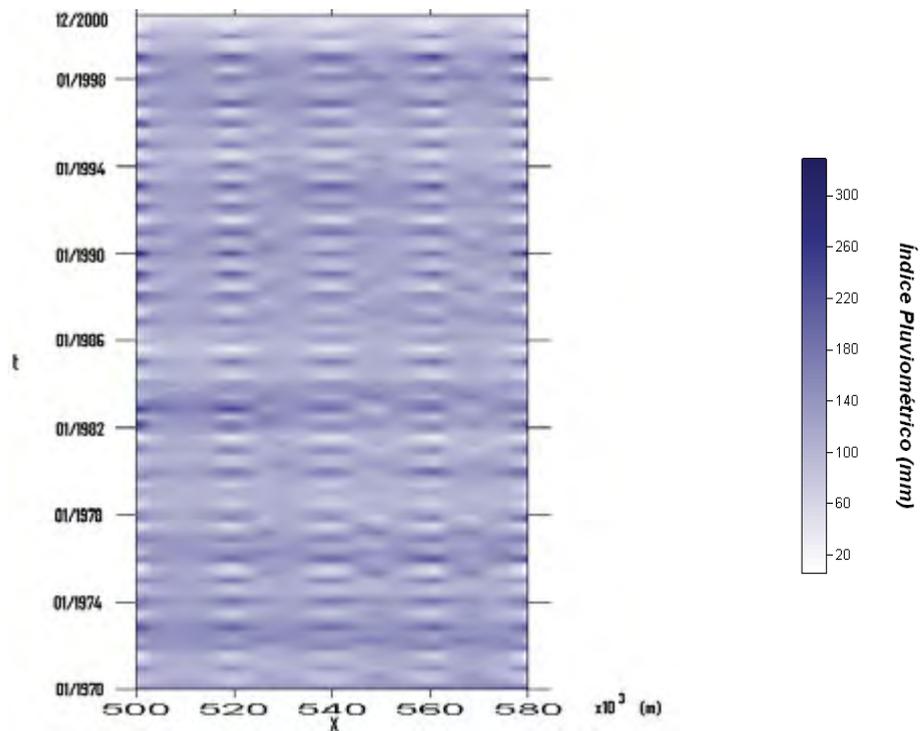


Figura 70: Mapa de superfície espaço-tempo da coordenada Y=7585000 (Faixa 6).

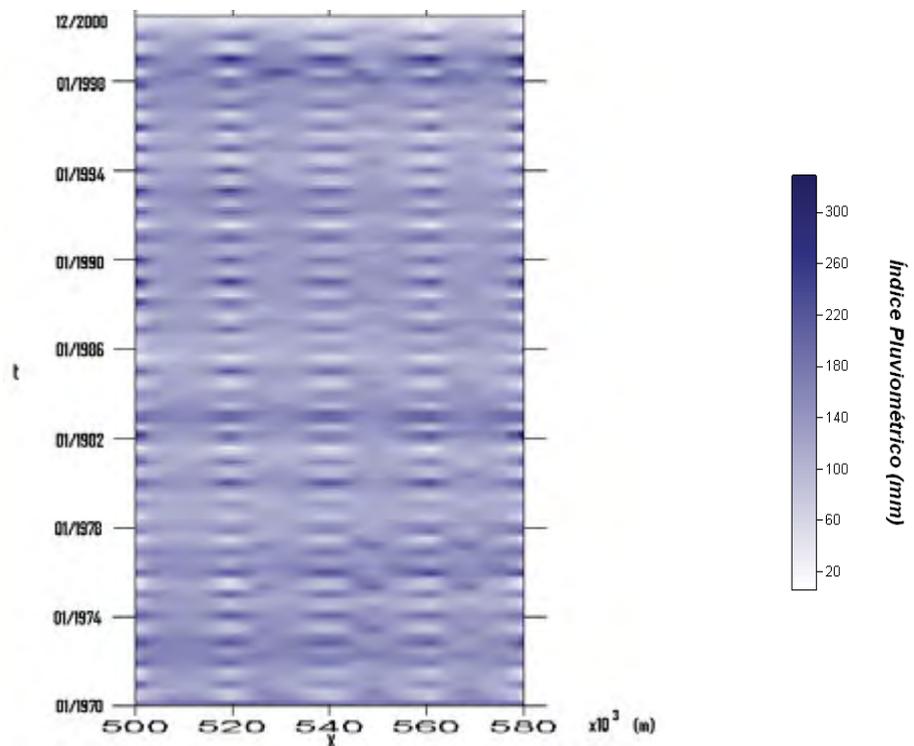


Figura 71: Mapa de superfície espaço-tempo da coordenada Y=7625000 (Faixa 7).

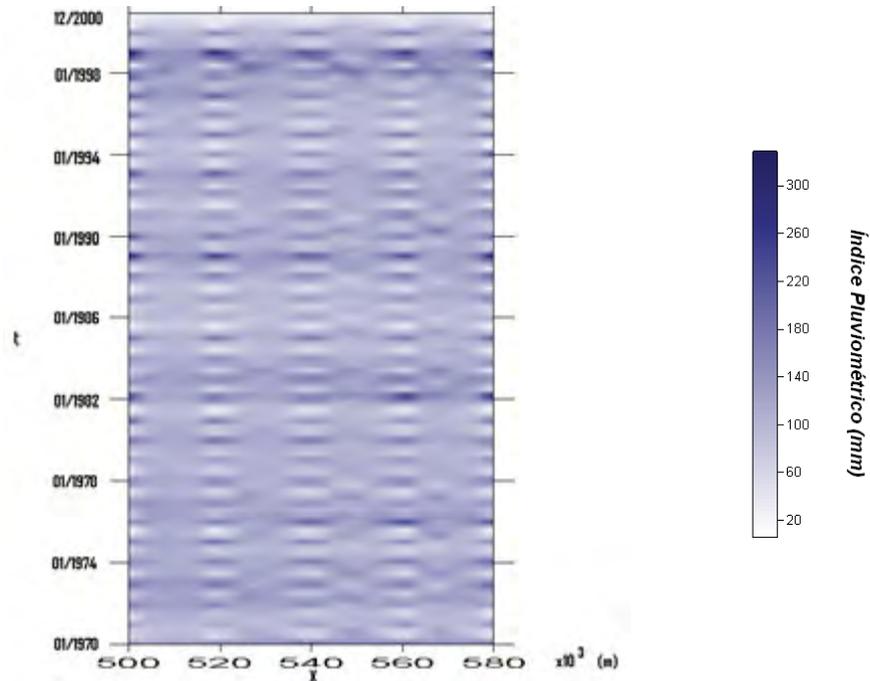


Figura 72: Mapa de superfície espaço-tempo da coordenada Y=7665000 (Faixa 8).

Outra informação que se pode extrair dos mapas de superfície espaço-tempo é a quantidade de chuva. Os anos mais chuvosos, considerando as Faixas 4, 5, 6, 7 e 8, foram os correspondentes a 1983, 1989 e 1998.

### i) geoestatística no espaço

Esta etapa consiste em visualizar o comportamento do regime pluviométrico para todos os 31 anos. Neste caso, foram gerados mapas de superfície espacial através da técnica de interpolação por krigagem ordinária em cada ano. A variável selecionada para a geração dos mapas foi à precipitação total anual que consiste na soma das precipitações totais mensais em cada ano. (Figura 73 a Figura 76)

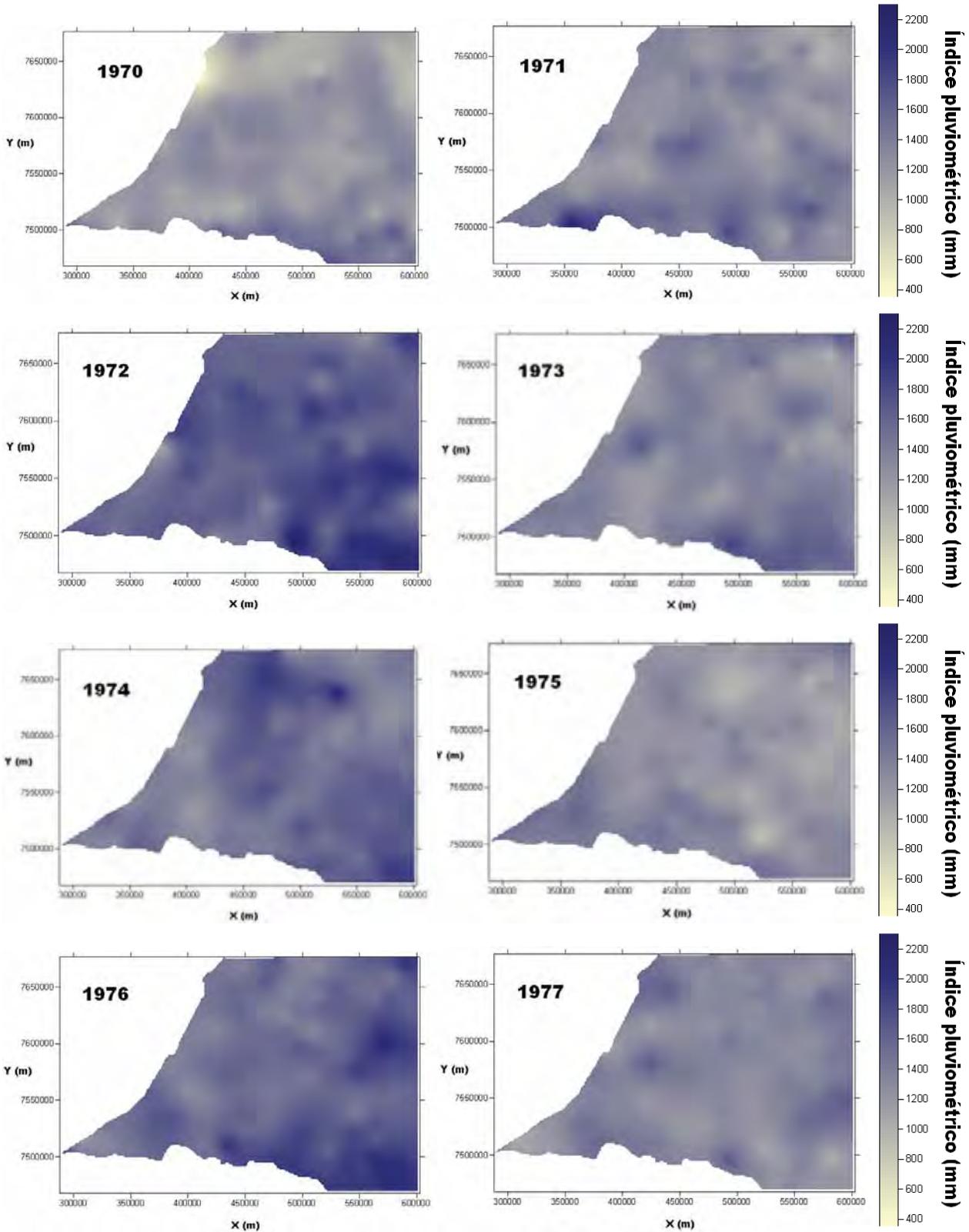


Figura 73: Mapas de superfície do período de 1970 a 1977.

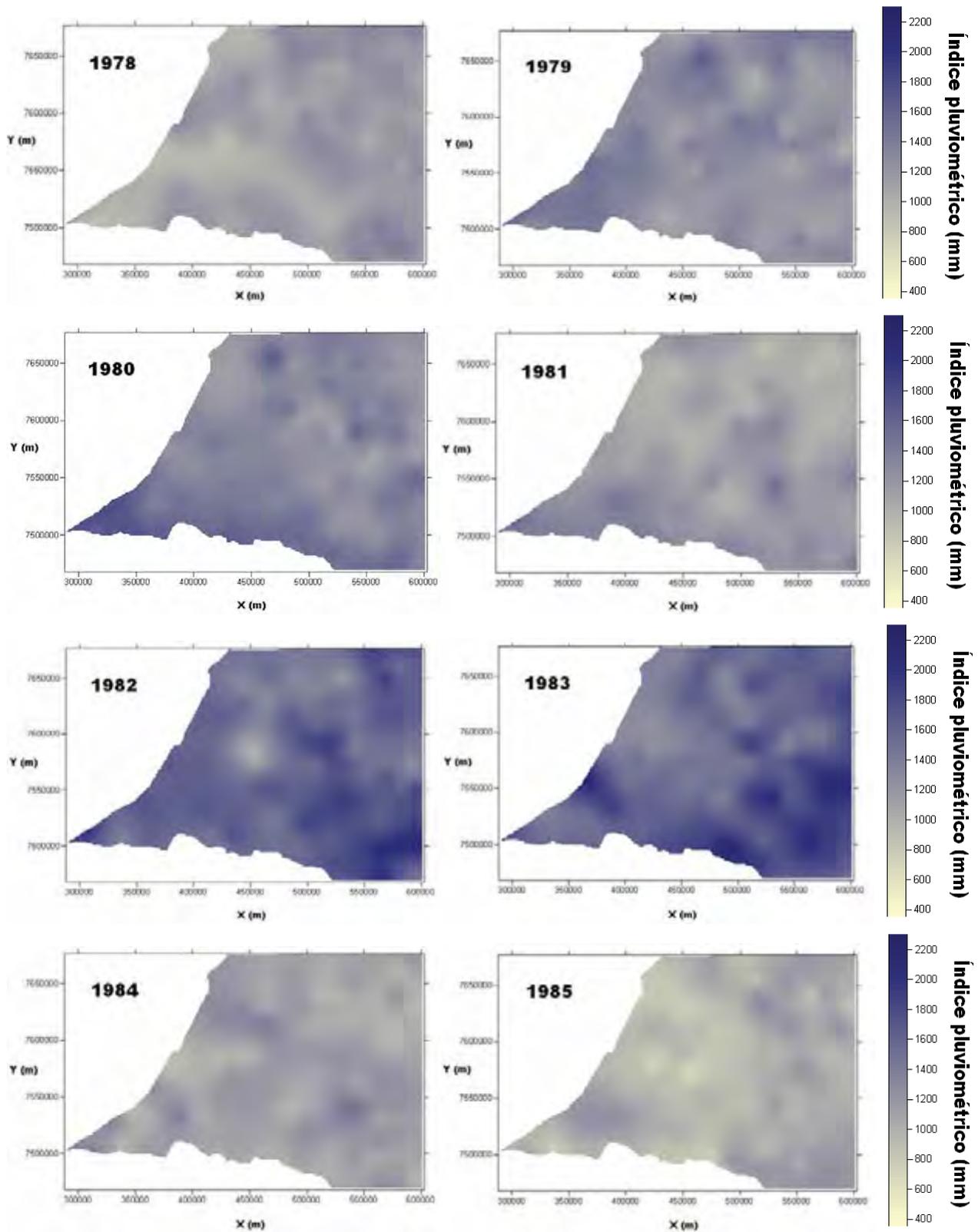


Figura 74: Mapas de superfície do período de 1978 a 1985.

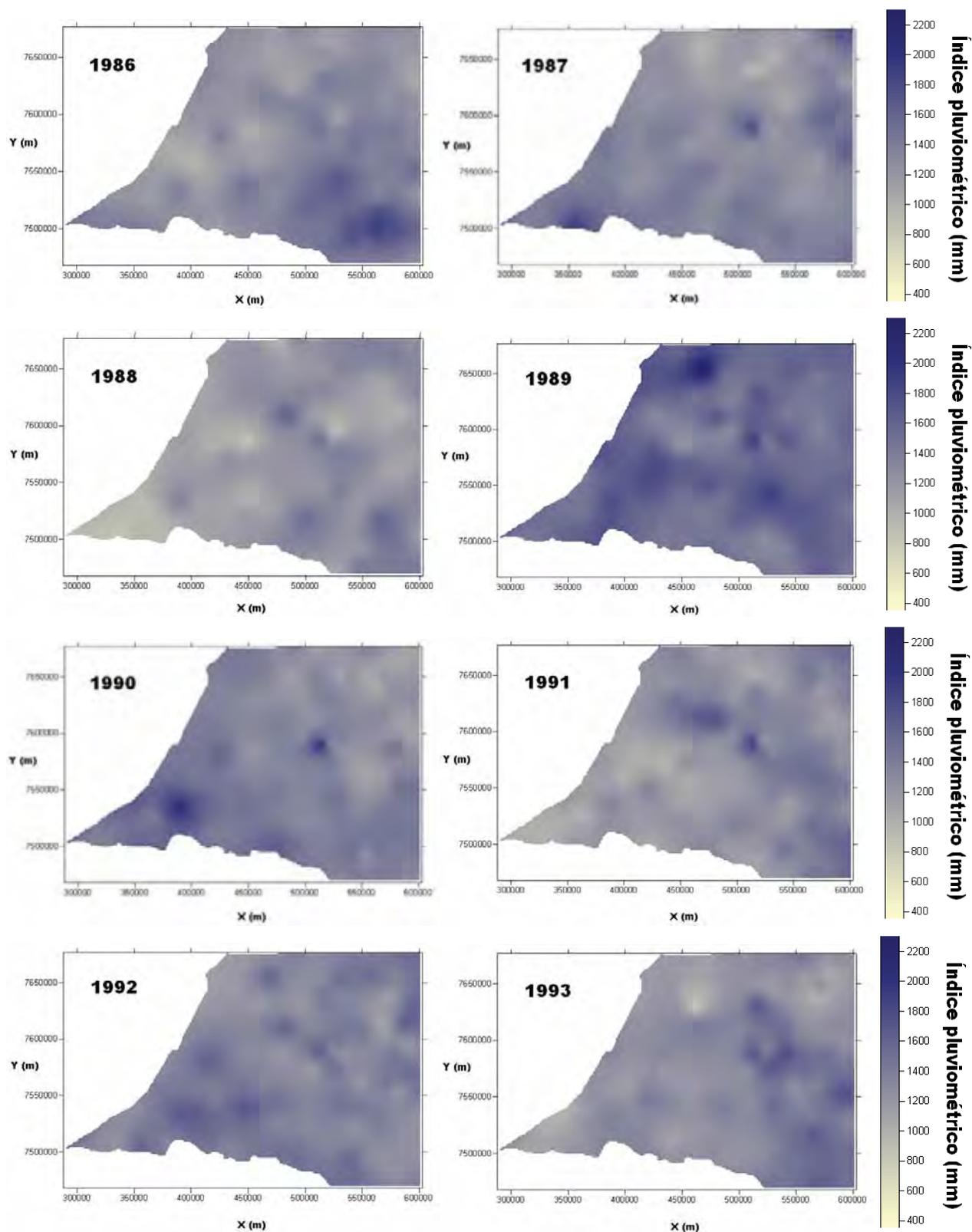


Figura 75: Mapas de superfície do período de 1986 a 1993.

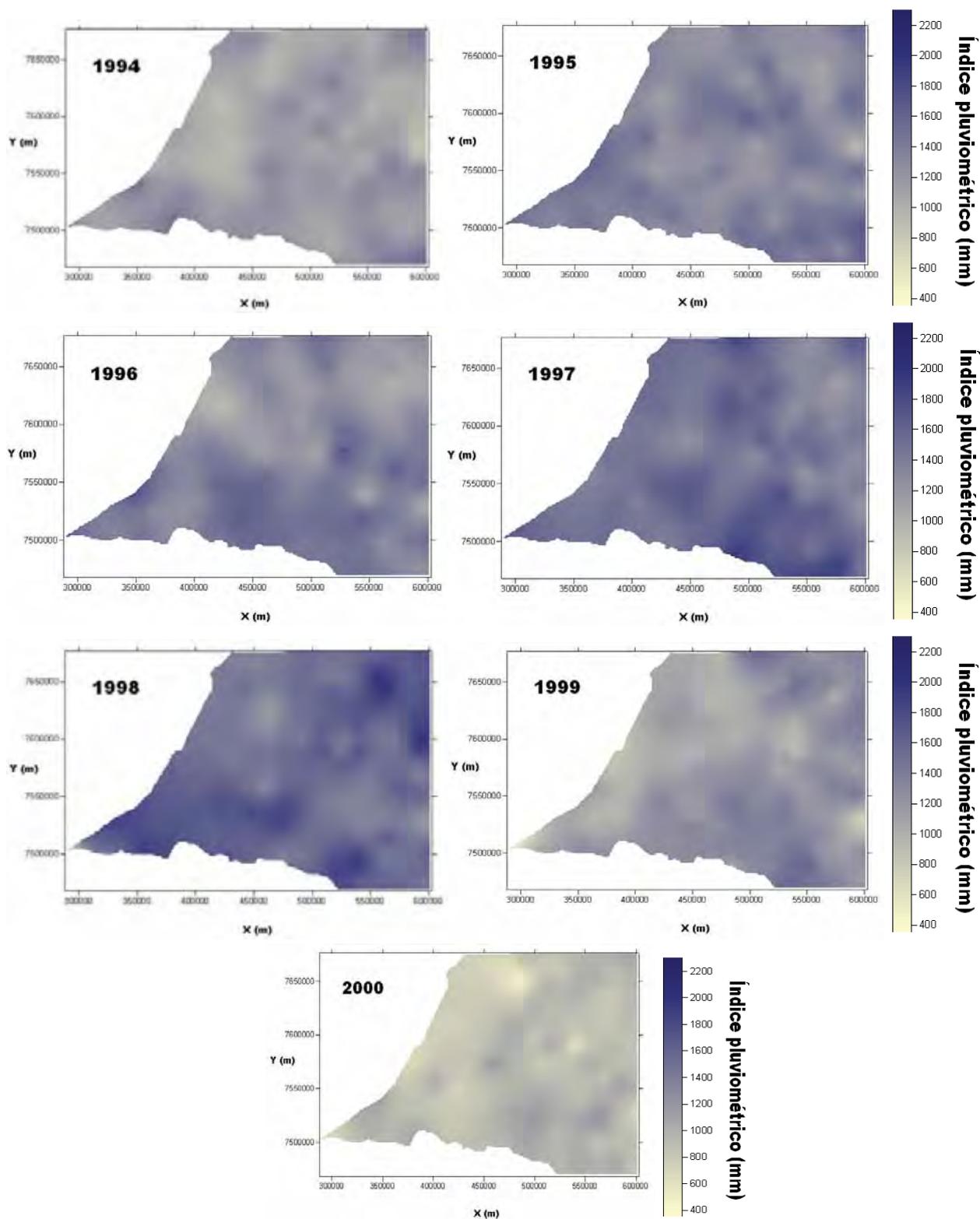


Figura 76: Mapas de superfície do período de 1994 a 2000.

As regiões com maiores índices pluviométricos (em torno de 2200 mm) estão situadas em pequenas regiões no oeste paulista, no decorrer dos anos, essas concentrações migram para regiões vizinhas e retornam para a mesma região ao passar dos anos como no exemplo ilustrado na Figura 77.

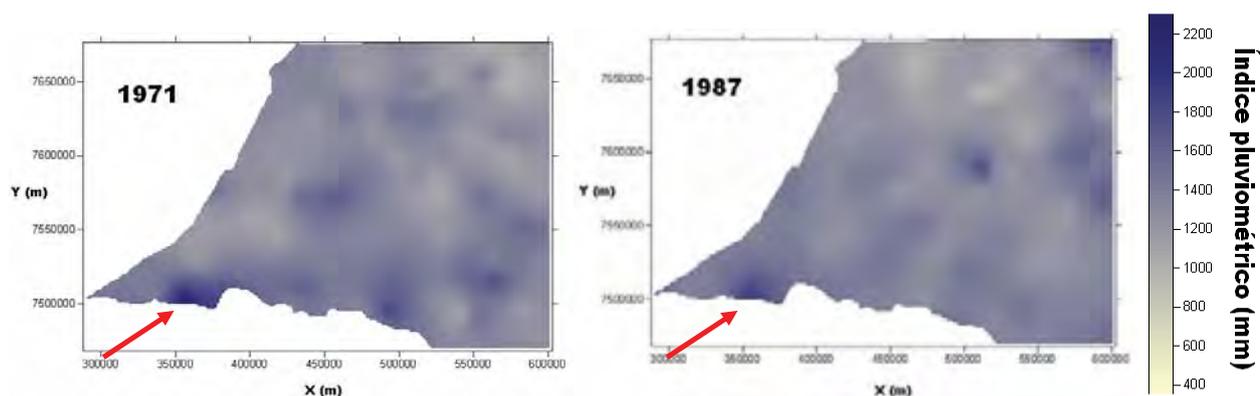


Figura 77: Retorno dos maiores índices pluviométricos na mesma localidade.

As análises dos períodos mais secos e mais chuvosos foram feitas visualmente pelos gráficos Box-plot dos 31 anos representados nas Figuras 78 e 79.

O período mais seco dos 31 anos foi o correspondente ao ano de 2000, pois contém os menores valores de precipitação pluviométrica total. A respeito do período mais chuvoso, foi o correspondente ao ano de 1972.

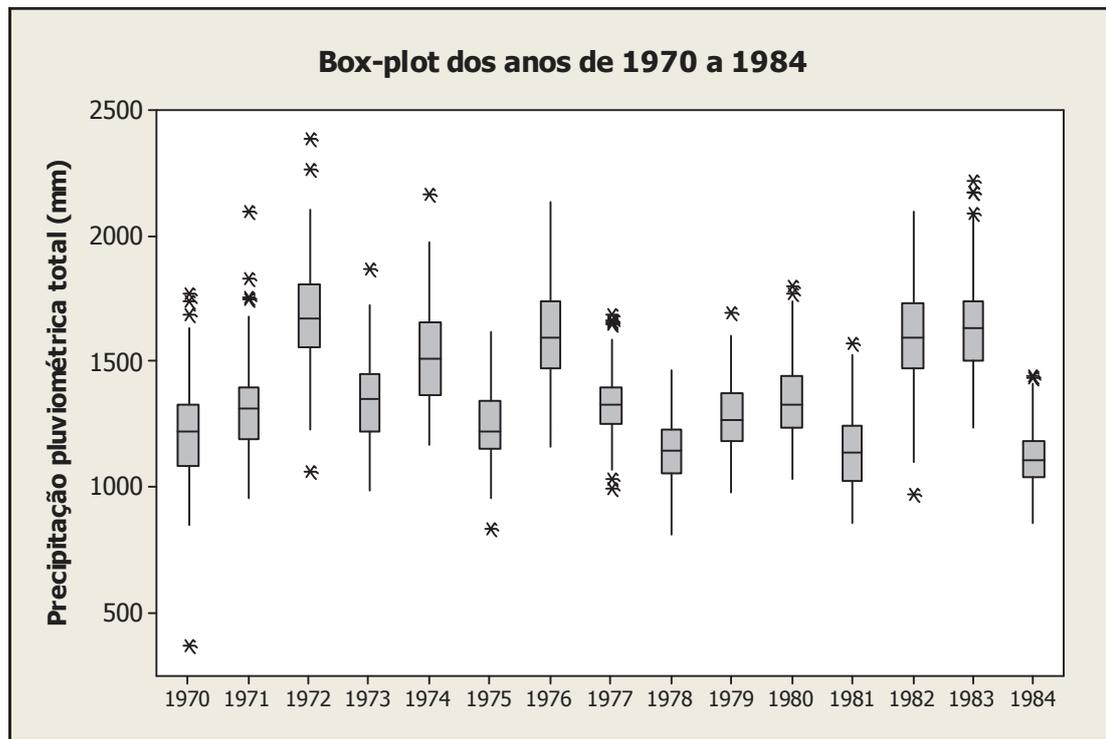


Figura 78: Box-plot da precipitação pluviométrica total para os anos de 1970 a 1984.

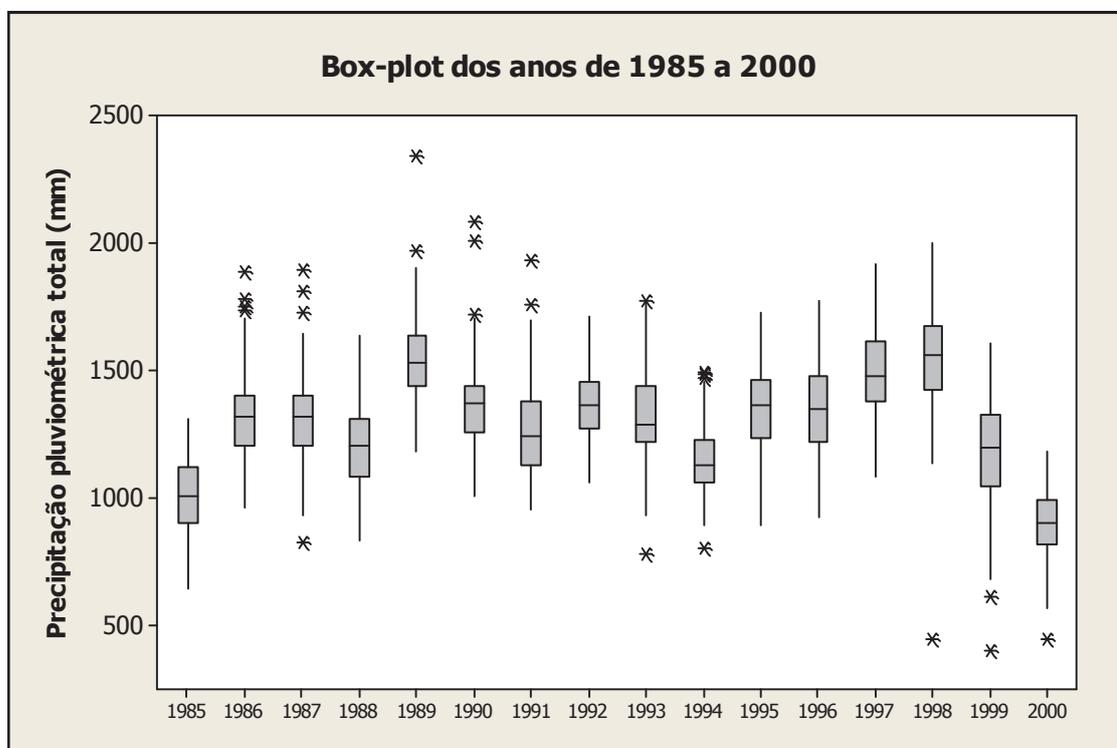


Figura 79: Box-plot da precipitação pluviométrica total para os anos de 1985 a 2000.

#### **4. CONCLUSÃO: Considerações finais, recomendações, trabalhos futuros**

Neste trabalho foram apresentados os principais conceitos de geoestatística clássica como o semivariograma, modelos teóricos, anisotropia e algumas interpolações por krigagem mais usuais como a simples, ordinária, universal e cokrigagem. Com relação à modelagem espaço-temporal, foram apresentados modelos espaço-tempo como o métrico, produto, linear, não-separável, produto-soma, produto-soma generalizado, geoestatístico espaço-tempo e a metodologia Máxima Entropia Bayesiana.

O conjunto de dados analisado refere-se a precipitação pluviométrica da região oeste do Estado de São Paulo no período de 1970 a 2000. Nesse conjunto, inicialmente, fez-se uma análise de agrupamentos, utilizando-se a técnica proposta por Ward para formação de grupos de estações similares e formação de anos similares. Com as estações formou-se 7 grupos homogêneos e para os anos, 3 grupos homogêneos.

Nessa análise foi possível localizar grupos com o mesmo comportamento, como a região mais seca, ao norte da região de Presidente Prudente e região de Araçatuba e a mais chuvosa que está localizada na região de Marília. Os anos mais secos são os correspondentes aos de 1970, 1978, 1981, 1984, 1985, 1986, 1988, 1991, 1993, 1994, 1999 e 2000 e os mais chuvosos são os de 1972, 1974, 1976, 1982, 1983, 1989, 1997 e 1998.

Os resultados apresentados através do método de Ward foram semelhantes aos resultados dos autores Sant'Anna Neto (1995) e Boin (2000) mesmo considerando somente a variável precipitação mensal.

As variáveis correspondentes a precipitação pluviométrica no caso da geoestatística no espaço apresentaram comportamento anisotrópico. As variáveis precipitação máxima mensal e precipitação média mensal foram ajustados modelos somente para o caso isotrópico, pois ocorreram os casos de tendência e ausência de dependência espacial para o caso anisotrópico.

O primeiro modelo espaço-tempo adotado no estudo do regime pluviométrico foi o modelo separável denominado modelo produto-soma generalizado. As componentes espacial e temporal foram geradas a partir da geoestatística clássica aplicadas separadamente no espaço e no tempo. O primeiro modelo espaço-tempo gerado corresponde às variáveis precipitação mensal média no caso espacial e precipitação média no caso temporal. O segundo corresponde as

variáveis precipitação mensal máxima e precipitação máxima, para os casos espacial e temporal respectivamente.

De acordo com o valor  $k$  que pondera a soma e produto dos modelos espacial e temporal, observa-se que o produto desses modelos não tem muita significância, pois o valor  $k$  está próximo de 0, assim o modelo produto-soma generalizado passa a ser considerado apenas como um modelo soma.

A respeito da metodologia geoestatística proposta por Høst, Omre and Switzer (1995) para os dados pluviométricos no oeste do Estado de São Paulo, o modelo geoestatístico espaço-tempo apresentou melhores estimativas para menores índices pluviométricos (período seco). Esse obteve melhores resultados com as precipitações anuais se comparando com as precipitações mensais, como estudado entre precipitação mensal e precipitação anual, apresentado no item c).

Uma das maneiras de comprovar a eficiência das estimativas do modelo geoestatístico espaço-tempo proposto por Høst, Omre and Switzer (1995) foi retirar a estação de coleta mais distante das demais estações e gerar o modelo sem essa estação comparando com os seus valores reais. Outra maneira foi comparando as estimativas com as médias aritméticas referentes às três estações mais próximas de cada estação retirada. Os resultados obtidos foram satisfatórios pois as medidas geradas a partir dessa estimativa apresentaram alta correlação, EAM baixo e a normalidade da distribuição dos erros.

As superfícies espaço-tempo geradas a partir das estimativas do modelo geoestatístico espaço-tempo descreveram o comportamento do regime pluviométrico ao longo do espaço e do tempo. Essas superfícies mostram as menores e maiores concentrações de chuva e tanto no eixo temporal quanto no eixo espacial, o regime pluviométrico descreveu um comportamento periódico.

As principais dificuldades encontradas no trabalho foram:

- elaboração do banco de dados: os dados do DAEE estavam organizados de modo diferente com o utilizado nessa dissertação e precisou ser reorganizado;
- dados ausentes e dados atuais: o banco de dados do DAEE possuíam campos com dados ausentes e não estavam atualizado, pois o maioria dos dados não passavam do ano de 2000;

- aplicação da Máxima Entropia Bayesiana (MEB): devido a poucos trabalhos publicados na área e a falta de dados *soft* optou-se não ser aplicado aos dados pluviométricos;

Ao mesmo tempo que houve a dificuldade em trabalhar com dados ausentes, a metodologia usada auxiliou a sanar este problema.

Algumas sugestões e recomendações para trabalhos futuros são:

- trabalhar com dados mais atuais;
- análise de tendência através de outras técnicas de krigagem, como a Krigagem Universal;
- aplicar outros métodos de modelagem espaço-temporal;
- comparar as estimativas geradas entre as diversas técnicas de modelagem espaço-temporal;
- visualizar os resultados obtidos por krigagem espaço-temporal através de blocos tridimensionais;

A utilização de modelos espaço-tempo pode ser útil para instituições que trabalham com bancos de dados temporal e espacial que contém dados ausentes, pois a utilização desses modelos, assim como o modelo proposto por Høst, Omre and Switzer (1995), auxilia no preenchimento desses dados com boa precisão.

## REFERÊNCIAS

BOGAERT, P., **Spatial Prediction of Categorical Variables: The Bayesian Maximum Entropy Approach**. In: Stochastic Environmental Research and Risk Assessment 16 (2002) 425–448.

BOIN, M. N. **Chuvas e erosões no Oeste Paulista: Uma Análise Climatológica Aplicada**. Rio Claro: UNESP, 2000, 264 p. Tese (Doutorado em Geociências e Meio Ambiente) - IGCE, UNESP, 2000.

CAMARGO, E.C.G. **Desenvolvimento, Implementação e teste de Procedimentos Geostatísticos (Krigagem) no Sistema de Processamento de Informações Georeferenciadas (SPRING)**. 1997. 105 f. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos.

CHANDLER, R. E. et al., *Space-Time Modelling of Rainfall for Continuous Simulation*. Taylor & Francis Group, LLC, 2007.

CHRISTAKOS, G., *Modern Spatiotemporal Geostatistics*. New York: Oxford University Press, 2000, 288p.

CHRISTAKOS, G., VYAS, V. M., **A composite spatiotemporal study of ozone distribution over eastern United States**. Atmospheric Environment, v.32, p.2845– 2857, 1998.

De IACO, S., MYERS, D. E., POSA, D. *Space-Time variograms and a functional form for total air pollution measurements*. Computational Statistics & Data Analysis, 41, 311-328, 2002.

DEUTSCH, C. V. & JOURNAL, A. G., **GSLIB-Geostatistical Software Library and Users' Guide**. Oxford University Press, 1998.

DOUAIK, A., et al., *Soil salinity using spatio-temporal kriging and Bayesian maximum entropy with interval soft data*. Geoderma 128 (2005) 234-248. Disponível em <www.sciencedirect.com>. Acesso 10 jun. 2009.

FLORES, E.F. **Modelagem em Climatologia Geográfica: Um Ensaio Metodológico Aplicado ao Oeste Paulista**. 2000. 237 f. Tese (Doutorado em Geociências e Meio Ambiente) – Universidade do Estadual Paulista, Rio Claro.

GOOVAERTS, P., SONNET, P., *Study of spatial temporal variation of hydrogeochemical variables using factorial kriging analysis*. In: Soares, A., ed. Geostatistical Tróia 92. Dordrecht: Kluwer, 1993. v.2, p 745-756.

GNEITING, T. et al. , *Geostatistical Space-Time Models, Stationarity, Separability, and Full Symmetry*. Taylor & Francis Group, LLC, 2007.

HOOSBEEK, M. R., **Incorporating scale into spatio-temporal variability: applications to soil quality and yield data**. Geoderma, v.85, p.113– 31, 1998.

HØST, G., OMRE, H., SWITZER, P., *Spatial interpolation errors for monitoring data*. In: Jour. Am. Statistical Assoc., V. 90, no. 431, p. 853-861. 1995.

ISAAKS, E. H., SRIVASTAVA R. M., *An introduction to applied Geostatistics*. New York, Oxford University. 1989. 560p.

JOHNSON, R. A., WICHERN, D. W. *Applied Multivariate Statistical Analysis*, 6<sup>th</sup> ed. Upper Saddle River, NJ : Pearson Prentice Hall, 2007

JOURNEL, A.G. *Fundamentals of Geostatistics in Five Lessons*. Stanford: Center of Reservoir Forecasting Applied Earth Sciences Department, 1988.

KYRIAKIDIS, P.J., JOURNEL, A. G. *Geostatistical space-time models: a review*. Mathematical Geology, 1999.

LANDIM, P.M.B. **Análise Estatística de dados Geológicos**. São Paulo: Editora da UNESP, 1998.

LIMA, D. L., **Sensoriamento Remoto e Geoestatística na Caracterização Espaço-Temporal de Plantas Aquáticas**. 2006. 99 f. Dissertação (Mestrado em Ciências Cartográficas) – Faculdade de Ciências e Tecnologia, UNESP, Presidente Prudente.

MONEY, E. S., et al., *Modern Space/Time Geoestatics Using River Distances: Data Integration of Turbidity and E. Coli Measurements to Assess Fecal Contamination Along the Raritan River in New Jersey*. In: Environ.Sci.Technol. 2009, 43, 3736–3742.

MONTEIRO, C. A. F. **Análise Rítmica em Climatologia**. São Paulo: USP/IG, 1977. 21 p. (Série Climatologia, 1)

MOTA, V. C., **Metodologia Geoestatística para Caracterizar a Variabilidade Temporal de Elementos Climáticos de Juiz de Fora – MG**. 2008. 122 f. Dissertação (Mestrado em Estatística e Experimentação Pecuária) - UFLA, Lavras, MG.

RIBEIRO, R. L. F., MINGOTI, S. A. **Apresentando um programa computacional em R e Geo-R para análise espaço-temporal via geoestatística**. In: XIII SIMPEP, 2006, Bauru.

RIBEIRO, R. L. F., MINGOTI, S. A. **Modelos geoestatísticos de previsão espaço-temporal: Implementação computacional no software R**. In: XXVII Encontro Nacional de Engenharia de Produção, 2007, Foz do Iguaçu.

ROUHANI, S., WACKERNAGEL, H., *Multivariate geostatistical approach to space-time data analysis*. Water Resources Research, v. 26, n. 4, p. 585–591, 1990.

SANT'ANNA NETO, J.L. **As chuvas no Estado de São Paulo: Contribuição ao estudo da variabilidade e tendência da pluviosidade na perspectiva da análise geográfica**. São Paulo: USP, 1995. 251 p. Tese (Doutorado em Geografia – Geografia Física) – FFLCH, USP, 1995.

SIQUEIRA, H. R. et al., **Comportamento da Precipitação Pluviométrica Mensal do Estado de Minas Gerais: Análise Espacial e Temporal**. Revista Horizonte Científico, 2007.

SOARES, A., PATINHA, P. J., and PEREIRA, M. J. *Stochastic simulation of space-time series: Application to a river water quality modeling*, In: SRIVASTAVA, R. M., ROUHANI, S., CROMER, M. V., JOHNSON, A. I., and DESBARATS, A. J., eds., **Geostatistics for environmental and geotechnical applications**. American Society for Testing and Materials, West Conshohocken, PA, p. 146–161, 1996.

SNEPVANGERS, J. J. J. C., HEUVELINK, G. B. M., HUISMAN, J. A., *Soil water content interpolation using spatio-temporal kriging with external drift*. Geoderma, v.112, p.253-71, 2003.

VALERIANO, M. M.; SANTOS, A. O. *Mapping of plant-soil data with high random variability*. Acta Scientiarum, v. 24, n. 5, 2002. p. 921-928.