

ET-11



UNIVERSIDADE EDUARDO MONDLANE
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA E INFORMÁTICA
Curso de Licenciatura em Informática

Trabalho de Licenciatura

**Geoestatística como Método Automático de Previsão em Séries Temporais:
uma Comparação com os Métodos de Holt-Winters e Box-Jenkins**

Supervisor: dr. Rodrigues Zicai Fazenda

Estudante: Alberto Santos Capece

Maputo, Março de 2007

R.E. 10.959

Caricido

UNIVERSIDADE EDUARDO MONDLANE
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA E INFORMÁTICA
Curso de Licenciatura em Informática

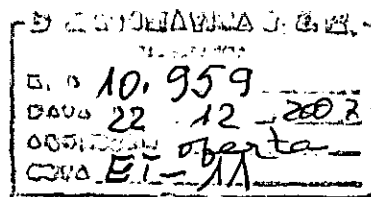
Trabalho de Licenciatura

Tema:

Geoestatística como Método Automático de Previsão em Séries Temporais:
uma Comparação com os Métodos de Holt-Winters e Box-Jenkins

Supervisor: dr. Rodrigues Zicai Fazenda

Estudante: Alberto Santos Capece



Maputo, Março de 2007

DEDICATÓRIA

Aos meus pais,
Santos Capece e
Lídia Greia Castiano.

AGRADECIMENTOS

Uns consideram que escrever um trabalho de licenciatura é como enfrentar um leão, mas, na verdade, os meses que passei para escrever este trabalho foram mais como uma alcateia inteira de leões. Por sorte, não tive de enfrentá-los sozinho. Houve muitas pessoas que contribuíram de inúmeras maneiras para a realização deste trabalho. Eu gostaria de agradecer a cada uma delas.

A Deus, a quem devo minha maior gratidão: *a vida!*

Ao dr. Rodrigues Zicai Fazenda pelos ensinamentos, orientação e motivação, bem como na confiança que depositou em mim durante o desenvolvimento deste trabalho.

Aos meus pais pelo apoio, carinho e incentivo em todas as horas. Meus pais muitas vezes se sacrificaram em prol de minha educação, sem eles não seria o que sou hoje.

Agradeço a Sra. Ivone Maria Armando que soube dizer “fica comigo” no momento certo.

Aos Bancos de Sangue dos Hospitais Central de Maputo e Geral José Macamo, por proporcionarem condições para a realização desta pesquisa, e em especial o Dr. Joel, a Sra. Helena Mendes e a Sra. Luísa Pombuane pela concessão de dados necessários ao estudo.

Agradeço ainda a Sueli Aparecida Mingoti pelos artigos científicos concedidos.

Ao dr. Valentim Raposo que sempre mostrou ser alguém disposto para me ajudar.

Aos meus irmãos, pelo apoio imensurável que me dão, vai o meu muito obrigado.

Finalmente, agradeço a todas as pessoas que, de alguma forma, contribuíram para que chegasse onde estou; seria impossível citar todas.

Muito obrigado pelo apoio de todos...

DECLARAÇÃO DE HONRA

Declaro por minha honra, que este trabalho é resultado da minha investigação, e que não foi submetido para outro grau que não seja o indicado, “Licenciatura em Informática”, pela Universidade Eduardo Mondlane.

Maputo, Março de 2007

O Estudante

Alberto Santos Capece
(Alberto Santos Capece)

RESUMO

A análise de séries temporais permite identificar e descrever o comportamento de uma variável no tempo. O presente trabalho tem como objectivo comparar os métodos aplicados para prever o comportamento de variáveis avaliadas no tempo, nomeadamente: Geoestatística, Box-Jenkins e Holt-Winters.

Para alcançar o objectivo proposto, inicialmente, fez-se o estudo dos principais aspectos teóricos sobre os métodos em comparação neste trabalho. Em seguida foi desenvolvido o processo de estimação dos modelos dos métodos propostos aplicados às séries históricas disponibilizadas pelos Bancos de Sangue dos Hospitais Central de Maputo e Geral José Macamo no que concerne à unidades de sangue doadas e índice de prevalência de HIV. Posterior a isso, empregou-se uma estratégia de teste e selecção de modelos mais fiáveis e consequentes métodos para um horizonte de previsão de 12 (doze) meses, de natureza quantitativa, para subsidiar as conclusões apresentadas neste trabalho.

Os resultados obtidos usando esta estratégia dão maior vantagem ao método automático de geoestatística. E em todos os casos analisados o método Holt-Winters mostrou ser o segundo classificado no ranking de potencialidade de previsão. Alias os métodos ARIMA posicionaram-se em último lugar excepto na série ^{de?} unidades de sangue infectadas pelo vírus HIV colectadas no BSHGJM (Banco de Sangue do Hospital Geral José Macamo) onde mostrou ser o melhor método.

Não sendo palavra portuguesa, parece-me ser mais ideal escrevê-la em itálico ou entre vírgulas altas.
(Para quem não saiba Inglês, o que significa ranking?)

LISTA DE TABELAS

Tabela 01.	Potencialidades de previsão do EPAM	35
Tabela 02.	Estatística descritiva das séries	38
Tabela 03.	Coeficientes de previsão do modelo de vizinhança 11 da série de unidades de sangue dos dadores voluntários do BSHCM	39
Tabela 04.	Coeficientes de previsão do modelo de vizinhança 16 da série de unidades de sangue dos dadores voluntários do BSHGJM	40
Tabela 05.	Coeficientes de previsão do modelo de vizinhança 12 da série de unidades de sangue dos dadores repositores do BSHCM	40
Tabela 06.	Coeficientes de previsão do modelo de vizinhança 3 da série de unidades de sangue dos dadores repositores do BSHGJM	40
Tabela 07.	Coeficientes de previsão do modelo de vizinhança 41 da série de unidades de sangue infectadas por HIV do BSHCM	40
Tabela 08.	Coeficientes de previsão do modelo de vizinhança 30 da série de unidades de sangue infectadas por HIV do BSHGJM	41
Tabela 09.	Valores dos Coeficientes do método Holt-Winters para as séries do BSHCM ...	41
Tabela 10.	Valores dos Coeficientes do método Holt-Winters para as séries do BSHGJM .	42
Tabela 11.	Parte do teste ADF para a série de unidades de sangue dos voluntários do BSHCM	43
Tabela 12.	Parte do teste ADF para a série de unidades de sangue dos voluntários do BSHGJM	43
Tabela 13.	Parte do teste ADF para a série de unidades de sangue dos dadores repositores do BSHCM	45
Tabela 14.	Parte do teste ADF para a série de unidades de sangue dos dadores repositores do BSHGJM	46
Tabela 15.	Parte do teste ADF para a série de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHCM	47
Tabela 16.	Parte do teste ADF para a série de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHGJM	48
Tabela 17.	Parte do teste ADF para a série transformada de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHCM	48
Tabela 18.	Comparação entre os diferentes modelos aplicados a série da figura 1a)	50
Tabela 19.	Comparação entre os diferentes modelos aplicados a série da figura 1b)	52
Tabela 20.	Comparação entre os diferentes modelos aplicados a série da figura 2a)	53

Tabela 21.	Comparação entre os diferentes modelos aplicados a série da figura 2b)	53
Tabela 22.	Comparação entre os diferentes modelos aplicados a série de número de unidades de sangue colectados no BSHCM	54
Tabela 23.	Comparação entre os diferentes modelos aplicados a série de número de unidades de sangue colectados no BSHGJM	55
Tabela 24.	Comparação entre os diferentes modelos aplicados a série da figura 3a)	56
Tabela 25.	Comparação entre os diferentes modelos aplicados a série da figura 3b)	57

LISTA DE FIGURAS

Figura 01.	Unidades de sangue doadas por doadores Voluntários	38
Figura 02.	Unidades de sangue doadas por doadores Repositores	38
Figura 03.	Unidades de sangue infectadas por HIV	39
Figura 04.	a) função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades dos doadores voluntários do BSHCM	43
Figura 05.	a) função de auto-correlação e b) função de auto-correlação parcial para o número de unidades dos doadores voluntários do BSHGJM	43
Figura 06.	a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo ARMA(3,0,2)	44
Figura 07.	a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo SARIMA(0,0,0)(1,0,0) ₁₂	44
Figura 08.	a) Função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades dos doadores Repositores do BSHCM	45
Figura 09.	a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo SARIMA(0,1,0)(1,1,0) ₁₂	45
Figura 10.	Função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades dos doadores Repositores do BSHGJM	46
Figura 11.	a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo ARIMA(6,0,0)	46
Figura 12.	a) Função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades de sangue inutilizadas por infecção com o vírus do HIV do BSHCM	47
Figura 13.	a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo SARIMA (1,0,0)(1,1,0) ₁₂	47
Figura 14.	a) Função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHGJM	48
Figura 15.	Gráfico da série da figura 3b) a pós a sua primeira diferença	48
Figura 16.	a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo ARIMA (2,1,0)	49
Figura 17.	Previsões para as unidades de sangue dos voluntários do BSHCM	51
Figura 18.	Previsões para as unidades de sangue dos voluntários do BSHGJM	52
Figura 19.	Previsões para as unidades de sangue dos repositores do BSHCM	53

Figura 20.	Previsões para as unidades de sangue dos repositores do BSHGJM	54
Figura 21.	Previsões para as unidades de sangue colectados no BSHCM	55
Figura 22.	Previsões para as unidades de Sangue Colectados no BSHGJM	55
Figura 23.	Previsões para as unidades de sangue das unidades de sangue infectados pelo vírus HIV do BSHCM	56
Figura 24.	Previsões para as unidades de sangue das unidades de sangue infectados pelo vírus HIV do BSHGJM	57

LISTA DE ABREVIATURAS E SIGLAS

ACF	– Autocorrelation Function
ADF	– Augmented Dickey-Fuller
AIC	– Akaike Information Criterion
AR	– Auto-Regressão
ARIMA	– Auto-Regressive Integrated Moving Average
ARMA	– Auto-Regressive Moving Average
B	– Backward shift operator
BJ	– Box-Jenkins
BSHCM	– Banco de Sangue do Hospital Central de Maputo
BSHGJM	– Banco de Sangue do Hospital Geral José Macamo
EAM	– Erro Absoluto Médio
EM	– Erro Médio
EPAM	– Erro Parcial Absoluto Médio
EQM	– Erro Quadrático Médio
GEO	– Geoestatística
HCM	– Hospital Central de Maputo
HGJM	– Hospital Geral José Macamo
HIV	– Vírus de Imunodeficiência Adquirida
HWA	– Holt-Winters Aditivo
HWM	– Holt-Winters Multiplicativo
I	– Integrated
MA	– Moving Average
MS	– Microsoft
PACF	– Partial Autocorrelation Function
PREVGEO	– Software para fazer previsões em séries temporais via geoestatística
SAR	– Seasonal Auto-Regressive
SARIMA	– Seasonal Auto-Regressive Integrated Moving Average
SBC	– Schwartz Bayesian Criterion
SITRAM	– Software usado no Banco de Sangue do Hospital Central de Maputo
SMA	– Seasonal Moving Average
SOLVER	– Software de optimização contido no Microsoft Excel.
U	– Estatística U de Theil

ÍNDICE

Resumo	iv
Lista de Tabelas	v
Lista de Figuras	vii
Lista de Abreviaturas e Siglas	ix
Capítulo I: INTRODUÇÃO	1
1. Introdução	1
2. Definição do Problema	2
3. Pergunta de Pesquisa	2
4. Objectivos	2
4.1 Objectivo Geral	2
4.2 Objectivos Específicos	3
5. Fronteira	3
Capítulo II: MATERIAL E MÉTODOS	4
1. Material	4
2. Métodos	4
Capítulo III: REVISÃO DE LITERATURA	6
1. Séries Temporais	6
1.1 Tendência	7
1.2 Sazonalidade	7
2. Objectivos do Estudo das Séries Temporais	7
3. Previsão de Séries Temporais	8
4. Métodos de Previsão de Séries Temporais	9
4.1 Método Trivial de Previsão	9
4.2 Metodologia de Geoestatística	10
4.3 Metodologia de Geoestatística para Séries Temporais	11
4.3.1 Variograma	12
4.3.2 Previsões	12
4.4 Métodos de Suavização (Alisamento) Exponencial	13
4.4.1 Métodos de Suavização Exponencial Simples	14

4.4.2 Método Holt	15
4.4.3 Método de Holt-Winters	15
4.4.3.1 Método Holt-Winters Multiplicativo	16
4.4.3.2 Método Holt-Winters Aditivo	17
4.5 Método de Box-Jenkins	18
4.5.1 Modelos Estacionários	19
4.5.1.1 Modelo Auto-Regressivo (AR)	19
4.5.1.2 Modelos de Média Móvel (MA)	21
4.5.1.3 Modelos Auto-Regressivos de Médias Móveis (ARMA)	23
4.5.2 Modelos Não-Estacionários	23
4.5.2.1 Modelo Auto-Regressivos Integrados de Médias Móveis (ARIMA)	24
4.5.2.2 Modelos Sazonais	25
4.5.3 Coeficiente de Auto-correlação	26
4.5.4 Raiz Unitária e Teste ADF	27
4.5.5 Etapas da Metodologia Box-Jenkins	28
5. Uma Discussão sobre a Comparação dos Métodos de Previsão	31
6. Critérios de Comparação de Qualidades de Previsões	32
Capítulo IV: RESULTADOS E SUA ANÁLISE	37
1. Análise do Comportamento das Variáveis Respostas	37
1.1 Modelação via Geoestatística	39
1.2 Modelação por Método Holt-Winters	41
1.3 Modelação por Método de Box-Jenkins	42
1.3.1 Modelação das Séries de Unidades de Sangue dos Voluntários	43
1.3.2 Modelação das Séries de Unidades de Sangue dos Repositores	45
1.3.3 Modelação das Séries de Unidades de Sangue Infectadas por HIV	47
Capítulo V: DISCUSSÃO DOS RESULTADOS	50
1. Unidades de Sangue dos Voluntários do BSHCM	50
2. Unidades de Sangue dos Dadores Voluntários do BSHGJM	51
3. Unidades de Sangue dos Dadores Repositores do BSHCM	52
4. Unidades de Dadores Repositores do BSHGJM	53
5. Número de Unidades de Sangue Colectados no BSHCM	54
6. Número de Unidades de Sangue Colectados no BSHGJM	55
7. Número de Unidades de Sangue Infectados pelo Vírus HIV Colectados no BSHCM	56

8. Número de Unidades de Sangue Infectados pelo Vírus HIV Colectados no BSHGJM	57
Capítulo VI: CONCLUSÕES E RECOMENDAÇÕES	58
1. Conclusões	58
2. Recomendações	59
Capítulo VII: BIBLIOGRAFIA	60
Anexo A. Valores Reais	63
Anexo B. Resultados de Análise da Consistência Interna dos Dados	66
Anexo C. Valores Previstos	67
Anexo D. Dados Brutos	70

CAPÍTULO I

Introdução

1. Introdução

Actualmente a tomada de decisão em algumas áreas científicas tem sido influenciada, de algum modo, pela rapidez com que as mudanças ocorrem e pelo volume crescente de informação, o que conduz a situações de incerteza cada vez maiores.

Os métodos estatísticos desempenham um papel importante em quase todas as fases da pesquisa humana (Spiegel, 1994). No que diz respeito à previsão, Martins e Laugeni (1999, p.173) citados por Lopes (2002), afirmam tratar-se de um processo metodológico para a determinação de dados futuros baseados em modelos estatísticos, matemáticos ou econométricos ou ainda em modelos subjectivos apoiados numa metodologia de trabalho clara e previamente definida.

Este desejo de compreender o passado e prever o futuro impulsiona a procura de leis que expliquem o comportamento de determinados fenómenos ou acontecimentos (Cortez, 1997). No caso da previsão de séries temporais, é necessário usar métodos capazes de suportar e oferecer uma precisão exaustiva da previsão do futuro. Estes métodos têm-se tornando cada vez mais automáticos, facilitando e flexibilizando, deste modo, o processo de previsão de futuras ocorrências. A utilização desses métodos varia de acordo com as características das séries temporais a que se aplicam. Segundo Breseghello (2005) a elaboração de um sistema de previsão requer organização, conhecimento e habilidade em quatro áreas básicas:

1. Identificação e definição dos problemas a serem tratados (ou previstos);
2. Aplicação dos métodos de previsão;
3. Procedimentos para selecção do método apropriado a situações específicas;
4. Suporte organizacional para adaptar e usar os métodos de previsão requeridos.

As previsões devem caracterizar-se pela precisão dos seus resultados, pela simplicidade dos métodos usados e, sobretudo, pela confiabilidade estatística dos modelos empregues para gerar as tais previsões (Melo, 2001). Ao longo das últimas décadas, vários métodos convencionais para a previsão de séries temporais foram desenvolvidos.

O presente trabalho tem como objectivo a comparação de três métodos de previsão de séries temporais: Geoestatística, Holt-Winters e Box-Jenkins. A comparação será feita aplicando a cada método as seguintes séries temporais reais: *número de unidades de sangue e número de unidades de sangue infectados pelo vírus HIV colectados nos Bancos de Sangue dos Hospitais Central de Maputo e Geral José Macamo.*

2. Definição do Problema

→ Muitos métodos quantitativos têm sido, historicamente, os mais utilizados na previsão de séries temporais. Alguns desses métodos têm apresentado um baixo grau de precisão como resultado da escassa fundamentação teórica dos modelos incorporados nesses métodos e como consequência temos baixa precisão dos resultados oferecidos por estes modelos (Souza, 2005).

→ Os métodos de previsão baseiam-se em opiniões de especialistas (Souza, 2005), sendo vulneráveis à tendências que podem comprometer a confiabilidade de seus resultados.

Muitas vezes, especialistas na área de previsão, deparam-se com a questão que concerne à escolha do método de previsão que pode-se usar para o estudo de um caso, porque à "prior" a série não apresenta características relevantes, que ditam os mecanismos a serem usados no prognóstico dos futuros resultados.

3. Pergunta de Pesquisa

É preciso melhorar a redacção desta pergunta.
É possível, fazendo comparação de diferentes métodos de previsão, e diminuir o nível de dificuldade de escolha de um método fiável para a previsão de uma série temporal?

Talvez colocar uma vírgula

4. Objectivos

Para este trabalho foram definidos os seguintes objectivos:

eliminando a conjunção conjuntiva "e".

4.1 Objectivo Geral

Realizar um estudo comparativo entre três métodos de previsão de séries temporais nomeadamente Geoestatística, Holt-Winters e Box-Jenkins.

4.2 Objectivos Específicos

- Identificar, através da revisão bibliográfica, a diferença entre a Geoestatística e Séries Temporais; *(Em que momento se vão identificar os métodos Holt-Winters e Box-Jenkins - não será na rev. bibliográfica?)*
- Obter a partir da Geoestatística e Séries Temporais os critérios de comparação de métodos de previsão;
- Analisar o comportamento das variáveis respostas;
- Avaliar o desempenho dos métodos de previsão para a identificação do melhor método de previsão.

5. Fronteira

Quanto à fronteira do trabalho, pode-se afirmar que está relacionada, principalmente, à questão de trabalha-se com apenas três métodos de previsão: Geoestatística, Holt-Winters e Box-Jenkins ficando outros métodos como uma sugestão a estudos posteriores. Estes métodos são aplicados única e exclusivamente a dados de doações de sangue em dois Bancos de Sangue pertencentes aos Hospitais Central de Maputo e Geral José Macamo, ambos na cidade de Maputo.

minúsculo → O tamanho temporal usado para a aplicação dos métodos respeitou, inteiramente, a disponibilidade dos dois Bancos de Sangue aquando da colecta de dados para o presente trabalho.

Melhorar esta frase

"...à questão de trabalha-se com..."

(melhorar a frase inteira em rever o uso da forma verbal "trabalha-se" antecedida da preposição "de").

CAPÍTULO II

Material e Métodos

1. Material

Foi usado para esta pesquisa o seguinte material:

- Para modelação por meio da Geoestatística foi utilizado o programa PREVGE0 que funciona como um complemento do popular software estatístico *Minitab for Windows* versão 14.20.
- No processo de modelação através do método Box-Jenkins (identificação, estimação dos parâmetros, verificação e previsão) foi realizado mediante o uso do pacote estatístico SPSS 11.0.
- O processo de modelação por meio da utilização do método Holt-Winters (inicialização, optimização dos parâmetros e previsão) foi realizado mediante a utilização do suplemento SOLVER do software *Microsoft Excel 2003*.

foram?

- Por fim, todos os gráficos foram^{em} construídos mediante o uso do pacote estatístico *Minitab for Windows*, com a excepção dos gráficos comparativos de previsão^{em} que foi usado o *Microsoft Excel 2003*.

2. Métodos

Para a elaboração deste trabalho, começou-se por estudar os três métodos em questão com base na literatura.

Após a revisão da literatura, passou-se à recolha de dados de sangue nos Bancos de Sangue dos Hospitais Central de Maputo e Geral José Macamo.

As séries de dados (*número de unidades de sangue doadas e número de unidades de sangue infectadas pelo vírus HIV*) do Banco de Sangue do Hospitais Central de Maputo (BSHCM) foram fornecidas mediante os relatórios anuais produzidos pelo software SITRAM. E por sua

vez, as do Banco de Sangue do Hospital Geral José Macamo (BSHGJM) foram disponibilizadas em relatórios manuais produzidos pelos técnicos do mesmo banco.

A aplicação dos métodos de previsão propostos no trabalho (Geoestatística, Holt-Winters e Box-Jenkins) foi feita sobre as séries supracitadas de forma individual, visto que cada série tem suas peculiaridades, o comportamento de cada um é distinto no que se refere à tendência e à sazonalidade, exigindo modelos diferentes para suas representações. Aliás, para as séries das unidades de sangue colectadas no BSHCM e BSHGJM foram fragmentadas em quatro (duas para cada), nomeadamente, *unidades de sangue dos dadores voluntários* e *unidades de sangue dos dadores repositores* para explorar de melhor maneira as diferentes características que as séries destes dois grandes grupos de dadores apresentam.

Os dados foram fornecidos em valores mensais de unidades de sangue no período de Janeiro de 2000 a Dezembro de 2005 perfazendo um total de 72 observações, portanto, as previsões serão realizadas também mensalmente. A escolha deste tamanho temporal deveu-se, inteiramente, à disponibilidade dos bancos de sangue do HCM e do HGJM. Esta totalidade dos dados está dentro dos requisitos de alguns métodos de previsão como é o caso de Box-Jenkins que para a sua aplicação necessita de pelo menos 50 observações.

A escolha dessa série temporal para o estudo valerá as instituições, principalmente por dois motivos: a modelação permitirá maior entendimento do comportamento dos dados ao longo do tempo; e os resultados poderão ser assimilados pelas instituições (Bancos de Sangue do HCM e do HGJM), constituindo, deste modo, fonte adicional de informação no suporte às decisões referentes a investimentos e dimensionamento da equipe técnica. Lembrar que este, não é o objectivo central da presente pesquisa.

Para se encontrar a consistência interna dos dados foi calculado o *coeficiente alfa de Cronbach* usando o pacote estatístico *SPSS for Windows* versão 11.0. E como pode ser visto no item 1 do capítulo IV é bastante baixo (é muito inferior a 1) e o desvio padrão para cada série que também é bastante baixo em relação a média.

A avaliação de desempenho dos modelos de métodos de previsão é baseada no resultado do cálculo do EPAM (Erro Percentual Absoluto Médio) e do U de Theil (Coeficiente de Desigualdade) descritos no ponto 6 de análise dos erros de previsão do capítulo de revisão de literatura. O primeiro critério é usado para estimar os modelos e o último é utilizado para determinar o melhor método, isto é, aquele com maior grau de desempenho.

CAPÍTULO III

Revisão de Literatura

1. Séries Temporais

A classe de fenômenos cujo processo de observação e consequente quantificação numérica gera uma sequência de dados distribuídos no tempo é denominada *série temporal* (Souza, 2005). Quando o processo de observações é feito em função do tempo, estamos perante a uma série temporal contínua. Ao passo que, se o processo de observações dos valores é feito em tempos específicos, geralmente equidistantes, a série temporal diz-se discreta e pode ser representada por $z_t = \{z_1, z_2, z_3, \dots, z_n\}$. Cada observação discreta está associada a um instante de tempo distinto, existindo uma relação de dependência serial entre essas observações que possibilita prever os valores futuros da série (Cortez, 1997).

As séries temporais podem exibir até quatro características diferentes no seu comportamento: média (regularidade), tendência, sazonalidade e ciclo (Breseghello, 2005). A característica de *média* observa-se quando os valores da série temporal oscilam na vizinhança de um número constante. A série possui *sazonalidade* no caso em que os padrões cíclicos de variação se repetem em intervalos relativamente constantes no tempo. O *ciclo* de uma série temporal existe quando a série exibe variação crescente ou decrescente, porém, em intervalos de tempo não regulares. Finalmente, a característica de *tendência* ocorre quando a série apresenta comportamento ascendente ou descendente por um longo período de tempo. Todavia, a variação, em séries temporais, que não pode ser explicada pelas características acima, é devido ao *ruído aleatório* ou *erro* que compreende a variabilidade intrínseca dos dados e não pode ser matematicamente modelado.

Das quatro características acima mencionadas, importa referir neste trabalho a tendência e a sazonalidade, sendo elas as principais características que influenciam as observações de uma

série temporal, alias, conforme Cortez (1997) os modelos dos métodos de previsão de séries temporais são criados baseando-se na tendência e/ou na sazonalidade.

1.1 Tendência

Uma parte importante do estudo de uma série temporal consiste na análise da tendência. Ela ocorre quando a série apresenta um comportamento de crescimento ou declínio dos seus valores por um longo período de tempo. Uma série que possui este tipo de comportamento é chamada de *não estacionária* (Cortez, 2002). As mais comuns são: Linear, Quadrática e Exponencial.

1.2 Sazonalidade

A análise de uma série temporal não só se limita ao cálculo da tendência dos seus valores ou de uma expressão analítica que a represente. Com efeito, tão importante como isolar a tendência é isolar os efeitos das flutuações periódicas presentes na série. Em séries temporais componentes de particular importância, quando se pretende fazer previsões a curto prazo, são as flutuações sazonais que se caracterizam pela ocorrência de padrões cíclicos de variação, que se repetem em intervalos relativamente constantes de tempo.

2. Objectivos do Estudo das Séries Temporais

São vários os objectivos do estudo das séries temporais. A seguir são descritos os principais objectivos (Murteira *et. al.*, 1993):

- a) A *descrição* de uma série temporal é um objectivo modesto, mas quase sempre presente, no estudo de uma série e envolve: cálculo de estatísticas sumárias da série; representação gráfica dos dados da série e de funções (empíricas) dos mesmos; e o ajustamento de características da série.
- b) A *modelação* consiste na explicação de uma série como tendo sido gerada a partir de um certo modelo e envolve a estimação de parâmetros e a avaliação da qualidade do diagnóstico, em termos da qualidade estatística e do ajustamento, do(s) modelo(s) proposto(s).
- c) A *previsão* consiste na “estimação” de valores futuros de uma série temporal. A sua importância advém do facto de ser habitualmente vantajoso, dentre tantas áreas, no plano económico, conhecer a evolução provável de uma série no futuro.
- d) O *controlo* consiste na monitorização dos valores da série tendo em vista a detecção de alterações nas características da mesma, as quais podem indiciar que: o modelo proposto não é válido; o modelo que estava a ser usado deixou de ser válido (ex.: a série correspondente a

um índice pode sofrer uma ruptura por mudança do modo de cálculo do índice); e alguns dos parâmetros do modelo mudaram de valor.

De entre estes, a ênfase neste trabalho será dada aos métodos que monitoram o processo de previsões. Segundo Cortez (2002) os modelos de previsão tratam o sistema como se fosse “caixa negra”, não tendo como objectivo determinar os factores que afectam o seu comportamento. Existem, possivelmente, algumas razões para esta postura:

- A primeira deve-se ao facto de o sistema não ser determinístico;
- A segunda reside no facto de se desejar prever o que acontece e não como acontece.

É importante anotar que o analista estatístico deverá retornar à fase de *modelação* sempre que na fase de *controlo* chegue à conclusão que o modelo que estava a ser usado não se ajusta ao conjunto de valores observados da série.

3. Previsão de Séries Temporais

Qual? Uma das técnicas de previsão em uso socorre-se da *previsão de séries temporais*, que se baseia em observações passadas de uma dada variável distribuída no tempo (Cortez, 2002). Não se deve encarar a previsão como uma profecia, mas sim como a melhor maneira de extrapolar as relações existentes nos dados, para fazer previsões (Cortez, 1997). Ao fazermos uma previsão, estamos sujeitos a erros inevitáveis, contudo é objectivo do investigador minimizar os tais erros aplicando técnicas adequadas para a sua análise.

Dependendo do valor assumido pelo *horizonte de previsão*¹ Souza (1989) classifica as previsões dos valores futuros numa série temporal como de curto, médio ou longo prazos. Desse modo, diante da possibilidade de existência de diferentes horizontes de previsão, Refenes (1993) especifica técnicas distintas para prognosticar os valores futuros de uma série temporal, são elas:

- ponto? ou vírgula?*
- a) *Previsão múltiplos passos*: esta abordagem, adoptada para longos horizontes de previsão, procura identificar as tendências gerais e os pontos de inflexão mais relevantes na série temporal. Na previsão múltiplos passos, os valores correntes são empregues na realização da previsão para determinado instante; esta previsão é, então, introduzida entre as observações passadas, compondo, desta forma, um novo conjunto de dados, sobre o qual será obtida a previsão no tempo subsequente;

¹ O *horizonte de previsão* é o número de períodos futuros cobertos pela previsão.

- b) Previsão *simples passo*: nesta técnica não há incorporação de previsões aos dados utilizados para encontrar a próxima previsão, sendo esta, independente dos valores posteriormente previstos. A previsão é feita apenas para o *período* (unidade básica de tempo em que a previsão é requerida) imediatamente anterior ao actual, a partir das observações da série temporal.

A garantia da eficiência das previsões de uma série temporal, somente, é alcançada adoptando como horizonte de previsão, o instante de tempo imediatamente subsequente à origem t (Souza, 1989).

4. Métodos de Previsão de Séries Temporais

Define-se um método de previsão como sendo o conjunto de procedimentos usados no desenvolvimento de uma determinada previsão (Souza, 1989). Para este autor, a maioria dos métodos de previsão de séries temporais se baseiam na suposição de que observações passadas contêm todas as informações sobre o padrão de comportamento da série temporal e esse padrão é recorrente no tempo. Portanto, o propósito dos métodos de previsão consiste em distinguir o padrão de qualquer ruído que possa estar contido nas observações e usar esse padrão para prever os valores futuros da série temporal, assim, pela identificação dessa componente, as previsões para períodos de tempo subsequentes ao observado pode ser desenvolvidas.

Os métodos de previsão variam conforme o grau de sofisticação e fins a que se destinam, estando divididos em duas classes: qualitativos e quantitativos. Os métodos qualitativos preocupam-se em prever o futuro a partir de julgamentos pessoais e subjectivos de quem esteja ^{a fazer} fazendo a previsão, sendo vulneráveis a tendências que podem comprometer a confiabilidade de seus resultados. São utilizados quando não se dispõe de informação sobre o passado. Os métodos quantitativos preocupam-se, basicamente, em ajustar um modelo formalizado de previsão aos dados históricos de forma a projectar os dados futuros.

Os métodos quantitativos são mais eficazes e a sua aceitação generalizada é uma consequência próxima do uso maciço de computadores. Estes são úteis não só pelo seu poder de cálculo mas também pela sua capacidade de armazenamento de dados.

4.1 Método Trivial de Previsão

É o método mais simples de previsão, pois utiliza a última observação ocorrida como previsão do próximo período. Serve principalmente como um método base de comparação com outros modelos mais sofisticados (Wold, 1998). A formulação matemática é dada por:

$$\hat{z}_t = z_{t-1} \quad (1)$$

Onde \hat{z}_t é a previsão para o período t e z_{t-1} é o valor real observado no período $t-1$. A previsão trivial é normalmente utilizada quando a série de dados possui um comportamento altamente imprevisível.

4.2 Metodologia de Geoestatística

Um estudo feito por Daniel G. Krige² usando dados de concentração de ouro, concluiu que somente com a informação dada pela variância seria insuficiente para interpretar o fenômeno em estudo visto que seria necessário levar em consideração a distância entre as observações (Issaks *et. al.* 1989). Para solucionar este problema, surge a *geoestatística*, que leva em consideração a *localização geográfica* e a *dependência espacial* entre os dados. A geoestatística utiliza um modelo probabilístico e a posição espacial das observações, para estudar a variabilidade dos valores observados.

Mais tarde, George Matheron³ baseado nas observações de Krige desenvolveu a *teoria das variáveis regionalizadas*, a partir dos fundamentos da geoestatística. Na teoria das variáveis regionalizadas, $z(x)$ pode ser definida como uma variável aleatória que assume diferentes valores z em função da posição x dentro de uma certa região S , que pode ser unidimensional, bidimensional ou tridimensional, com certa aparência de continuidade. O conjunto de variáveis $z(x)$ medidas em toda a área S pode ser considerado uma função aleatória $Z(x)$, uma vez que são variáveis aleatórias regionalizadas e considera-se que a dependência entre elas é regida por algum mecanismo probabilístico (Armstrong *et. al.*, 1993).

Mesmo tendo sua origem ligada a problemas específicos encontrados na mineração, a geoestatística, pela generalidade de seu desenvolvimento e por lidar com dados distribuídos espacialmente, pode ser utilizada criteriosamente na interpretação e estudo de grande parte dos fenômenos que aparecem nos domínios das ciências da natureza. Em linhas gerais, suponhamos que tem-se uma amostra aleatória de uma variável de interesse z e que esta amostra tenha sido colectada em várias localizações distintas de uma determinada região. Neste caso, considerando-se o comportamento de variação da variável z em relação às coordenadas de localização da amostra, é possível construir um modelo estatístico para prever-se o valor de z para localizações que não foram amostradas previamente. Este modelo incorpora a informação da relação existente entre os valores de z naquelas localizações diferentes via uma função chamada de *Semi-*

² Daniel G. Krige, o pioneiro em introduzir o uso de médias móveis para evitar a super-estimação sistemática de reservas em mineração (Issaks *et. al.* 1989)

³ George Matheron, o pai da teoria das variáveis regionalizadas, concebida e idealizada na década 60.

Variograma e que tem um papel fundamental no procedimento de previsão espacial chamado de *Kriging*.

Oliveira (2002) adianta que a geoestatística constitui um critério científico e moderno de interpretação e estudo da maior parte dos fenómenos naturais. Esta ideia vem ser generalizada pelo centro de geoestatística de Fontainebleau na França⁴ ao concluir que a geoestatística estuda qualquer fenómeno no qual pode ser quantificado, desenvolvido e interpretado por uma variável espacial e/ou temporal.

Por sua vez, Mingoti e Fidelis (2001) salientam que a geoestatística tem sido aplicada para avaliação de taxas de mortalidade infantil, abundância de espécies e previsão de número de casos de doenças em várias áreas no mundo. Lembrar que para este trabalho, a geoestatística será utilizada na perspectiva temporal para prever o número de unidades de sangue no Banco de Sangue do HCM e HGJM.

4.3 Metodologia de Geoestatística para Séries Temporais

O uso de geoestatística em \mathbb{R} é possível pela própria formulação da metodologia, embora sua maior aplicação seja no \mathbb{R}^d , $d \geq 2$ (Mingoti e Gloria, 2002). Dentro do contexto de séries temporais o objectivo é estudar o comportamento da variável resposta $z(t)$ que é medida no tempo t . Portanto, o tempo faz papel da coordenada de localização da unidade amostral.

Para se proceder a análise dos dados através da geoestatística, duas suposições são necessárias sobre os valores observados: a estacionaridade intrínseca e a isotropia (Mingoti e Gloria, 2002):

$$a) E[z(t)] = \mu, \forall t \in T \quad (2)$$

$$b) E\{[z(t_i) - z(t_j)]^2\} = 2\gamma(|t_i - t_j|), \forall t_i \neq t_j \in T \text{ e } i, j \in \mathbb{N} \quad (3)$$

isto é, o processo $z(t)$ tem média constante e não depende do tempo t ; e para cada $t_i \neq t_j$ a esperança matemática do quadrado das diferenças $[z(t_i) - z(t_j)]$ deve ser uma função que depende apenas da distância no tempo das observações $z(t_i)$ e $z(t_j)$, isto é, $\|t_i - t_j\|$. Sendo T o conjunto de indexação de tempo. $2\gamma(|t_i - t_j|)$ representa uma função conhecida como variograma, sendo que $\gamma(|t_i - t_j|)$ representa a função Semi-Variograma (metade do variograma).

⁴ <http://cg.ensmp.fr/HomePageEnglish.html> [extraído no dia 20/10/2006]

4.3.1 Variograma

A isotropia do processo permite definir a função intrínseca chamada Variograma, uma ferramenta básica de suporte às técnicas de geoestatística. É imprescindível a sua análise, visto que as equações de previsão para a geoestatística são todas funções dos valores do Variograma do processo estocástico gerador da série.

De acordo com a equação 3, tem-se que o variograma é definido como sendo a esperança matemática do quadrado da diferença entre dois períodos, ou seja:

$$2\gamma(|t_i - t_j|) = E\{[z(t_i) - z(t_j)]^2\} \quad (4)$$

Se as condições especificadas pela hipótese intrínseca são contempladas, então o Variograma pode ser estimado através de um conjunto de valores $z(t)$ (Mingoti e Neves, 1999), ou seja:

$$2\hat{\gamma}(h) = \frac{\sum_{t=1}^{(n-h)} [z(t) - z(t-h)]^2}{(n-h)} \quad (5)$$

Sendo $2\hat{\gamma}(h)$ o Variograma experimental estimado a partir dos dados da série; $h \in T = \{1, 2, \dots, n-1\}$ e $(n-h)$ é o número de pares $(z(t_i), z(t_j))$, $i \neq j$, que estão a uma distância de h unidades da origem fixa.

4.3.2 Previsões

Quando se está a estudar o comportamento de uma variável, é obviamente necessário recolher uma quantidade considerável de dados. Porém o objectivo é geralmente compreender o comportamento dessa variável, a partir dos dados existentes.

A continuidade e a dependência da variável temporal verificada pelo Variograma e levando-se em consideração as hipóteses de variância mínima e não tendenciosidade condicionam a estimação valores da variável em estudo em períodos futuros com ajuda da técnica Kriging (Issaks *et. al.* 1989). Entende-se por Kriging como um método de interpolação semelhante a médias móveis que permite estimar valores de $z(t)$ em períodos que não foram observados. A técnica de Kriging assume que os dados existentes se encontram correlacionados no tempo, isto é, se considerarmos uma variável $z(t)$ num determinado origem t é muito provável que se encontrem resultados muito próximos de t quanto mais próximos estiverem de t (princípio da geoestatística). Porém, a partir de uma certa distância de t , provavelmente não se encontre valores aproximados de t porque a correlação pode deixar de existir.

Penso que este é um erro do tipo-não do computador.

As equações de previsão para tempo t_k qualquer não pertencente à série observada, são obtidas utilizando a técnica de predição espacial denominada *Ordinary Kriging* (Mingoti, 1996). Neste caso, a equação de previsão, $z(t_m)$, é dada por:

$$\hat{Z}(t_m) = \sum_{i=1}^n \lambda_i Z(t_i) \quad t_m > t_n \quad (6)$$

$$\text{(onde } \sum_{i=1}^n \lambda_i = 1 \text{)}$$

e $z(t_1), z(t_2), \dots, z(t_n)$ representam os valores observados da série. Para o caso específico de séries temporais, os coeficientes λ_i (pesos associados a cada valor $z(t)$) são obtidos através da seguinte equação (Mingoti, 1996):

$$\lambda_0 = \Gamma_0^{-1} \gamma_0 \quad (7)$$

$$\gamma_0 = (\gamma(t_m - t_1), \gamma(t_m - t_2), \dots, \gamma(t_m - t_n), 1)' \quad (8)$$

e Γ_0 é a matriz quadrada $(n+1) \times (n+1)$ dada por:

$$\Gamma_0 = \begin{cases} \gamma(t_i - t_j), & i = 1, \dots, n, j = 1, \dots, n, \\ 1 & , i = n+1, j = 1, \dots, n, \\ 0 & , i = n+1, j = n+1, \\ 1 & , i = 1, \dots, n, j = n+1 \end{cases} \quad (9)$$

onde γ representa o valor do Semi-Variograma. Neste caso, o estimador linear $\hat{Z}(t_m)$ minimiza a variância do erro de previsão.

Mingoti (1996) refere-se que em séries temporais se estamos interessados em fazer previsão de h passos à frente a partir de uma origem fixa (origem de previsão), considerando-se sempre o mesmo número de valores vizinhos v , as previsões que serão geradas terão todas a mesma variância de erro de previsão. O mesmo autor, refere-se ainda que quando os valores da equação (5) são próximos dos valores teóricos do variograma do processo pode-se utilizar estes valores na resolução das equações (7) e (8) de modo a se encontrar os coeficientes λ_i da equação de previsão.

4.4 Métodos de Suavização (Alisamento) Exponencial

Os métodos de suavidade exponencial para séries temporais formam uma grande classe de métodos que procuram capturar as informações contidas nas próprias observações da série. Estes métodos usam uma ponderação distinta para cada valor observado na série, de modo que valores

mais recentes recebem pesos maiores. Assim, os pesos formam um conjunto que decai exponencialmente a partir de valores mais recentes.

4.4.1 Método de Suavização Exponencial Simples

Quando os dados da série temporal mantêm-se constante sobre um nível médio, isto é, não apresentam tendência nem sazonalidade é usual o método de *suavidade exponencial simples* (Breseghello, 2005). O valor previsto de uma suavização exponencial simples tem duas componentes: o valor previsto para o período t , \hat{z}_t , e uma correção feita para compensar os erros de previsão ($z_t - \hat{z}_t$) cometidos nos períodos anteriores.

A sua representação matemática é dada por:

$$\hat{z}_{t+1} = \hat{z}_t + \alpha(z_t - \hat{z}_t) \quad (10)$$

Onde \hat{z}_{t+1} é a previsão para o período $t+1$, feita no período actual t . O coeficiente α é a percentagem de ajuste que será aplicada ao erro de previsão, podendo assumir valores entre 0 e 1; z_t é o valor observado na série temporal para o tempo t .

A equação (10) pode ser reformulada da seguinte maneira:

$$\hat{z}_{t+1} = \alpha z_t + (1 - \alpha)\hat{z}_t \quad (11)$$

Com esta equação é fácil demonstrar porquê o método é chamado de *exponencial*.

Seja $\hat{z}_t = \alpha z_{t-1} + (1 - \alpha)\hat{z}_{t-1}$ e substituindo \hat{z}_t na equação (11) teremos:

$$\hat{z}_{t+1} = \alpha z_t + (1 - \alpha)[\alpha z_{t-1} + (1 - \alpha)\hat{z}_{t-1}] \quad (12)$$

$$\hat{z}_{t+1} = \alpha z_t + \alpha(1 - \alpha)z_{t-1} + (1 - \alpha)^2 \hat{z}_{t-1} \quad (13)$$

Podemos agora substituir \hat{z}_{t-1} , depois \hat{z}_{t-2} e assim por diante obtemos a seguinte equação matemática:

$$\hat{z}_{t+1} = \alpha z_t + \alpha(1 - \alpha)z_{t-1} + \alpha(1 - \alpha)^2 z_{t-2} + \alpha(1 - \alpha)^3 z_{t-3} + \dots + \alpha(1 - \alpha)^{t-1} z_1 + (1 - \alpha)^t \hat{z}_1 \quad (14)$$

Os coeficientes relacionados aos valores observados decaem exponencialmente, mostrando que as observações mais recentes têm mais peso para a previsão do próximo período.

Para fazer previsões utilizando este método deve-se escolher um valor para α , que pode ser aquele que minimize os erros de previsão. Este valor, pode ser encontrado utilizando um

programa computacional não linear de otimização, como por exemplo, o suplemento SOLVER do MS Excel. Outro valor a ser estimado é o primeiro valor de \hat{z}_t . Quando dados históricos estão disponíveis, pode-se usar a média aritmética das N observações mais recentes (Souza, 2005). Caso contrário pode-se utilizar a última observação, ou fazer uma estimativa subjectiva (Breseghello, 2005).

4.4.2 Método Holt

No ano de 1957 Charles C. Holt expandiu o modelo de suavização exponencial simples para poder lidar-se com dados que apresentavam tendência linear (Souza, 2005) e, assim poder realizar previsões que fossem mais precisas que as realizadas com o modelo de suavização exponencial simples, por ajustá-las a esta condição (Bezerra e Araújo, 2005); modelo este que veio a ser chamado de *Holt*.

Agora para se fazer previsões são necessárias duas constantes de suavização, α e β , que também variam entre 0 e 1. *As* previsões são estimadas utilizando as três equações abaixo:

Letra Maiscula → $L_t = \alpha z_t + (1 - \alpha)(L_{t-1} + T_{t-1})$ (15)

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (16)$$

$$\hat{z}_{t+m} = L_t + mT_t \quad (17)$$

A equação (15) representa o ajuste do nível⁵ da série e a equação (16) representa o ajuste da tendência. Já a equação (17) calcula as previsões para os próximos m períodos.

Como em outros métodos exponenciais existe uma necessidade de se estimar o primeiro valor para L_1 e T_1 . Pode-se fazer $L_1 = z_1$ e $T_1 = z_2 - z_1$ (Pellegrini, *at. el.*, 2000). Outra questão é escolher os valores de α e β , assim como na suavidade exponencial simples pode-se usar os valores que minimizem o erro cometido pelas previsões dentro da amostra.

4.4.3 Método de Holt-Winters

Anos depois, em 1960, Peter R. Winters estendeu o modelo de Holt, incluindo uma nova equação que possibilitasse ser acrescida nas previsões o comportamento da componente sazonal dos dados que se estivessem trabalhando, gerando assim o Método de *Holt-Winters* (Bezerra *et al.*, 2005).

⁵ O *nível*, segundo Enders (2005), refere-se a uma série temporal horizontal, ou seja, os dados históricos não possuem tendência, sazonalidade ou variações aleatórias. O mesmo autor diz ainda que uma série temporal caracterizada pelo padrão ou componente supracitada, os dados flutuam ao redor de uma média constante. Quando isso ocorre, tal série é denominada estacionária em relação à média.

A aplicação do método de Hot-Winters não exige que dados originais respeitem uma distribuição normal e/ou sejam estacionários tratando-se por tanto de um método determinístico ^{portanto} (Bezerra et al., 2005). Isso torna-o bastante popular, todavia, por ser determinístico, tem alguns inconvenientes no que diz respeito ao aumento do tamanho temporal considerado, pois, se o modelo apresentar discrepâncias consideráveis não é possível estudar seus resíduos (que dão variação não explicada dos dados) e com isso melhorar sua previsão de forma que o torne mais aderente.

Singular: ométodo

permite

O Método Holt-Winters ^{tem} permitem não só a projecção dos dados para m períodos, como também ajustar o nível, a tendência e a sazonalidade da série observada baseando-se em três equações de amortecimento. Os Modelos Holt-Winters dividem-se em dois grupos: Aditivo e Multiplicativo. No modelo aditivo, a amplitude da variação sazonal é constante ao longo do tempo enquanto que no modelo multiplicativo, a amplitude da variação sazonal aumenta ou diminui em função do tempo.

4.4.3.1 Método Holt-Winters Multiplicativo

É utilizado na modelação de dados sazonais onde a amplitude do ciclo sazonal varia com o passar do tempo. A sua representação matemática é dada por:

$$L_t = \alpha \frac{z_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (18)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (19)$$

$$S_t = \gamma \frac{z_t}{L_t} + (1 - \gamma)S_{t-s} \quad (20)$$

$$\hat{z}_{t+m} \doteq (L_t + mT_t)S_{t-s+m} \quad (21)$$

→ Onde L_t representa o nível (*level*) da série; T_t denota a tendência (*trend*); S_t é a componente sazonal e, \hat{z}_{t+m} corresponde à previsão para m períodos à frente. A constante α representa o coeficiente de ponderação, fixado pelos analistas que varia de 0 a 1 sendo que quanto maior for o valor, mais rapidamente o modelo reagirá à uma variação real dos dados observados; β representa a constante de amortecimento de tendência e γ permite suavizar a presença de dados aleatórios na sazonalidade (β e γ também variam de 0 a 1). A equação (18) difere da equação que trata do nível da série no modelo de Holt, já que o primeiro termo é dividido por uma componente sazonal, eliminando assim a flutuação sazonal de z_t . A equação (19) é exactamente

igual à equação da tendência no método de *Holt*. Já a equação (20), faz um ajuste sazonal nas observações z_t .

Todos os métodos de suavização exponencial incluindo o de Holt-Winters necessitam de valores iniciais de suas componentes (neste caso, nível, tendência e sazonalidade) para dar início aos cálculos. Breseghello (2005) cita Makridakis *et al.* (1998) que para a estimativa da componente sazonal, necessita-se, no mínimo, de uma estação completa de observações, isto é, s períodos. As estimativas iniciais de nível e da tendência são feitas no período s definido com a componente sazonal.

Segundo Makridakis *et al.* (1998) citados por Bezerra *et al.* (2005) para a inicialização do nível pode-se usar a média das observações do primeiro período sazonal, conforme a equação abaixo:

$$L_s = \frac{1}{s}(z_1 + z_2 + z_3 + \dots + z_s) \quad (22)$$

O cálculo da estimativa inicial para a tendência requer o uso de duas estações completas ($2s$):

$$T_s = \frac{1}{s} \left(\frac{z_{s+1} - z_1}{s} + \frac{z_{s+2} - z_2}{s} + \dots + \frac{z_{s+s} - z_s}{s} \right) \quad (23)$$

Para índices sazonais, utilizam-se s estimativas iniciais:

$$S_1 = \frac{z_1}{L_s}, S_2 = \frac{z_2}{L_s}, \dots, S_s = \frac{z_s}{L_s} \quad (24)$$

Os valores das constantes de suavização (α , β e γ) seguem a mesma lógica de determinação sugerida para os outros métodos de suavização exponencial anteriormente descritos.

4.4.3.2 Método Holt-Winters Aditivo

O método aditivo de Holt-Winters é utilizado na modelação de dados sazonais onde a amplitude do ciclo sazonal permanece constante com o passar do tempo. Este método compõe-se de seguintes equações de nível, tendência, sazonalidade e previsão respectivamente:

$$L_t = \alpha(z_t - s_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (25)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (26)$$

$$S_t = \gamma(z_t - L_t) + (1 - \gamma)S_{t-s} \quad (27)$$

$$\hat{z}_{t-m} = L_t + mT_t + S_{t-s+m} \quad (28)$$

Como se pode observar, a equação da tendência permanece a mesma utilizada para o método Holt-Winters Multiplicativo. Nas demais equações, a única diferença, é que a componente sazonal ^{está a efectuar ou efectua} (está efectuando) operações de soma e subtração, ao invés de multiplicação e divisão, respectivamente. Os valores iniciais de L_s e T_s são calculados de forma idêntica aos do modelo Multiplicativo. Mas as componentes sazonais são calculadas da seguinte maneira:

$$S_1 = z_1 - L_s, S_2 = z_2 - L_s, \dots, S_s = z_s - L_s \quad (29)$$

Os valores das constantes de suavização (α , β e γ) também seguem a mesma lógica de determinação sugerida para os outros métodos de suavização exponencial.

4.5 Método de Box-Jenkins

Segundo Pellegrini e Fogliatto (2000), os métodos quantitativos utilizam dados históricos para prever a série em períodos futuros. Uma das técnicas quantitativas mais popular é o método Box-Jenkins, também conhecido como *Modelos Auto-Regressivos Integrando a Média Móvel*, ou simplesmente ARIMA⁶, que foi proposto por George E. P. Box e Gwilym M. Jenkins no início da década 70 (Box *et al.*, 1994) citado por Breseghello (2005). O método de Box-Jenkins consiste na busca de um modelo ARIMA que represente o processo estocástico gerador da série temporal, a partir de um modelo ARMA⁷ aplicável na descrição de séries temporais estacionárias, estendendo esse conceito para séries temporais não estacionárias (Luz *et. tal.*, 2005).

Os modelos de Box-Jenkins partem da ideia de que os valores de uma série temporal são altamente dependentes, ou seja, cada valor pode ser explicado por valores prévios da série. Os modelos ARIMA representam a classe mais geral de modelos para a análise de séries temporais. Para este método as previsões dependem da determinação de cada um dos processos:

1. o processo Auto-Regressivo (AR)⁸ que representa a ordem de relacionamento com os dados passados;
2. o processo de Integração (I)⁹ ou diferenciação, que representa a ordem do processo integrado não-estacionário, ou seja, o número de vezes que o processo deve ser submetido a diferença para que se torne estacionário;

⁶ Auto-Regressive Integrated Moving Average

⁷ Auto-Regressive Moving Average

⁸ Auto-Regressive

⁹ Integrated

3. o processo Médias Moveis (MA)¹⁰ que representa o número de defasagens dos ruídos aleatórios não Autocorrelacionados que terá a série temporal como função linear.

Todavia, uma série pode ser modelada pelos três processos anteriormente descritos ou apenas por um, resultando deste modo em vários modelos ditados exclusivamente pelas características das próprias séries.

A grande vantagem do método ARIMA é em função dos cuidados teóricos e práticos que orientam a análise dos erros de previsão (Pellegrini *et. al*, 2000). No final, os erros devem ficar totalmente “pulverizados” no sentido de que eles não devem mostrar nenhum padrão de tendenciosidade ou sazonalidade e devem ter uma média zero e desvio padrão constante. Se porventura os erros demonstram algum tipo de padrão, isso significa que eles contêm mais informações úteis e essas informações podem e devem ser transferida para a estrutura do modelo adequado. Para a melhor compreensão dos modelos Box-Jenkins, descreve-se a seguir cada componente, divididas em dois grandes grupos de modelos (estacionários e não-estacionários):

4.5.1 Modelos Estacionários

Um processo estocástico é caracterizado por uma família de variáveis aleatórias que descrevem a evolução de algum fenômeno de interesse. Processos estocásticos que caracterizam o estudo de séries temporais descrevem a evolução *temporal* de um fenômeno de interesse. Uma importante classe de processos estocásticos utilizados na representação de séries temporais são os *Modelos estacionários*, que pressupõem um processo sob equilíbrio, onde a família de variáveis se mantém a um nível constante médio. A primeira característica requerida pelo método de ARIMA é que a série deve seguir uma regularidade chamada *estacionariedade*. Uma série estacionária se caracteriza por uma média e um desvio padrão sempre constante no tempo. Quando essa constância nos dados não é confirmada, terá que aplicar algum tipo de transformação antes que a análise de ARIMA seja aplicada aos dados. Teoricamente todos os outros perfis da distribuição dos dados, como a assimetria e curtose, também devem ser constantes, inclusive as covariâncias.

4.5.1.1 Modelo Auto-Regressivo (AR)

Um modelo estocástico útil na representação de um grande número de séries temporais é o *modelo auto-regressivo*. Neste modelo, o valor corrente do processo é expresso como uma combinação linear finita de valores prévios do processo e um ruído aleatório ε_t .

¹⁰ Moving Average

Para a sua modelação matemática, definem-se os valores observados de um processo em espaços de tempos igualmente divididos $t, t-1, t-2, \dots$ por $z_t, z_{t-1}, z_{t-2}, \dots$

Definem-se também $\tilde{z}_t, \tilde{z}_{t-1}, \tilde{z}_{t-2}, \dots$ como sendo os desvios da média μ , ou seja,

$$\tilde{z}_t = z_t - \mu, \tilde{z}_{t-1} = z_{t-1} - \mu, \tilde{z}_{t-2} = z_{t-2} - \mu, \dots$$

Assim, a equação do *modelo auto-regressivo de ordem p*, ou simplesmente $AR(p)$, é dada por:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + \varepsilon_t \quad (30)$$

Onde, os coeficientes ϕ_i descrevem como um valor corrente \tilde{z}_t se relaciona com os valores passados \tilde{z}_{t-i} para $i = 1, 2, \dots, p$. O coeficiente auto-regressivo de ordem p , pode ser expresso usando a definição do operador B ¹¹:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (31)$$

Simplificando a representação matemática do modelo auto-regressivo obtemos:

$$\phi(B)z_t = \varepsilon_t \quad (32)$$

A igualdade (32), segundo Box *et al* (2004) citado por Breseghello (2005), condiciona a estacionaridade do processo, desde que o polinómio $\phi(B)$ de grau p , tenha todas as suas raízes maiores que 1 em valores absolutos e que as auto-covariâncias (γ_k) sejam independentes. O modelo $AR(p)$ contém $p+2$ parâmetros desconhecidos ($\mu, \phi_1, \phi_2, \dots, \phi_p, \sigma_\varepsilon^2$), os quais podem ser estimados a partir de valores observados na série temporal. σ_ε^2 é a variância do processo de ruído aleatório ε_t . O processo auto-regressivo possui dois importantes casos especiais: os processos de primeira e segunda ordens.

A versão mais simples de um modelo AR é aquela em que \tilde{z}_t depende somente de \tilde{z}_{t-1} e de ε_t . Diz-se, nesse caso, que o modelo é auto-regressivo de ordem 1, o que se indica compactamente por $AR(1)$. A representação algébrica desse modelo é a seguinte:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \varepsilon_t \quad (33)$$

Para que o modelo seja considerado de estacionário é necessário que $|\phi_1| < 1$ e as auto-covariâncias (γ_k) sejam independentes. As auto-covariâncias são dadas por:

$$\gamma_k = \phi_1^k \gamma_0 \quad (34)$$

¹¹ backward shift operator

e as autocorrelações, ρ_k , são dada pela equação:

$$\rho_k = \phi_1 \rho_{k-1}, \text{ para } k > 0, \text{ ou } \rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k, k \geq 0. \quad (35)$$

A função de autocorrelação decai exponencialmente quando ϕ_1 é positivo; quando ϕ_1 é negativo, a função de autocorrelação também decai exponencialmente, mas apresenta alternância de sinal (Werner e Ribeiro, 2003).

Quando o $p=2$, tem-se um processo auto-regressivo da segunda ordem, designado por AR(2) e descrito por:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \varepsilon_t \quad (36)$$

Novamente, para que o processo seja estacionário, as raízes da equação $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 = 0$ devem estar fora do círculo unitário ($|\phi_1| < 1$). Isto obriga aos parâmetros ϕ_1 e ϕ_2 que satisfaçam as seguintes condições (Breshegello, 2005):

- i) $\phi_1 + \phi_2 < 1$;
 - ii) $\phi_1 - \phi_2 < 1$;
 - iii) $-1 < \phi_2 < 1$.
- (37)

A função de autocorrelação do processo AR(2) é dada por:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \text{ com } (k > 0). \quad (38)$$

Para $k=1$ e 2 temos as equações de Yule-Walker, respectivamente:

$$\rho_1 = \frac{\phi_1}{1 - \phi_2} \text{ e } \rho_2 = \phi_2 + \frac{\phi_2}{1 - \phi_2} \quad (39)$$

4.5.1.2 Modelos de Média Móvel

Para este modelo a série z_t , resulta da combinação dos ruídos brancos¹² ε do período actual com aqueles ocorridos em períodos anteriores. Assim um modelo de médias móveis de ordem q ou MA(q) é dado por:

$$\tilde{z}_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \quad (40)$$

¹² O termo de erro é chamado de "ruído branco" se ele possui média zero, variância constante σ^2 , e é não-autocorrelacionado

onde $\tilde{z}_i = z_i - \mu$; θ_i é o parâmetro que descreve como z_i se relaciona com o valor ε_{i-i} para $i = 1, 2, \dots, q$.

A equação 40 pode ser igualmente reformulada utilizando o operador B:

$$\tilde{z}_i = 1 - \theta_1 B - \theta_1 B^2 - \dots - \theta_q B^q = \theta(B)\varepsilon_i \quad (41)$$

a qual contém $q-2$ parâmetros desconhecidos $(\mu_c, \phi_1, \phi_2, \dots, \phi_p, \sigma_c^2)$, estimáveis a partir dos valores observados na série temporal (Brescghello 2005).

A função de autocorrelação de um processo MA(q) é:

$$\rho_k = \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, \quad (42)$$

com $k = 1, 2, \dots, q$ e $\rho_k = 0$ quando $k > q$.

O modelo MA(1) é a versão mais simples dessa classe de modelos. Sua representação matemática é dada pela equação:

$$\tilde{z}_i = \varepsilon_i + \theta_1 \varepsilon_{i-1}, \quad (43)$$

$$\text{ou } \tilde{z}_i = (1 + \theta_1 B)\varepsilon_i, \quad (44)$$

sendo estacionário para qualquer θ_1 .

A função de autocorrelação ρ_k do processo MA(1) é dada por:

$$\rho_k = \frac{-\theta_1}{1 - \theta_1^2}, \quad (45)$$

quando $k = 1$, e $\rho_k = 0$, quando $k \geq 2$. A primeira autocorrelação será positiva se $\theta_1 < 0$ e negativo se $\theta_1 > 0$. A obtenção da equação de invencibilidade, isto é, a transformação dum modelo MA(1) num modelo AR(∞), é condicionada pela restrição de que $|\theta_1| < 1$ (Werner e Ribeiro, 2003). Para que o processo se torne invertível, é necessário que as raízes da equação $\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) = 0$ sejam maiores que 1. O processo de médias móveis MA(2), representado por: $\tilde{z}_i = \varepsilon_i + \theta_1 \varepsilon_{i-1} + \theta_2 \varepsilon_{i-2}$, é estacionário para quaisquer valores de θ_1 e θ_2 .

A função de autocorrelação do processo MA(2) é dada por:

$$\rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}; \quad (46)$$

$$\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}; \quad (47)$$

$$\rho_k = 0, \text{ para } k \geq 3. \quad (48)$$

4.5.1.3 Modelos Auto-Regressivos de Médias Móveis

Em alguns casos, pode ser necessário utilizar um grande número de parâmetros em modelos puramente AR ou puramente MA (Werner e Ribeiro, 2003). Neste caso, é válido misturar as componentes de um modelo AR com as componentes de um modelo MA, gerando, assim, um modelo ARMA. Modelo ARMA (p, q) exigirá um número menor de termos e pode ser expresso conforme a equação:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}, \quad (49)$$

ou, utilizando a notação do operador de defasagem B e rearranjando os termos da equação (49) obteremos:

$$\phi(B)\tilde{z}_t = \theta(B)\varepsilon_t \quad (50)$$

O modelo possui $p+q+2$ parâmetros desconhecidos ($\mu, \phi_1, \phi_2, \dots, \phi_p; \theta_1, \dots, \theta_q, \sigma_\varepsilon^2$), que podem ser estimados a partir dos valores observados na série temporal. Na prática, Box *et al.* (1994) citados por Werner e Ribeiro (2003) dizem que os valores de p e q são geralmente menores que 2 (dois) para séries temporais estacionárias. As condições de estacionaridade estabelecidas para os processos AR(p) e MA(p) se mantêm nos modelos ARMA(p, q).

O modelo ARMA mais simples é o ARIMA (1,1) dado pela equação:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}. \quad (51)$$

A função de autocorrelação do modelo ARMA (1,1) é dada por:

$$\rho_1 = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2+2\phi_1\theta_1}; \quad (52)$$

$$\rho_k = \phi_1 \rho_{k-1} \text{ para } k > 1 \quad (53)$$

4.5.2 Modelos Não-Estacionários

Quando uma série temporal apresenta média e variância dependentes do tempo, é porque ela não é estacionária. Muitas são as séries que não possuem média constante, isto é, em nenhum dado intervalo de tempo, as observações da série se comportam como as observações de um intervalo

de tempo distinto. Tais séries são chamadas de *não-estacionárias na média* (Werner e Ribeiro, 2003). Da mesma forma, é possível uma série temporal exibir *não-estacionariedade* na média e no declive.

4.5.2.1 ^{Modelos} (Modelo) Auto-Regressivos Integrados de Médias Móveis (ARIMA)

Como a maioria dos procedimentos de análise estatística de séries temporais supõe que estas sejam estacionárias, será necessário transformá-las caso ainda não o sejam. Segundo Breseghello (2005) a transformação mais usual consiste em tomar diferenças sucessivas da série original até obter uma série estacionária. A primeira diferença de z_t é definida por $(z_t - z_{t-1})$. Essa diferença faz com que a série apresente um comportamento não-estacionário apenas na média. Após a segunda diferenciação dada por $(z_t - z_{t-1}) - (z_{t-1} - z_{t-2}) = z_t - 2z_{t-1} + z_{t-2}$, a série torna-se estacionária (Werner e Ribeiro, 2003).

Séries não-estacionárias podem geralmente ser representadas por um *operador auto-regressivo generalizado* $\varphi(B)$, no qual uma ou mais raízes do polinómio $\varphi(B)$ são iguais a 1 em módulo. Em particular, se existirem d raízes unitárias, o operador $\varphi(B)$ assumirá a forma abaixo (Box *et al.*, 1994) citados por (Pellegrini e Fogliatto, 2000):

$$\varphi(B) = \varphi(B)(1 - B)^d \quad (54)$$

onde $\varphi(B)$ é um operador estacionário. Em contrapartida um modelo que apresenta comportamento *homogéneo não-estacionário* é representado de seguinte forma:

$$\varphi(B)z_t = \varphi(B)(1 - B)^d z_t = \theta(B)\varepsilon_t \quad (55)$$

$$\text{ou } \varphi(B)w_t = \theta(B)\varepsilon_t \quad (56)$$

$$\text{onde } w_t = \nabla^d z_t = z_t - z_{t-d} \quad (57)$$

Assim, um comportamento homogéneo não-estacionário pode ser representado por um processo estacionário, com d níveis de diferenciação. Na prática, d pode ser 0, 1, ou, no máximo, 2.

Assim o processo definido pelas equações (55) e (57) produz um eficiente modelo para descrever séries temporais não-estacionárias. Este modelo é conhecido de processo Auto-Regressivos Integrados de Médias Móveis (ARIMA) de ordem (p, d, q) , onde p corresponde ao componente auto-regressivo (AR), d ao número de diferenciações e q ao componente de médias móveis.

O modelo ARIMA (p, d, q) é dado pela seguinte equação:

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (58)$$

A palavra “integrado” do modelo ARIMA tem o sentido de “somado”, já que a equação (58) pode ser escrita por:

$$z_t = S^d w_t, \quad (59)$$

sendo $S = \nabla^{-1} = (1 - B)^{-1}$ o operador soma, definido por:

$$S w_t = \sum_{j=0}^{\infty} w_{t-j} = w_t + w_{t-1} + w_{t-2} + \dots \quad (60)$$

Assim, o processo geral auto-regressivo integrado a médias móveis (ARIMA) pode ser gerado somando-se (integrando-se) o processo estacionário ARMA (w_t), d vezes.

A seguir, apresentaremos alguns casos especiais do modelo ARIMA (Box *et al.* 1994) referenciado por Breseghello (2005):

- Modelo ARIMA (0,1,1):

$$\nabla z_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} = (1 - \theta_1 B) \varepsilon_t, \quad (61)$$

onde $p = 0$, $d = 0$, $q = 0$, $\phi(B) = 1$, $\theta(B) = 1 - \theta_1 B$.

- Modelo ARIMA (0,2,2):

$$\nabla^2 z_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} = (1 - \theta_1 B - \theta_2 B^2) \varepsilon_t, \quad (62)$$

onde $p = 0$, $d = 2$, $q = 2$, $\phi(B) = 1$, $\theta(B) = 1 - \theta_1 B - \theta_2 B^2$.

- Modelo ARIMA (1,1,1):

$$\nabla z_t - \phi_1 z_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1}, \text{ ou} \quad (63)$$

$$(1 - \phi_1 B) \nabla z_t = (1 - \theta_1 B) \varepsilon_t, \quad (64)$$

onde $p = 1$, $d = 1$, $q = 1$, $\phi(B) = 1 - \phi_1 B$, $\theta(B) = 1 - \theta_1 B$.

4.5.2.2 Modelos Sazonais

Durante a modelação de uma série pode-se notar que ela apresenta variações sazonais. Isto ocorre quando a série exibe uma característica periódica que se repete a cada s intervalos de tempo. Por exemplo, em séries compostas por observações mensais e sazonalidade anual, s é igual 12. Define-se $\nabla_s = (1 - B^s)$ como sendo o *operador de diferença sazonal*. Assim $\nabla_s z_t = (1 - B^s) z_t = z_t - z_{t-s}$ é a primeira diferenciação sazonal. Em geral, D diferenciações sazonais podem ser requeridas para produzir uma série estacionária. Neste caso, o operador de diferenciação sazonal de ordem D é $\nabla_s^D = (1 - B^s)^D$. Assim, a forma geral do modelo sazonal auto-regressivo da média móvel de ordem (P, D, Q) é:

$$\Phi(B^s)\nabla_s^D z_t = \Theta(B^s)\varepsilon_t \quad (65)$$

Onde $\Phi(B^s)$ e $\Theta(B^s)$ são polinômios em B^s de graus P e Q , respectivamente, que satisfazem as condições de séries não-estacionárias. A representação (P,D,Q) é feita em letras maiúsculas, para diferenciá-las da representação feita nos modelos não sazonais (Breseghello, 2005).

No modelo da equação (65), as componentes de erro ε_t estão geralmente correlacionadas. Assim ε_t estaria relacionado com ε_{t-1} , ε_{t-2} , etc. Para tratar tal relacionamento, introduz-se um segundo modelo:

$$\phi(B)\nabla^d \varepsilon_t = \theta(B)\varepsilon_t \quad (66)$$

Onde ε_t é um processo de ruído aleatório. $\phi(B)$ e $\theta(B)$ são polinômios em B de grau p e q , respectivamente, que satisfazem as condições de série não-estacionária.

Subtraindo-se a equação (65) na (66), obtém-se um modelo multiplicativo geral,

$$\phi_p(B)\Phi_p(B^s)\nabla^d\nabla_s^D z_t = \theta(B)\Theta_Q(B^s)\varepsilon_t, \quad (67)$$

chamado de processo multiplicativo de ordem $(p,d,q) \times (P,D,Q)_s$.

4.5.3 Coeficiente de Autocorrelação

Uma estatística importante na análise de séries temporais é o coeficiente de autocorrelação ρ . A autocorrelação é usada para descrever a correlação entre dois valores da mesma série temporal, em diferentes períodos de tempo. Assim, um coeficiente de autocorrelação ρ_1 mede a correlação entre dois valores adjacentes na série, e a autocorrelação, neste caso, é dita *autocorrelação de lag* (ou desfasagem) 1. De maneira genérica, o coeficiente de autocorrelação ρ_k mede a correlação entre as observações distantes k períodos de tempo (ou seja, uma autocorrelação de *lag k*). A autocorrelação de *lag k* é expressa pela seguinte equação:

$$\rho_k = \frac{E[(z_t - \mu) \cdot (z_{t-k} - \mu)]}{\sqrt{E[(z_t - \mu)^2] \cdot E[(z_{t-k} - \mu)^2]}}; \quad (68)$$

ou

$$\rho_k = \frac{E[(z_t - \mu) \cdot (z_{t-k} - \mu)]}{\sigma_z^2}; \quad (69)$$

onde σ_z^2 é a variância da série temporal. Uma estimativa do coeficiente de autocorrelação populacional ρ_k nas equações (68) e (69) é dada pelo coeficiente de *autocorrelação amostral*:

$$r_k = \frac{\sum_{t=k+1}^n (z_t - \bar{z})(z_{t-k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2}, \text{ com } k = 0, 1, 2, \dots, n. \quad (70)$$

Onde $\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$ *estimativa?*

Na prática, para se obter uma boa estimativa do coeficiente de autocorrelação, deve-se dispor de pelo menos 50 observações da variável z . Segundo Enders (1995) o número de autocorrelações de lags diferentes que se calcula para a análise de série temporal deve ser de $N/4$, onde N é o número total de observações na série. Similarmente à autocorrelação, a autocorrelação parcial também permite analisar o relacionamento entre os valores de uma série temporal. Porém, a autocorrelação parcial mede o grau de associação entre z_t e z_{t-k} quando o efeito de outros lags (1, 2, 3, ..., $k-1$) são removidos. A autocorrelação parcial é representada por ϕ_{kk} .

O coeficiente de autocorrelação parcial ϕ_{kk} é o k -ésimo coeficiente num processo auto-regressivo de ordem k .

4.5.4 Raiz Unitária e Teste ADF

Augmented Dickey - Fuller
(Deve evitar-se abreviaturas nos cabeçalhos)

A presença de tendência em séries de tempo compromete a aplicação de algumas metodologias de previsão de séries temporais. Os modelos de regressão linear só têm suas propriedades asseguradas se os valores da variável em análise forem estacionários, ou seja, se as suas média e variância se apresentam constantes ao longo do tempo e o valor da covariância entre dois períodos de tempo depender apenas da distância entre os dois períodos.

Vários testes podem ser utilizados para detectar se a série é estacionária. Neste trabalho utilizou-se o teste ADF (Augmented Dickey-Fuller) para verificar a presença de raiz unitária nas séries consideradas.

Se uma série deve ser diferenciada d vezes antes de tornar-se estacionária, então ela contém d raízes unitárias e é dita ser integrada de ordem d , denotada por $I(d)$.

Neste contexto um processo $I(d)$, z_t , $t = 1, \dots, n$, é definido por

$$(1 - B)^d z_t = \varepsilon_t, \quad t = 1, 2, \dots, n \quad (71)$$

onde B é um operador da forma $Bz_t = z_{t-1}$ e ε_t é $I(0)$.

Os testes de raízes unitárias são capazes de detectar se a série foi suficientemente diferenciada para se tornar estacionária. Para tão, testa-se:

H_0 : a série possui raiz unitária, e

H_1 : a série não possui raiz unitária.

No caso mais simples o teste é realizado por mínimos quadrados num modelo AR(1), $z_t = \rho z_{t-1} + \varepsilon_t$, e testa-se $\rho = 1$ e $\rho \neq 1$.

Como na maioria dos casos ε_t não é ruído branco, utiliza-se uma correcção. Neste caso, será utilizado o teste com a correcção paramétrica, conhecido como teste ADF (Vasconcellos e Alves, 2000), que sugere aumentar a regressão adicionando termos suficientes em Δz_{t-1} para “pulverizar” os resíduos. Então o teste ADF é a estatística-t usada para testar o coeficiente de z_{t-1} em:

$$\Delta z_t = (\rho - 1)z_{t-1} + \sum \theta_i \Delta z_{t-i} + \varepsilon_t \quad (72)$$

que tem a mesma distribuição assintótica de Dickey-Fuller e não depende de θ_i . No entanto, é necessário truncar a ordem da auto-regressão por regra, por um método automático como o AIC (Akaike Information Criterion) ou por alguma outra especificação.

A hipótese nula de uma raiz unitária ($H_0: \rho - 1 = 0$) é baseada no teste-t com uma distribuição não normal. Mackinnon (1991) apresenta a tabela com os valores críticos apropriados para alguns tamanhos amostrais (Vasconcellos e Alves, 2000).

4.5.5 Etapas da Metodologia Box-Jenkins

A metodologia Box-Jenkins ^{tem por objectivo} (objectiva) determinar um modelo estocástico linear da classe ARIMA (ou SARIMA) que, por hipótese, tenha gerado a série temporal z_t (Granger *et al.*, 1986). Para esta determinação, quatro etapas iterativas devem ser cumpridas.

1ª etapa: Identificação do Modelo

De forma resumida, esta etapa consiste em identificar quais dos “filtros” AR, I e MA (e SAR, I e SMA) fazem parte do processo estocástico que gerou a série temporal z_t , bem como identificar as suas respectivas ordens (Vasconcellos e Alves, 2000). Esta é a etapa mais complexa da metodologia, não havendo consenso em relação à melhor estratégia a ser usada (Granger *et al.*, 1986).

De acordo com Enders (1995), deve-se, primeiro, analisar o gráfico da série temporal z_t e suas funções de autocorrelação (ACF) e de autocorrelação parcial (PACF). No caso da análise gráfica

da série, a presença de forte tendência ou a aparência de variância e média que não sejam constantes servem como indicativo da não estacionaridade da série. Esta primeira verificação deve ser complementada pela análise da ACF e da PACF.

Não existe a secção 3.5.4 neste capítulo

O teste final da não estacionaridade da série pode ser feito mediante o teste da raiz unitária aumentado de Dickey-Fuller (teste ADF) descrito na ^{(4-54)?} secção 3.5.4 deste capítulo. Para a aplicação deste teste, deve-se definir, primeiro, o número óptimo da defasagem temporal da série pelos critérios de Schwarz (SBC) e Akaike (AIC). Caso seja validada a hipótese de não estacionaridade pelo teste ADF, a série temporal z_t deve ser diferenciada quantas vezes forem necessárias, até que se torne estacionária.

Só com a série temporal estacionária é que se busca a selecção de valores de p , d , q (e P , D , Q , no caso da presença da sazonalidade). Para esta finalidade são utilizados, de novo, os instrumentais da ACF e da PACF.

2ª etapa: Estimação do Modelo

Considerando válida a condição de estacionaridade, e verificadas ^{as} possíveis configurações de modelos na etapa da identificação, a etapa de estimação leva em consideração três aspectos: a parcimónia do modelo, as condições de estacionaridade e de invencibilidade e a qualidade do ajuste (Enders, 1995).

O aspecto da parcimónia considera que a incorporação de coeficientes adicionais se, por um lado, aumenta o grau de ajuste, por outro reduz os graus de liberdade. Modelos que sejam parcimoniosos geram melhores projecções que as configurações sobre-parametrizadas. O objectivo, portanto, é se aproximar do verdadeiro processo gerador sem que, no entanto, ^{se simule} ~~simule~~ se o exacto processo (Enders, 1995).

A abordagem de Box-Jenkins pressupõe que o modelo seja tanto estacionário como invertível. Do ponto de vista formal, z_t é invertível se puder ser representado por um processo auto-regressivo de ordem infinita ou convergente. Este aspecto é relevante pois a ACF e a PACF pressupõem, de forma implícita, que a série em questão possa ser representada por um modelo auto-regressivo.

Para se verificar a qualidade do ajuste do modelo, utilizam-se, de forma complementar à ACF e à PACF, os critérios AIC e SBC. Estes critérios têm como principal atributo a "penalização" dos modelos com muitos parâmetros, escolhendo-se aqueles mais parcimoniosos (Vasconcellos e

Alves, 2000). Na prática este procedimento pode ser aplicado de forma semi-automática já que muitos pacotes estatísticos (ex. MINITAB e SPSS) fornecem estes valores. Porém, após um modelo ser seleccionado a análise residual ainda deve ser feita antes de se passar à etapa das previsões.

3ª etapa: Diagnóstico do Modelo

Para se fazer a verificação da correcta especificação do modelo ARIMA(p,d,q) ou SARIMA(p,d,q)(P,D,Q), analisa-se o termo aleatório ε_t , que deve apresentar as propriedades de um ruído branco. Caso estas propriedades não sejam verificadas, deve-se reiniciar a tentativa de selecção do modelo pela 1ª etapa, o que ressalta a característica interactiva desta metodologia.

Com este objectivo, a prática comum é a construção de correlogramas (ACF e a PFAC) dos resíduos para verificar a presença de *lags* significativos. Entende-se por *lags* significativos aqueles que ultrapassam os limites de $\pm 2S(r_k)$ para a ACF e $\pm 2S(\hat{\phi}_{kk})$ para PACF onde

$$S(r_k) \approx \sqrt{\frac{1 + 2(r_1^2 + r_2^2 + \dots + r_q^2)}{n}} \quad e \quad (73)$$

$$S(\hat{\phi}_{kk}) \approx \sqrt{\frac{1}{n}} \quad (74)$$

onde n é o número de observações da série.

Neste contexto, torna-se importante que os resíduos do modelo estimado não sejam serialmente correlacionados. Isto pode ser verificado pela estatística de Ljung-Box (estatística Q) se todas as autocorrelações residuais ou autocorrelações parciais são estatisticamente significantes.

De acordo com Enders (1995), não existe um nível de significância que possa ser considerado o “mais apropriado”. Dois pontos, porém, ajudam a definir se um modelo está bem ajustado: a presença de correlações residuais marginais que sejam significantes; e valores da estatística Q que sejam significantes ao nível de 10%. Nestas circunstâncias, torna-se possível formular um modelo que ajusta melhor os dados da série.

4ª etapa: Previsão

Uma vez determinado o melhor modelo para a série temporal em estudo, pode-se usar para gerar previsões de observações futuras. Partindo-se do período presente t , e supondo que se deseja prever a série num período futuro, \hat{z}_{t+k} representa a previsão para um período $t+k$ feita em t . A previsão para período $t+k$ é normalmente construída a partir de sucessivas previsões para os

períodos $t+1, t+2, \dots, t+k-1$. Neste procedimento, o valor de z_{t+j} , o qual não se conhece no tempo t , é substituído pela sua previsão \hat{z}_{t+j} . O valor de ε_{t+j} o qual também não se conhece no tempo t , é substituído por zero, e $\varepsilon_{t-j} = z_{t-j} - \hat{z}_{t-j}$. No início do processo de previsão deve-se assumir que $\varepsilon_{t-j} = 0$ para $t-j \leq 0$.

5. Uma Discussão sobre a Comparação dos Métodos de Previsão

Existem informações aparentemente contraditórias em relação aos métodos mais eficientes de ajuste e previsão de séries temporais. Pellegrini et. al. (2000) referem-se que Makridakis fez “competições” em 1979 e em 1982 com séries temporais e os métodos automáticos tiveram, em média, o melhor desempenho dentre todos os métodos. Já Morettin e Toloí (1981) afirmam que estudos empíricos (da época) apontam para uma grande vantagem dos métodos Box-Jenkins em relação aos demais.

Luz, et. al. (2005) cita Makridakis e Hibon (2000) que publicaram resultados da mais recente competição entre séries temporais, relativas a diferentes áreas tais como a industrial, a financeira, a demográfica e outras. Esta competição feita com 3003 séries temporais, confirmou as conclusões obtidas anteriormente, de que métodos mais simples (os automáticos) tem um desempenho tão bom quanto, e em muitos casos melhor que outros métodos estatísticos sofisticados, desde que a natureza da série seja desconhecida para o analista.

Mueller (1996), concluiu em seus estudos sobre a previsão do mercado, que o Alisamento Exponencial Simples é mais satisfatório entre os métodos de previsão avaliados naquele trabalho, dentre os quais os modelos Box-Jenkins.

Refenes (1993) cita grande popularidade dos modelos Box-Jenkins em muitas áreas onde se confirma a força e flexibilidade destes modelos. Comenta também que esta flexibilidade torna esta técnica de difícil aplicação, exigindo grande experiência do pesquisador.

Luz, et. al. (2005) comentam que os modelos Box-Jenkins tiveram *frutíferas* (aplicações) em previsões sociais, económicas, engenharia, comércio internacional e problemas de stock. Os mesmos autores afirmam que a grande vantagem destes modelos está na previsão para períodos curtos, e que sua limitação é que necessitam no mínimo de 50 observações e preferencialmente 100 ou mais para serem usados.

Referências similares às citadas anteriormente são facilmente encontradas em artigos, livros e outras publicações, tais como dissertações e teses. Tais referências apontam para o melhor tipo de modelo, de acordo com uma situação específica. É preciso compreender que grau de acuidade, horizonte de precisão desejado e a disposição (ou não) de se trabalhar com um maior rigor matemático influenciarão directamente na escolha do “melhor” modelo.

6. Critérios de Comparação de Qualidades de Previsões

melhorar esta frase
Todo o processo de previsão seja realizado utilizando métodos qualitativos ou quantitativos tem uma incerteza associada. Assim, torna relevante a estimação desta incerteza através do cálculo do *erro de previsão* (Lopes, 2002).

Para comparar diferentes métodos de previsão, ou para ajudar na selecção de modelos, podem ser usados diversos critérios de comparação da qualidade de previsões. Geralmente, estes critérios são função dos erros de previsão e eventualmente (no caso de serem critérios que usem medidas relativas) dos valores da série em estudo.

Os critérios podem envolver: diferentes números de passos para as previsões e diferentes pesos para os erros de previsão, em função do instante de base para a previsão e do número de passos - dando, em geral, maior importância a previsões mais recentes e/ou a um menor número de passos.

A estimativa e acompanhamento dos erros de previsão são necessários para o processo de manutenção e monitoramento das previsões geradas. Porém, existe a necessidade de acompanhamento do desempenho dos modelos de previsão para que se possa confirmar suas validades e mantê-los actualizados. Este monitoramento é realizado através de cálculo do erro nas previsões e tem os seguintes objectivos:

1. verificar a precisão dos valores previstos;
2. identificar, isolar e corrigir variações anormais;
3. permitir a escolha de técnicas, ou parâmetros, mais eficientes.

Segundo Davis et al. (2001), os erros de previsão podem ser classificados em erros aleatórios e distorções. Os primeiros podem ser definidos como aqueles que não podem ser explicados. As distorções, por sua vez, incluem:

1. falha na inclusão de valores correctos;
2. utilização de relacionamentos errados entre variáveis;
3. emprego da linha de tendência errada;

4. localização da aderência sazonal em pontos diferentes de onde ela normalmente ocorre.

Por fim, torna-se claro que deve-se priorizar a utilização de modelos de previsão que gerem o menor erro no tempo. Esse procedimento conduz a maior segurança na selecção de modelos com fins preditivos que possuam maior grau de precisão. O erro de previsão é a diferença entre o valor ocorrido e o valor previsto, sendo o *erro* igual à:

$$e_t = z_t - \hat{z}_t \quad (75)$$

Onde:

e_t : erro de previsão no período t ;

z_t : observação actual para o período t ;

\hat{z}_t : previsão para o período t .

É aconselhável na formulação do modelo separar os dados em duas partes: uma chamada de *conjunto inicialização* e outra *conjunto teste*. Os dados considerados de conjunto inicialização são utilizados para estimar o modelo e os dados chamados de conjunto teste servem para validar o modelo.

Uma vez que o cálculo dos erros pode resultar em valores positivos ou negativos, o que, em algumas situações, resulta em somatório nulo, diferentes manipulações da estatística de erro (e_t) devem ser empregues. Estas diferentes manipulações consistem em critérios de escolha de modelos mais apropriados às séries temporais. Os critérios mais utilizados são:

a) Erro Médio (EM)

O erro médio constitui-se na média dos erros de previsão. O valor fornecido pelo erro médio poderá ser pequeno, uma vez que erros positivos e negativos encontrados durante os períodos tendem a se anular. O seu mérito é informar se a previsão foi sistematicamente acima ou abaixo do observado. A sua equação é:

$$EM = \frac{1}{n} \sum_{i=1}^n e_i \quad (76)$$

Onde: n corresponde ao número de previsões efectuadas.

O erro médio não funciona como critério de avaliação da precisão das previsões quando comparado com métodos diferentes de previsão, pois o erro individual pode ter valores grandes ou pequenos e ainda o erro médio será próximo ^{de} zero. Nesse caso os métodos mais precisos não podem ser distinguidos dos métodos piores. ←

b) Erro Quadrático Médio (EQM)

Erro Quadrático Médio define-se como sendo a média do quadrado do desvio entre a previsão e a observação real. É considerado uma medida importante para a escolha do modelo, pois segundo Lewis (1997) o EQM é utilizado para determinar o modelo de previsão mais preciso, isto é, considera-se o melhor modelo, aquele que possuir o menor EQM.

Sua definição matematicamente é:

$$EQM = \frac{1}{n} \sum_{t=1}^n e_t^2 \quad (77)$$

c) Erro Absoluto Médio (EAM)

Erro Absoluto Médio é considerado como a média dos erros absolutos. Esse procedimento supera a característica de cancelamento dos erros positivos e negativos no erro médio. A equação a seguir pode ser utilizada para cálculo da média em questão:

$$EAM = \frac{1}{n} \sum_{t=1}^n |e_t| \quad (78)$$

O erro é tratado como uma distância (entre a previsão \hat{z}_t e o valor observado z_t) que, então, não tem sinal. Todas as distâncias são somadas e a média calculada. Esse critério é utilizado para distinguir a precisão entre vários métodos de previsão. Desta forma, um método adequado deve ter um EAM pequeno.

d) Erro Percentual Absoluto Médio (EPAM)

O Erro Percentual Absoluto Médio é a média de todos erros absolutos percentuais. Para Lewis (1997), o EPAM é considerado como uma das medidas de erro mais usadas para se avaliar os métodos de previsão. Este erro fornece uma indicação do tamanho médio do erro, expresso como uma percentagem do valor observado; independentemente do erro ser positivo ou negativo.

Pode-se utilizar a equação matemática abaixo para o cálculo do EPAM:

$$EPAM = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{z_t} \times 100 \right| \quad (79)$$

O denominador da expressão anterior é o valor observado, mas pode ser utilizado ou z_t ou \hat{z}_t . Não existe nenhuma regra ou raciocínio mais adequado. Às vezes z_t ou \hat{z}_t pode conter valores nulos que torna sua colocação no denominador inapropriado (Samohyl, 2004). Na prática, utiliza-se a variável que dará os melhores resultados, mais interessantes, conforme o objectivo do pesquisador.

De acordo com Lewis (1997), os valores do Erro Percentual Absoluto Médio podem ser relacionados com a potencialidade das previsões segundo a tabela abaixo:

Tabela 1. Potencialidades de previsão do EPAM

EPAM	POTENCIALIDADE DE PREVISÃO
Menor que 10%	A previsão é potencialmente muito boa
10% – 20%	A previsão é potencialmente boa
20% – 30%	A previsão é potencialmente razoável
Maior que 30%	A previsão é potencialmente inexacta

e) Estatística U de Theil

Nenhuma medida citada anteriormente dá uma boa base de comparação ordinal quanto aos ganhos de precisão auferidas pelo uso de determinados métodos de previsão em detrimento de outros. Por exemplo, será que pelo facto de se observar um valor de cinco unidades para a estatística EQM, ou uma percentagem de 3,2% para EPAM já permite dizer que um determinado método é bom predictor para um processo qualquer?

Dentro deste contexto a estatística U de Theil (ou coeficiente de desigualdade) é a adequada, pois permite uma comparação relativa entre os métodos de previsão formais com o método trivial de previsão.

Matematicamente, a estatística U de Theil é definida como:

$$U = \frac{\sqrt{\sum_{t=1}^{n-1} \left(\frac{z_{t+1} - z_t}{z_t} \right)^2}}{\sqrt{\sum_{t=1}^{n-1} \left(\frac{z_{t+1} - z_t}{z_t} \right)^2}} \quad (80)$$

A característica positiva da utilização da estatística U de Theil, como medida de precisão, é segundo Breseghello (2005), a sua interpretação intuitiva:

Se $U < 1$: a técnica de previsão testada é mais acurada que o método trivial, portanto, é vantajosa a utilização do método testado.

Se $U = 1$: a técnica de previsão testada é tão precisa quanto o método trivial de previsão. Logo, não existe diferença, no que tange a qualidade de previsão, se for empregue tanto o testado quanto o modelo trivial para a tarefa de previsão.

Se $U > 1$: a técnica de previsão testada é inferior (menor precisão), em relação ao método trivial. Sendo assim, é vantajoso utilizar o método trivial para prever o processo em estudo.

→ No entanto há que se fazer uma ressalva sobre a estatística U de Theil: a interpretação, da forma como foi descrita acima, na qual o valor de U para o Modelo Trivial, por definição é igual a um, só é válida quando o critério de decisão ^{se} baseia-se, eminentemente, sobre os erros de ajuste (erros gerados para o conjunto de inicialização).

Uma explicação intuitiva para esse critério diz que quando a previsão para o período seguinte $t+1$, \hat{z}_{t+1} , é perfeitamente no alvo e igual ao valor observado z_{t+1} , então $z_{t+1} = \hat{z}_{t+1}$, e o numerador da expressão é nulo, conseqüentemente, o valor de $U = 0$, impossível encontrar na prática. Deve ficar claro agora porquê o critério é chamado de coeficiente de desigualdade; quando for igual a zero, não há desigualdade nenhuma entre o valor previsto e o valor observado; a previsão não pode ser melhor.

→ Entretanto, Breseghello (2005) refere-se ainda que, quando o critério de decisão se baseia sobre os erros de previsão (erros gerados para o conjunto de teste), a interpretação acima não é válida (continua-se desejando o menor valor de U , pois quanto menor for o valor de U maior será a indicação ^{de} que o modelo tem uma boa capacidade de previsão), porém o valor de U para o modelo trivial não é igual a um, por definição. Para estes casos é necessário, então, calcular normalmente o valor de U , a partir daí proceder à análise comparativa.

O título deste capítulo (nesta posição do trabalho) parece tratar dos resultados do capítulo anterior.

CAPÍTULO IV

Acho mais conveniente adequar com um título diferente deste a pesquisa feita pelo estudante

Resultados e sua Análise

Neste capítulo realiza-se o processo de aplicação e estimação de modelos de previsão propostos em dados de alguns Bancos de Sangue: Banco de Sangue do Hospital Central de Maputo e Banco de Sangue do Hospital Geral José Macamo; ambos na cidade de Maputo. A justificativa da escolha desses Bancos de Sangue se dá pelo facto de que apenas esses dispunham de uma quantidade de dados que possibilitasse o estudo para este trabalho, resultados da exigência de alguns métodos aqui aplicados.

Em seguida, realiza-se, a partir de uma estratégia de avaliação e selecção nas quais os erros de previsão gerados são analisados, usando um processo de determinação do modelo com menor margem de erro.

1. Análise do Comportamento das Variáveis Respostas

A colecta da série de dados teve como intervalo de tempo o período de Janeiro de 2000 a Dezembro de 2005 segundo consta no anexo A (extraído do anexo D). De forma ^{para} que os objectivos idealizados pudessem ser atingidos, dividiu-se, então, estes dados em dois conjuntos distintos de dados: *conjunto inicialização* e *conjunto teste*. O conjunto inicialização (compreendido entre Janeiro 2000 e Dezembro 2004), foi utilizado para estimar o modelo, isto é, calibração dos coeficientes e geração das previsões. Já o *conjunto teste*, compreendido entre Janeiro de 2005 e Dezembro de 2005, foi utilizado para validar o modelo, isto é, como um meio para que a estratégia de avaliação escolhida pudesse ser implementada.

Para se avaliar a consistência interna dos dados das séries foi calculado o coeficiente *alfa de Cronbach* (vide no anexo B) que resultou em $\alpha = 0.080$ para as séries do BSHCM e $\alpha = 0.3551$ para as séries do BSHGJM. As estatísticas descritivas para as séries constam na tabela 2.

Nesta secção será feita a análise dos gráficos dos dados históricos disponibilizados para a modelação com o objectivo de verificar se as características possíveis de uma série temporal (ex.: tendência, sazonalidade, aleatoriedade ou todas) é directa e fortemente observada,

facilitando assim, o processo de entendimento pelo qual determinado modelo se mostra como o mais adequado para uma determinada situação.

Tabela 2. Estatística descritiva das séries

Séries	Média	Variância	Desvio Padrão
Unidades de Sangue do BSHCM	1.293,5000	90.118,6479	300,1977
Unidades Infectadas por HIV no BSHCM	66,1389	883,2480	29,7195
Unidades de Sangue do BSHGJM	252,4861	7.364,9576	85,8193
Unidades Infectadas por HIV no BSHGJM	27,3194	230,0514	15,1674

No estudo da série de número de unidades de sangue doadas teve-se em conta os comportamentos individuais de cada grupo de doadores. A modelação individual é essencial, na medida em que os diferentes segmentos apresentam suas especificidades referentes à tendência, à sazonalidade e à ordem dos modelos. No entanto, é necessário realçar que, neste trabalho vai-se aplicar todos os métodos propostos para as diferentes séries de dados e escolher dentre eles aquele com menor margem de erro para a previsão, independentemente das características que as series apresentarem.

Para que tal análise possa ser realizada, os gráficos referentes ao conjunto inicialização das séries históricas são ilustrados a seguir:

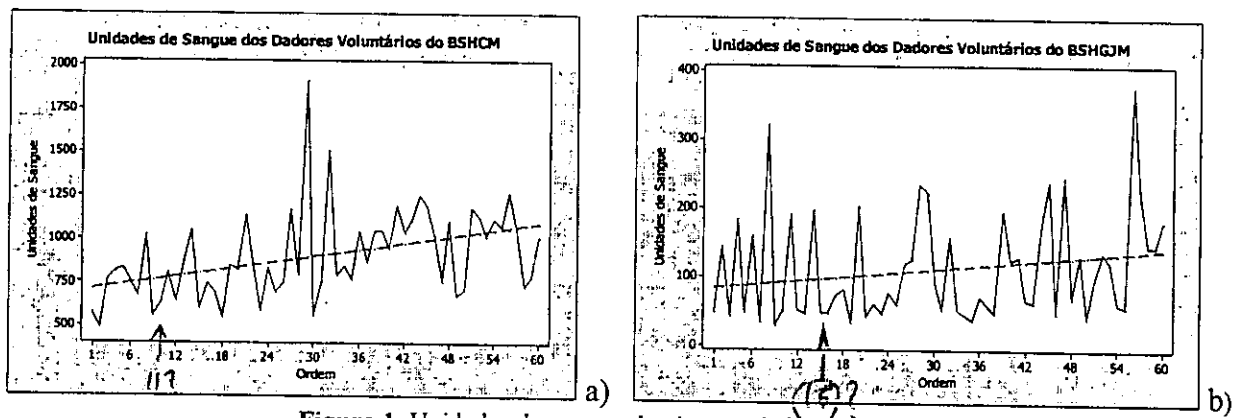


Figura 1. Unidades de sangue doadas por doadores Voluntários

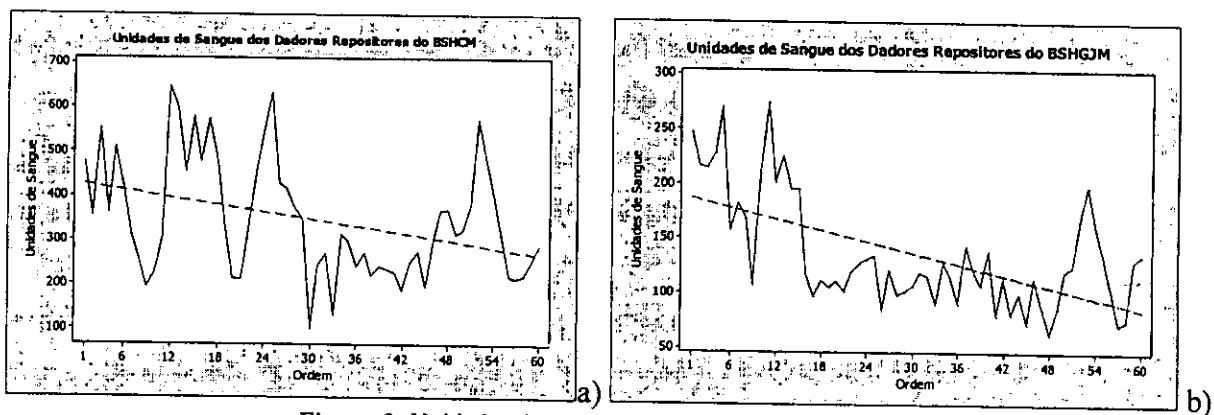


Figura 2. Unidades de sangue doadas por doadores Repositores

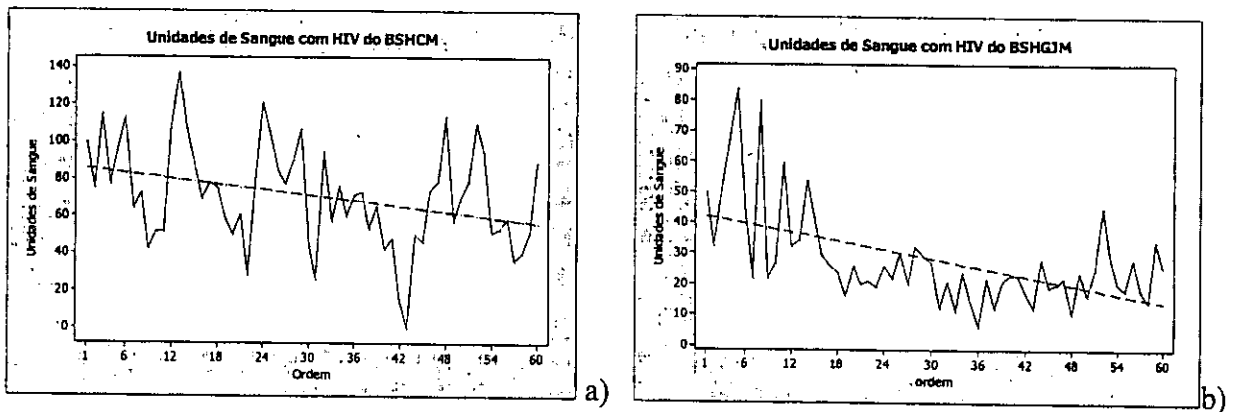


Figura 3. Unidades de sangue infectadas por HIV

1.1 Modelação via Geoestatística

Na análise dos dados usando a Geoestatística inicialmente foi definido a observação a partir da qual se deseja fazer a previsão, designado por *origem de previsão*, que foi fixada na sexagésima observação correspondente ao mês de Dezembro de 2004. A vizinhança de previsão será escolhida de tal modo que minimize o EPAM, critério que foi escolhido para a estimação dos modelos mas segundo Mingoti *et al.* (2002) valor da vizinhança ou deve ser menor ou igual ao valor escolhido como origem de previsão. Para se encontrar este valor (vizinhança de previsão óptimo) foram testadas os valores que variam de 2 a 60 e conseqüentemente escolhido o que minimiza o EPAM. O intervalo de confiança que foi usado para ajustar os modelos é de 95%.

A seguir está descrita a modelação para cada caso usando o programa PREVGE0 desenvolvido por Mingoti *et al* (2002):

A- Séries de Unidades de Sangue dos Voluntários

Refere-se à figura 1 a) e b)?

A partir da figura 1 a) e b) como verificar este resultado?

Para a série de unidades de sangue dos doadores voluntários do BSHCM, após testar todas as vizinhanças ($v = 2, \dots, 60$), concluiu-se que a vizinhança de previsão 11, é que produziu melhores resultados, isto é, para cada valor previsto e gerado são necessárias informações dos 11 valores imediatamente anteriores a origem de previsão. Com essa vizinhança obteve-se os seguintes coeficientes de previsão (lambdas):

Que expressão analítica usou para gerar os coeficientes de das tabelas 3 a 8?

Tabela 3. Coeficientes de previsão do modelo de vizinhança 11 da série de unidades de sangue dos doadores voluntários do BSHCM.

$\lambda_1 = 0.01$	$\lambda_2 = 0.04$	$\lambda_3 = 0.20$	$\lambda_4 = 0.21$	$\lambda_5 = 0.17$	$\lambda_6 = -0.21$
$\lambda_7 = 0.04$	$\lambda_8 = -0.17$	$\lambda_9 = 0.34$	$\lambda_{10} = 0.17$	$\lambda_{11} = 0.19$	

O mesmo critério foi usado para estimar os coeficientes de previsão da série de unidades de sangue dos voluntários do BSHGJM. A vizinhança que produziu melhores resultados é 16 e os valores dos coeficientes de previsão são:

Como é que se pode ver este resultado a partir do gráfico b) da figura 1?

Tabela 4. Coeficientes de previsão do modelo de vizinhança 16 da série de unidades de sangue dos dadores voluntários do BSHGJM.

$\lambda_1 = 0.244$	$\lambda_2 = 0.018$	$\lambda_3 = -0.076$	$\lambda_4 = -0.061$	$\lambda_5 = 0.353$	$\lambda_6 = 0.151$
$\lambda_7 = 0.105$	$\lambda_8 = 136$	$\lambda_9 = -0.021$	$\lambda_{10} = -0.012$	$\lambda_{11} = 0.005$	$\lambda_{12} = -0.065$
$\lambda_{13} = 0.56$	$\lambda_{14} = 0.116$	$\lambda_{15} = 0.129$	$\lambda_{16} = -0.080$		

B- Séries de Unidades de Sangue dos Repositores

A vizinhança que produz melhor resultado poderia ser ilustrada (evidenciada) no texto

Analogamente a análise anterior, foram testadas todas as vizinhanças ($v = 2, \dots, 60$). A vizinhança que melhor ajusta os dados da série foi a 12, isto é, cada valor previsto da série é gerado utilizando-se as informações dos 12 valores imediatamente anteriores ao ponto do valor que está sendo previsto. Com essa vizinhança obteve-se os seguintes coeficientes de previsão (lambdas):

Tabela 5. Coeficientes de previsão do modelo de vizinhança 12 da série de unidades de sangue dos dadores repositores do BSHCM.

$\lambda_1 = 0.194$	$\lambda_2 = -0.121$	$\lambda_3 = 0.162$	$\lambda_4 = 0.263$	$\lambda_5 = -0.199$	$\lambda_6 = 0.009$
$\lambda_7 = 0.044$	$\lambda_8 = 0.098$	$\lambda_9 = -0.145$	$\lambda_{10} = -0.086$	$\lambda_{11} = 0.127$	$\lambda_{12} = 0.653$

A mesma análise foi feita para estimar os coeficientes de previsão da série de unidades de sangue dos repositores no BSHGJM. A vizinhança que produziu melhores resultados é 3 e os valores dos coeficientes de previsão são:

Tabela 6. Coeficientes de previsão do modelo de vizinhança 3 da série de unidades de sangue dos dadores repositores do BSHGJM.

$\lambda_1 = 0.235$	$\lambda_2 = 0.127$	$\lambda_3 = 0.638$
---------------------	---------------------	---------------------

C- Séries de Unidades de Sangue infectadas por HIV

como dar crédito a esta afirmação?

Após testar todas as vizinhanças ($v = 2, \dots, 60$), a vizinhança que produziu melhores resultados para esta série foi 41, isto é, cada valor previsto é gerado utilizando-se as informações dos 41 valores imediatamente anteriores ao ponto do valor que está sendo previsto. Com essa vizinhança obteve-se os seguintes coeficientes de previsão (lambdas):

Tabela 7. Coeficientes de previsão do modelo de vizinhança 41 da série de unidades de sangue infectadas por HIV do BSHCM.

$\lambda_1 = 0.2059$	$\lambda_2 = -0.3694$	$\lambda_3 = -0.2199$	$\lambda_4 = 0.2010$	$\lambda_5 = -0.3968$	$\lambda_6 = 0.1429$
$\lambda_7 = 0.0879$	$\lambda_8 = -0.2441$	$\lambda_9 = 0.2026$	$\lambda_{10} = 0.0254$	$\lambda_{11} = 0.0185$	$\lambda_{12} = -0.1524$
$\lambda_{13} = 0.1367$	$\lambda_{14} = 0.1378$	$\lambda_{15} = 0.1112$	$\lambda_{16} = 0.3190$	$\lambda_{17} = -0.2059$	$\lambda_{18} = -0.0271$
$\lambda_{19} = 0.1940$	$\lambda_{20} = -0.1683$	$\lambda_{21} = 0.0729$	$\lambda_{22} = 0.1133$	$\lambda_{23} = -0.0219$	$\lambda_{24} = -0.0513$
$\lambda_{25} = -0.1427$	$\lambda_{26} = 0.0940$	$\lambda_{27} = -0.3258$	$\lambda_{28} = 0.2389$	$\lambda_{29} = 0.0331$	$\lambda_{30} = 0.1482$
$\lambda_{31} = 0.3759$	$\lambda_{32} = -0.0466$	$\lambda_{33} = -0.0959$	$\lambda_{34} = 0.0766$	$\lambda_{35} = 0.0632$	$\lambda_{36} = -0.1428$
$\lambda_{37} = 0.0297$	$\lambda_{38} = 0.3236$	$\lambda_{39} = -0.2175$	$\lambda_{40} = 0.1094$	$\lambda_{41} = 0.3667$	

O mesmo critério foi usado para estimar os coeficientes de previsão da série de unidades de sangue dos voluntários no BSHGJM. A vizinhança que produziu melhores resultados é 30 e os valores dos coeficientes de previsão são:

Tabela 8. Coeficientes de previsão do modelo de vizinhança 30 da série de unidades de sangue infectadas por HIV do BSHGJM.

$\lambda_1 = -0.9635$	$\lambda_2 = -0.4523$	$\lambda_3 = -0.3082$	$\lambda_4 = 0.4125$	$\lambda_5 = 0.6502$	$\lambda_6 = 0.6461$
$\lambda_7 = 0.8376$	$\lambda_8 = -0.7883$	$\lambda_9 = 0.7093$	$\lambda_{10} = 0.4126$	$\lambda_{11} = 0.1757$	$\lambda_{12} = -0.3260$
$\lambda_{13} = -0.5190$	$\lambda_{14} = -0.7805$	$\lambda_{15} = -0.8113$	$\lambda_{16} = -0.7216$	$\lambda_{17} = -0.8235$	$\lambda_{18} = -0.8268$
$\lambda_{19} = -0.5670$	$\lambda_{20} = -0.4078$	$\lambda_{21} = 0.1102$	$\lambda_{22} = 0.3383$	$\lambda_{23} = 0.6648$	$\lambda_{24} = 0.7612$
$\lambda_{25} = 0.8310$	$\lambda_{26} = 0.6602$	$\lambda_{27} = 0.6759$	$\lambda_{28} = 0.4824$	$\lambda_{29} = -0.2455$	$\lambda_{30} = -0.4033$

1.2 Modelação por Método Holt-Winters

Como já foi referido, o Método de Holt-Winters não exige qualquer manipulação dos dados originais para que, por exemplo, respeitem uma distribuição normal e/ou estejam estacionados. Portanto, sua aplicação resume-se apenas à utilização das fórmulas mencionadas e apresentadas no capítulo anterior. A aplicação deste método nos dados de doações de sangue e de unidades infectadas por HIV dos Hospitais Central de Maputo e Geral José Macamo foi realizada no aplicativo Microsoft Excel e os valores dos coeficientes de ajuste (α , β , e γ), foram estimados e escolhidos segundo o critério de minimização do valor do EPAM do conjunto de inicialização com o SOLVER, um programa de optimização contido no MS Excel.

Na modelação dos dados do conjunto de inicialização da série temporal de unidades infectadas por HIV do BSHCM, verificou-se que contém um valor igual a zero (observação do mês de Julho de 2003); razão pela qual para a estimação de valores de α , β e γ obedeceu o mesmo critério que os outros com a única diferença do denominador da fórmula do EPAM que é o *valor previsto* e não o *valor observado*. Esta alteração é defendida por Samohyl (2004) na medida em que o objectivo é produzir melhores resultados.

Usando o SOLVER obteve-se os resultados tabelados a seguir (tabelas 9 e 10):

Tabela 9. Valores dos Coeficientes do método Holt-Winters para as séries do BSHCM

HOSPITAL CENTRAL DE MAPUTO						
Série de Temporal	Alfa (α)		Beta (β)		Gama (γ)	
	H.W.A.	H.W.M.	H.W.A.	H.W.M.	H.W.A.	H.W.M.
Unidades dos Voluntários	0,00000	0,00000	0,07508	0,07803	0,18350	0,25318
Unidades dos Repositores	0,39091	0,41665	0,00211	0,09136	1,00000	0,78723
Unidades infectadas por HIV	0,00000	0,00000	0,00000	0,00000	0,08063	0,08769

Tabela 10. Valores dos Coeficientes do método Holt-Winters para as séries do BSHGJM

HOSPITAL GERAL JOSÉ MACAMO						
Série de Temporal	Alfa (α)		Beta (β)		Gama (γ)	
	H.W.A.	H.W.M.	H.W.A.	H.W.M.	H.W.A.	H.W.M.
Unidades dos Voluntários	0,022288	0,00116	0,06494	1,00000	0,97640	0,92108
Unidades dos Repositores	0,198125	0,22938	0,00000	0,00932	1,00000	0,17759
Unidades infectadas por HIV	0,28121	0,04361	0,03006	0,46277	0,92398	0,15539

1.3 Modelação por Método de Box-Jenkins

Como foi referenciado no item 4.5.5 do capítulo (II) a primeira etapa deste método é a identificação do modelo. Para este trabalho a modelação das séries foi feita inicialmente pela análise dos gráficos de tempo das séries em estudo.

A análise de gráfico ^{que análise é esta} pode indicar a presença de tendência ou alteração na variância, o que revelaria se a série é ou não estacionária. Esta análise vem a ser reforçada pelo teste de raiz unitária ADF (Augmented Dickey-Fuller) utilizando o aplicativo [WEB:REG] UNIT ROOT TEST (ADF-TEST)¹³. Para a aplicação deste teste, deve-se definir, primeiro, o número óptimo da defasagem temporal da série pelo critério AIC. Caso seja validada a hipótese de não estacionariedade pelo teste ADF, a série temporal z_t deve ser diferenciada tantas vezes quantas forem necessárias, até que se torne estacionária, alias, a diferenciação vai determinar o parâmetro d do modelo ARIMA(p, d, q).

A estimação de outros parâmetros (p e q) é feita a partir dos gráficos ACF e PACF donde são identificados vários modelos candidatos. Usando o pacote estatístico SPSS 11 foi possível seleccionar o modelo mais adequado aos dados de cada série com base no critério do menores valores de AIC, SBC e variância dos resíduos. Se existirem dois ou mais modelos candidatos com critérios mencionados muito próximos é introduzido o diagnóstico usando o critério do menor valor do EPAM.

Por ultimo são analisados os correlogramas dos resíduos se estes caracterizam o ruído branco. Caso contrário volta-se a incorporar outros modelos candidatos até que se encontre um, cujo resíduos são considerados ruído branco.

¹³ Disponível em <http://www.web-reg.de/> [extraído no dia 23/08/2006]

1.3.1 Modelação das Séries de Unidades de Sangue dos Voluntários

Quanto ao número de unidades de sangue dos doadores voluntários do BSHCM assim como do BSHGJM representadas na figura 1 apresentam fraca tendência crescente. Sendo assim, elas podem ser consideradas estacionárias. Esta suposição vem a ser reforçada pelo teste ADF feito com o modelo menos restrito (que inclui tendência e intercepto) a hipótese de não estacionaridade foi rejeitada, como se pode observar nas tabelas 11 e 12. Com isso as séries podem ser consideradas estacionárias, isto é, a ordem de integração é zero ($d = 0$).

Tabela 11. Parte do Teste ADF para a série de unidades de sangue dos voluntários do BSHCM

ADF Test Statistic	-8,401586	1% Critical Value	-4,121324
		5% Critical Value	-3,487851
		10% Critical Value	-3,172335

Tabela 12. Parte do Teste ADF para a série de unidades de sangue dos voluntários do BSHGJM

ADF Test Statistic	-8,479260	1% Critical Value	-4,121324
		5% Critical Value	-3,487851
		10% Critical Value	-3,172335

Devem ser também analisadas as respectivas funções de autocorrelações (ACF) e de autocorrelações parciais (PACF), como ilustradas nas figuras abaixo.

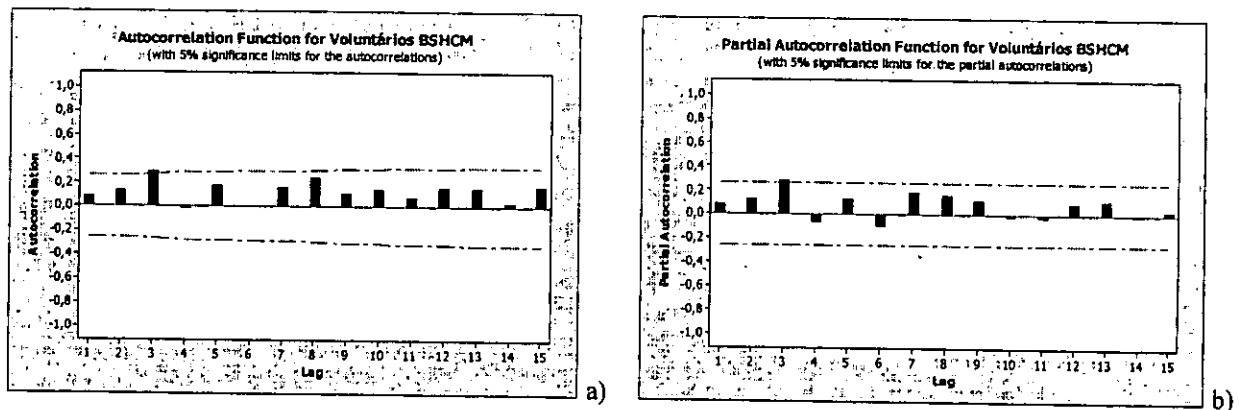


Figura 4. a) função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades dos doadores voluntários do BSHCM

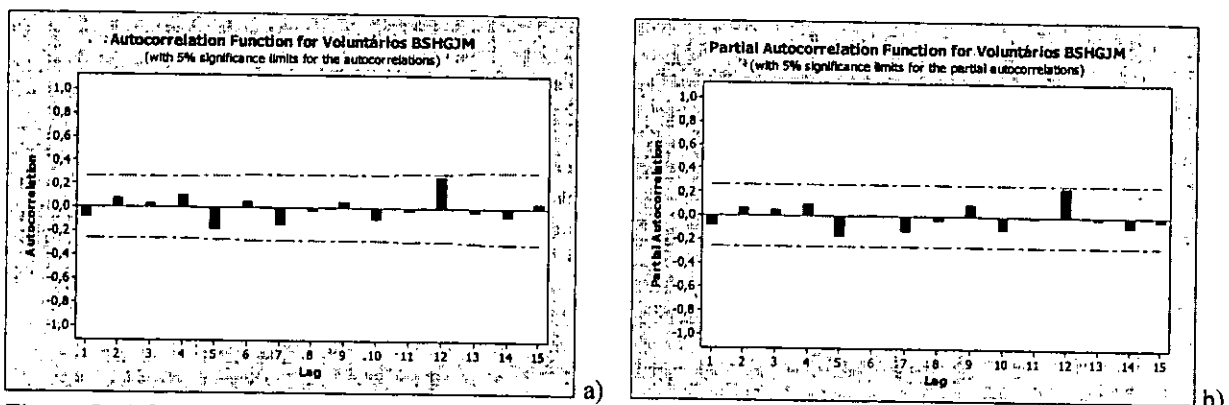


Figura 5. a) função de autocorrelação e b) função de autocorrelação parcial para o número de unidades dos doadores voluntários do BSHGJM

Para estimar outros parâmetros do modelo da série unidades de voluntários do HCM foram analisadas¹⁴ as funções de autocorrelações (ACF) e de autocorrelações parciais (PACF) donde se obteve vários modelos candidatos e com ajuda do pacote estatístico SPSS 11, optando-se pelo método de *forecasting de mínimos quadrados incondicional*, foi identificado o ARMA (3,0,2) que descreve melhor a série, por este ser o que produziu menores valores de AIC, SBC e variância residual, bem como EPAM.

No caso da série de unidades de voluntários do HGJM foi considerada a parte sazonal pelo facto das funções ACF e PACF apresentarem significância no *lag* sazonal 12. O modelo que produziu melhores valores de AIC e SBC e variância residual e conseqüente EPAM é o SARIMA (0,0,0)(1,0,0)₁₂.

A validação dos modelos apresentados acima foi feita mediante o teste dos resíduos. De acordo com a figura 6, os resíduos do modelo ARIMA(3,0,2) podem ser considerados de ruído branco. O mesmo acontece para os resíduos do modelo SAR(1)₁₂ mostrados na figura 7.

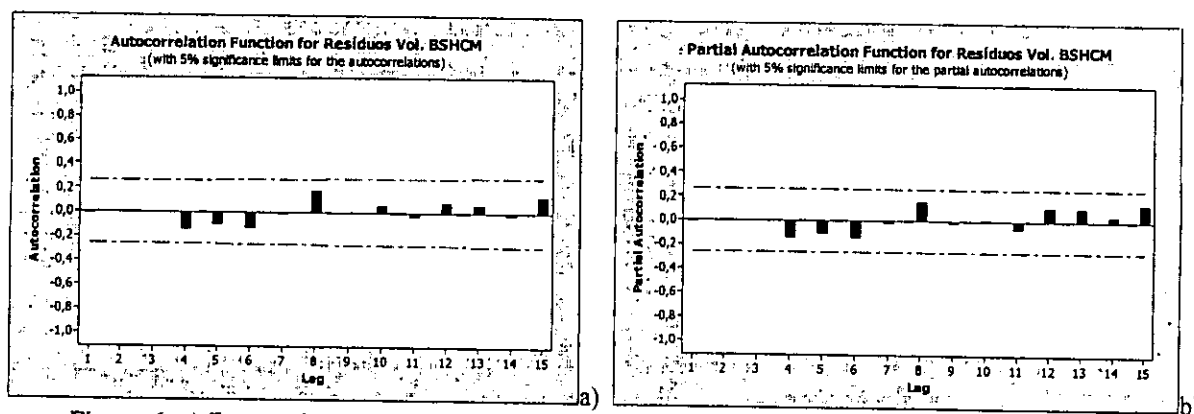


Figura 6. a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo ARMA(3,0,2)

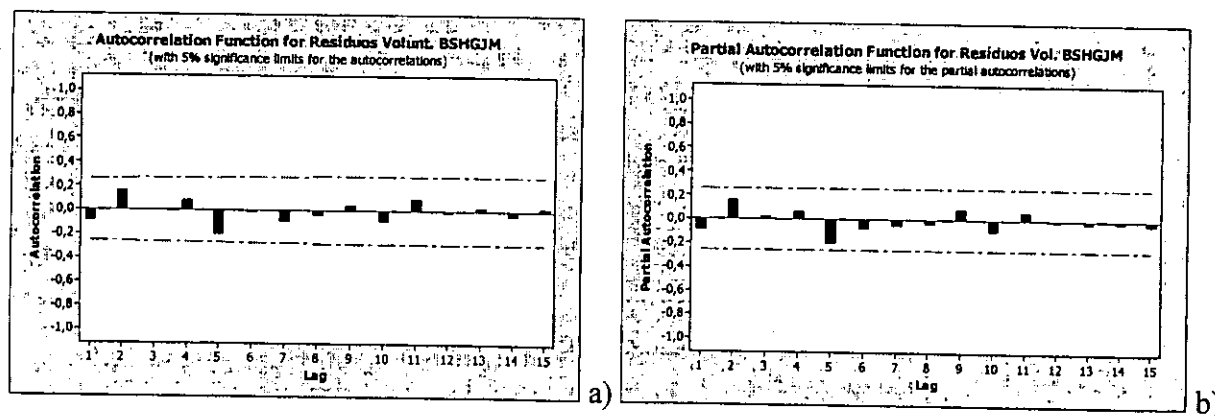


Figura 7. a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo SARIMA(0,0,0)(1,0,0)₁₂

¹⁴ Para a determinação do número de *lags* foi usado o critério descrito na secção 4.5.3 do capítulo III

1.3.2 Modelação das Séries de Unidades de Sangue dos Repositores

Para identificar os modelos apropriados, inicialmente, foi analisada a figura 2, que apresenta os gráficos de tempo do número de unidades de sangue dos dadores repositores. A análise do gráfico da figura 2a) indica a presença de tendência decrescente acentuada, indícios de série não estacionária. As funções ACF e PACF da figura 8 apresentam também indícios de uma série não-estacionária. Para tal se confirmar foi realizado o teste ADF de raízes unitárias. O resultado mostrado na tabela 13, revela que, para qualquer nível de significância, não foi rejeitada a hipótese nula, H_0 , isto é, a série possui raiz unitária, ou seja, não é estacionária.

Tabela 13. Parte do Teste ADF para a série de unidades de sangue dos dadores repositores do BSHCM

ADF Test Statistic	-1,822205	1% Critical Value	-4,165722
		5% Critical Value	-3,508499
		10% Critical Value	-3,184239

Para que a série da figura 2a) se tornar estacionária foi necessário transformar os dados achando a sua primeira diferença. Testes feitos aos possíveis modelos, constatou-se que seria necessário considerar a parte sazonal e assim o modelo que melhor interpreta os dados da série é o $ARIMA(0,1,0)(1,1,0)_{12}$ cujos correlogramas dos resíduos estão apresentados na figura 9.

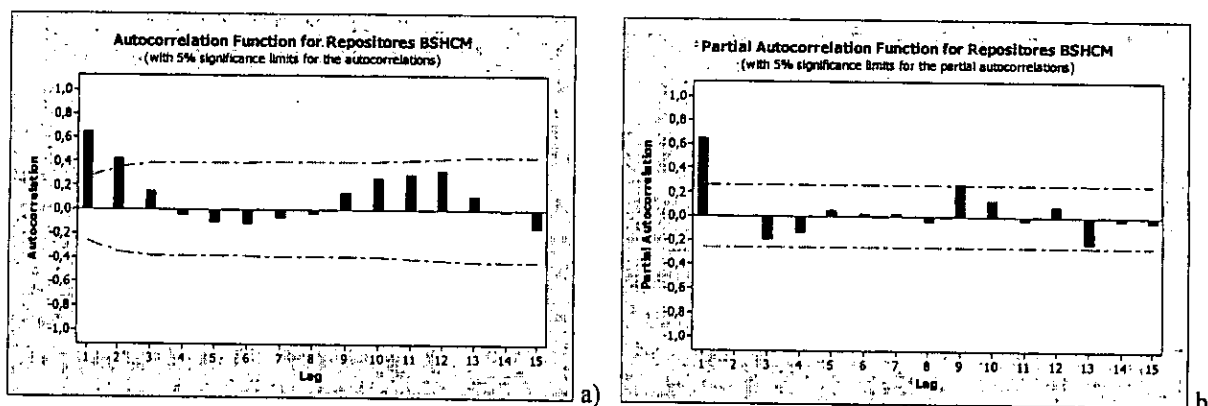


Figura 8. a) Função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades dos dadores Repositores do BSHCM

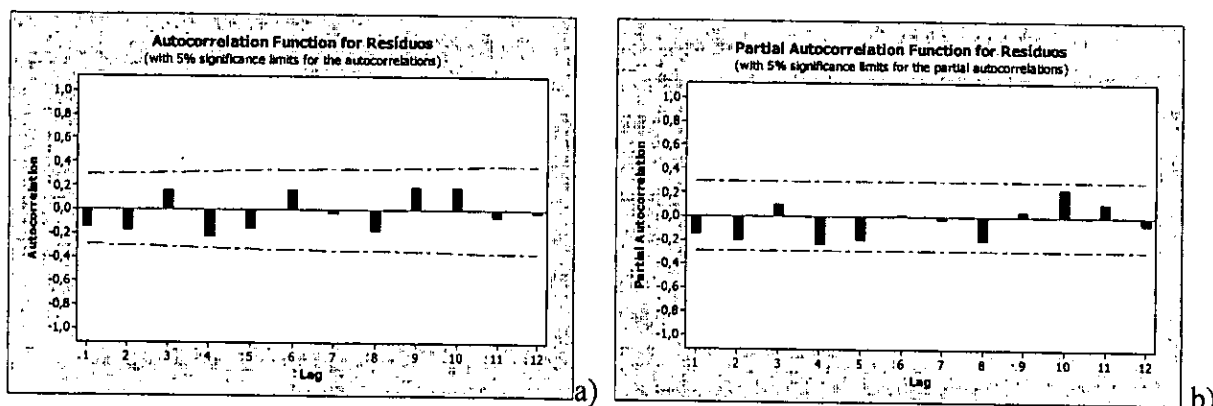


Figura 9. a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo $SARIMA(0,1,0)(1,1,0)_{12}$

A mesma análise foi feita para a série de unidades de sangue dos doadores repositores do BSHGJM donde se constatou que esta série é estacionária a todos os níveis de significância, veja a tabela 14.

Tabela 14. Parte do Teste ADF para a série de unidades de sangue dos doadores repositores do BSHGJM

ADF Test Statistic	-3,724301	1% Critical Value	-3,588543
		5% Critical Value	-2,929737
		10% Critical Value	-2,603087

A figura 10a) representa a ACF do número de unidades dos doadores repositores do HGJM. Pode-se observar ainda que a ACF decai exponencialmente e na PACF, figura 10b), pode-se constatar que a autocorrelação do lag 1 é mais significativa, indicando assim o modelo AR(1). Mas existe mais um lag significativo na defasagem 6, por isso os dois devem ser testados de modo a se identificar o melhor que interpreta os dados da série. O teste culminou com a identificação do ARIMA(6,0,0) por apresentar menores valores de AIC, variância dos resíduos e EPAM.

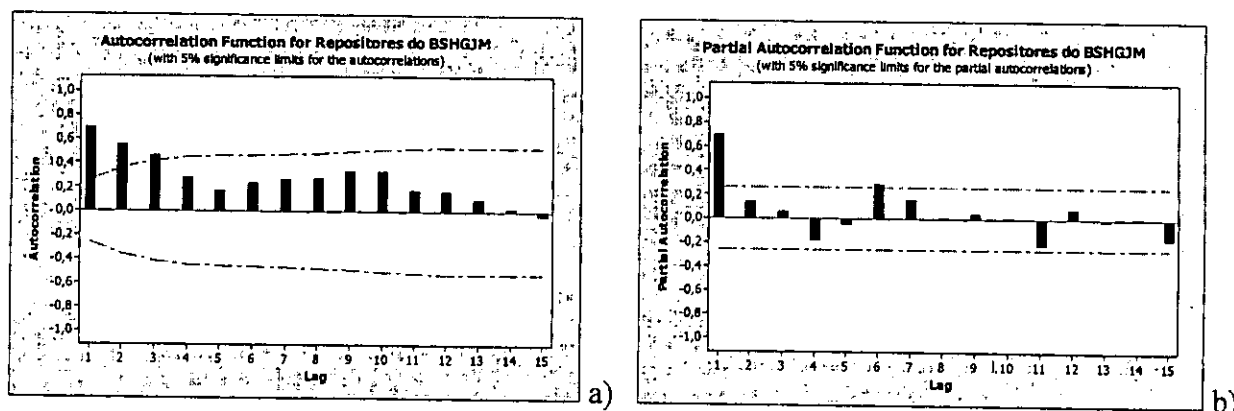


Figura 10. a) Função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades dos doadores Repositores do BSHGJM

Para a validação do modelo AR(6) foram construído a ACF e PACF dos resíduos, figura 11, de onde se constatou que as mesmas representam um ruído branco.

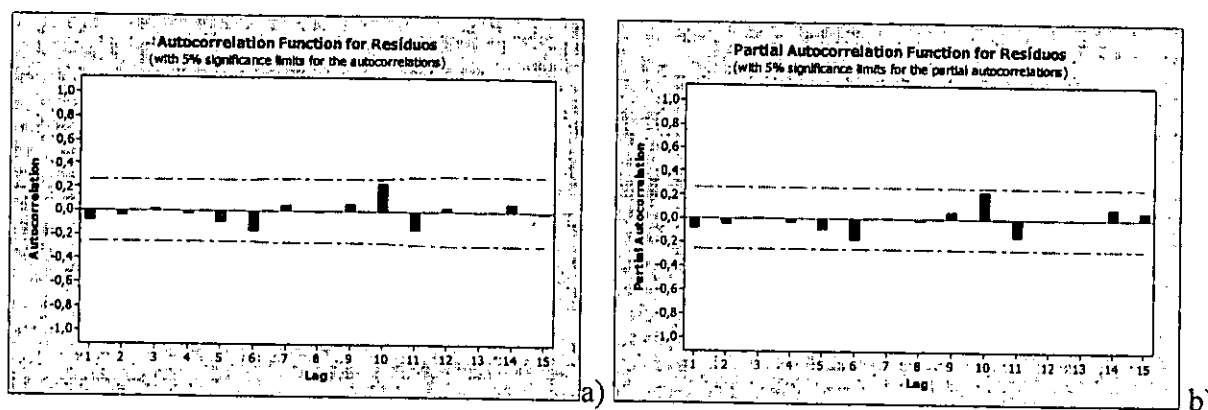


Figura 11. a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo ARIMA(6,0,0)

1.3.3 Modelação das Séries de Unidades de Sangue Infectadas por HIV

A série de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHCM apresenta fraca tendência decrescente que revela que a série provavelmente é estacionária. Esta estacionaridade se confirma após a aplicação do teste ADF sobre esta série, como pode-se confirmar na tabela 15. Os autocorrelações representadas na figura 12 indicam um modelo auto regressivos AR(1), mas ainda pode-se notar indícios de sazonalidade; razão pela qual modelos sazonais também foram testados e o que melhor os dados da série é o SARIMA (1,0,0)(1,1,0)₁₂. Como pode-se notar, houve necessidade de aplicar a diferença sazonal que permitiu a remoção da sazonalidade da série.

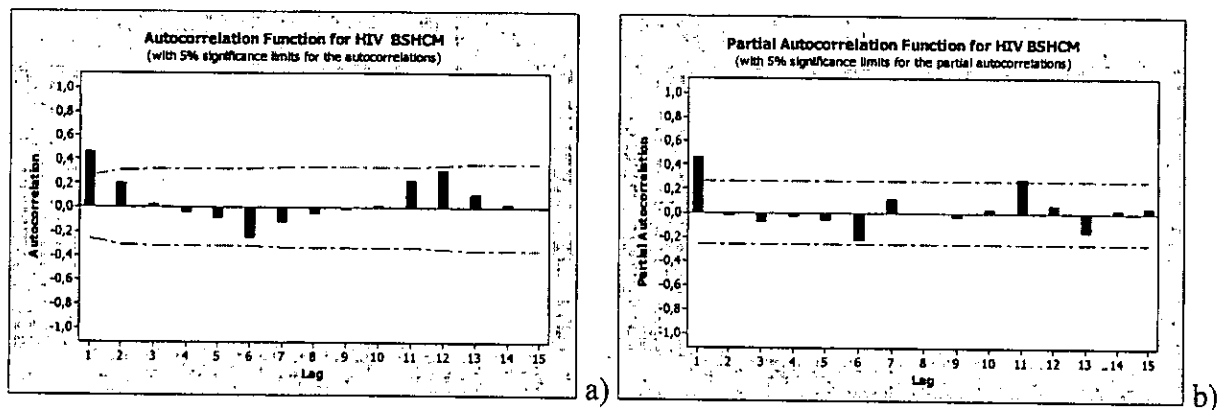


Figura 12. a) Função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades de sangue inutilizadas por infecção com o vírus do HIV do BSHCM

Tabela 15. Parte do Teste ADF para a série de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHCM

ADF Test Statistic	-4,798177	1% Critical Value	-4,121324
		5% Critical Value	-3,487851
		10% Critical Value	-3,172335

Para a validação deste modelo foram construídos os gráficos de Autocorrelações, figura 13, dos resíduos produzidos pelo modelo e constatou-se que estes não possuem lags significativos, confirmando-se assim que representam ruído branco.

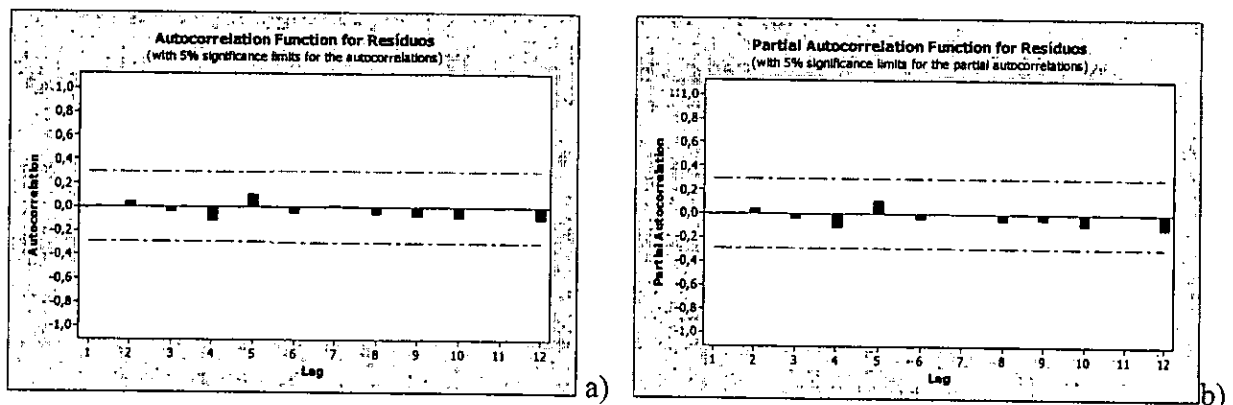


Figura 13. a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo SARIMA (1,0,0)(1,1,0)₁₂

Tabela 16. Parte do Teste ADF para a série de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHGJM

ADF Test Statistic	-2,741594	1% Critical Value	-4,170543
		5% Critical Value	-3,510729
		10% Critical Value	-3,185520

Mesma análise foi feita para a série de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHGJM donde com ajuda das funções de autocorrelações (figura 14) e teste ADF (tabela 16) verificou-se que a série não é estacionária; isto também pode observar-se no gráfico da figura 3b) onde a linha de tendência tem um declive relativamente maior.

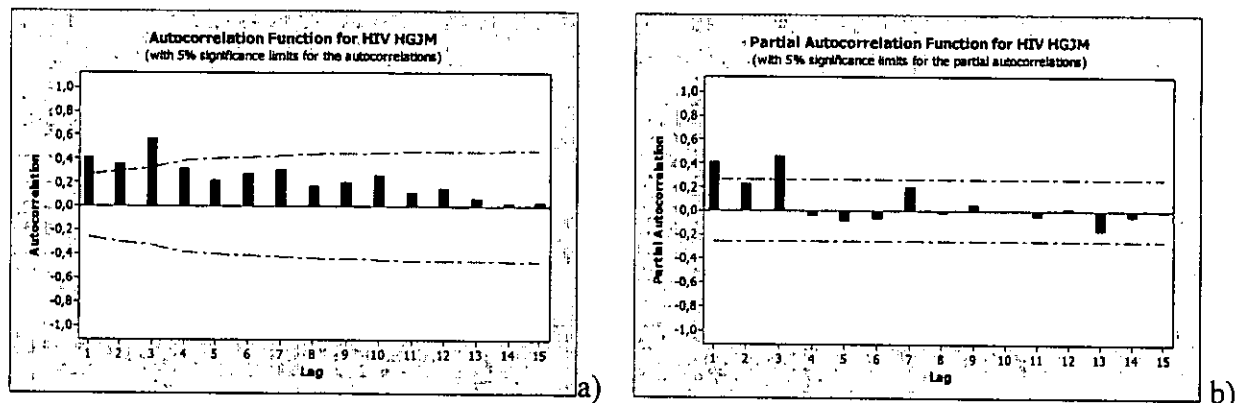


Figura 14. a) Função de Autocorrelação e b) função de autocorrelação parcial para o número de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHGJM

O gráfico da figura 15 representa a série transformada, após a primeira diferença, nela nota-se que os dados oscilam em torno da média zero, indicando deste modo que a série transformada é estacionária. Isto pode ser confirmado pelo teste ADF mostrado na tabela 17 abaixo:

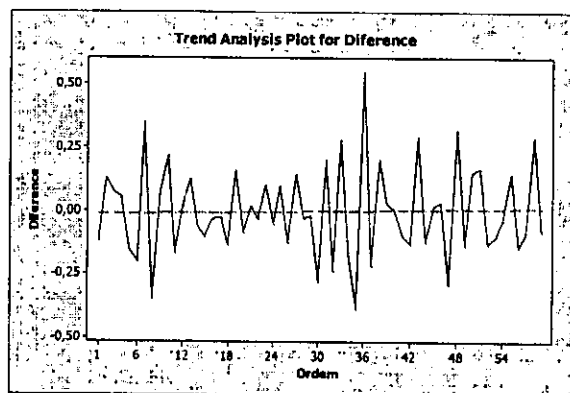


Figura 15. Gráfico da série da figura 3b) a pós a sua primeira diferença

Tabela 17. Parte do Teste ADF para a série transformada de unidades de sangue inutilizadas por infecção com o vírus do HIV no BSHCM

ADF Test Statistic	-4,839016	1% Critical Value	-4,140856
		5% Critical Value	-3,496960
		10% Critical Value	-3,177596

O modelo que resultou ruído branco representado na figura 16 é o ARIMA (2,1,0).

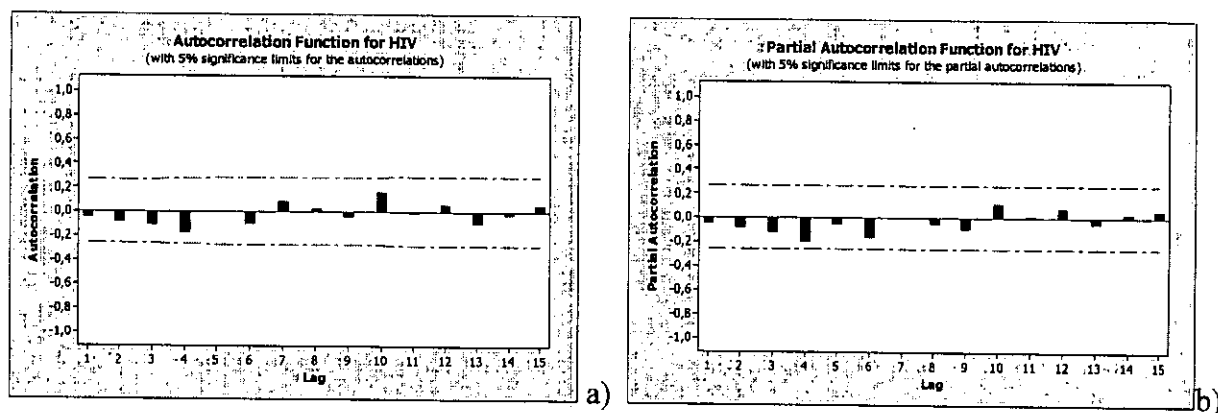


Figura 16. a) Função de Autocorrelação e b) função de autocorrelação parcial dos resíduos para o Modelo ARIMA (2,1,0)

CAPÍTULO V

Discussão dos Resultados

Neste capítulo, os modelos seleccionados no capítulo anterior são implementados para se avaliar o desempenho de cada um deles e conseqüentemente o seu método. Para tal avaliação considera-se a parte dos dados designada de conjunto teste que por meio da comparação do valor correspondente a estatística U de Theil. Entre os modelos aprovados foi seleccionado o que apresentou menor valor desta estatística. Os valores da estatística U de Theil estão apresentados em tabelas onde também podemos encontrar outros critérios de comparação tais como EAM, EQM e EPAM provocados pela discrepância das previsões. São também apresentados gráficos de linhas que ilustram claramente estas discrepâncias só para os modelos dos métodos em estudo. Os valores gerados pelos modelos seleccionados estão apresentados no anexo C.

1. Unidades de Sangue dos Voluntários do BSHCM

Na tabela 18 são mostrados os critérios de selecção do modelo mais apropriado. Como foi anteriormente referido, o critério usado para a selecção do modelo mais apropriado aos dados da série das unidades de sangue dos doadores voluntários do Banco de Sangue do Hospital Central de Maputo (HCM) e do Hospital Geral José Macamo (HGJM), é a estatística U de Theil.

Tabela 18. Comparação entre os diferentes modelos aplicados a série da figura 1a)

MODELOS	CRITÉRIOS			
	EAM	EQM	EPAM	U de Theil
Geoestatística	247,630	90346,520	19,046857	1,040787810
Holt-Winters Aditivo	247,667	78030,000	21,202605	1,07696732
Holt-Winters Multiplicativo	258,167	82976,000	22,634322	1,15396310
Box-Jenkins	296,572	134891,951	22,528740	1,26541180

Como se pode observar na coluna do critério U de Theil da tabela 18, o modelo de vizinhança 11 correspondente a Geoestatística é que apresenta o menor valor de U de Theil ($U = 1,04078781$).

Eis as equações usadas para o cálculo dos critérios patentes na tabela 18 e as restantes a seguir:

$$\text{Erro Absoluto Médio (EAM)} = \frac{1}{12} \sum_{t=61}^{72} |z_t - \hat{z}_t| \quad (81)$$

$$\text{Erro Quadrático Médio (EQM)} = \frac{1}{12} \sum_{t=61}^{72} (z_t - \hat{z}_t)^2 \quad (82)$$

$$\text{Erro Percentual Absoluto Médio (EPAM)} = \frac{1}{12} \sum_{t=61}^{72} \left| \frac{(z_t - \hat{z}_t)}{z_t} \times 100 \right| \quad (83)$$

$$\text{Estatística U de Theil (U)} = \sqrt{\frac{\sum_{t=61}^{71} \left(\frac{z_{t+1} - \hat{z}_{t+1}}{z_t} \right)^2}{\sum_{t=61}^{71} \left(\frac{z_{t+1} - z_t}{z_t} \right)^2}} \quad (84)$$

Onde: z_t é a valor observado no instante t e,
 \hat{z}_t é o valor previsto para o instante t .

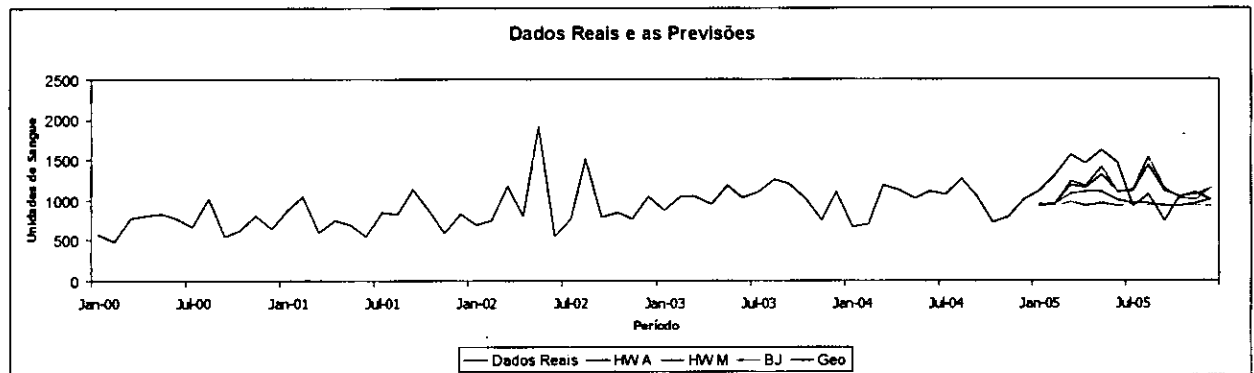


Figura 17. Previsões para as unidades de sangue dos voluntários do BSHCM (Onde: BJ – Box-Jenkins; Geo – Geoestatística; HWM – Holt-Winters Multiplicativo e HWA – Holt-Winters Aditivo)

No gráfico da figura 17 pode-se visualizar as diferenças entre os dados reais e as previsões proporcionadas pelos modelos implementados. Tendo em conta a significância dos valores de U pode-se notar também que o modelo aditivo é o segundo classificado.

2. Unidades de Sangue dos Dadores Voluntários do BSHGJM

Identicamente ao que foi feito no ponto 1 deste capítulo, a tabela 19 contém informações relevantes para a escolha do modelo mais apropriado para esta série.

Tabela 19. Comparação entre os diferentes modelos aplicados a série da figura 1b)

MODELOS	CRITÉRIOS			
	EAM	EQM	EPAM	U de Theil
Geoestatística	42,020	3692,669	26,044646	0,664606860
Holt-Winters Aditivo	74,583	9736,250	49,682308	0,967576822
Holt-Winters Multiplicativo	77,000	9775,167	51,897147	0,952650149
Box-Jenkins	48,135	5322,370	30,658642	0,803141596

Qual é o critério de uso de casas decimais?

Tendo em conta critério usado (U de Theil), verifica-se que o modelo mais eficiente para esta série é o de Geoestatística visto que apresenta menor valor da estatística U de Theil ($U = 0,664606860$).

EQM e EAM tem 3 casas decimais, mas EPAM tem 6 casas decimais e U tem 9 casas decimais!

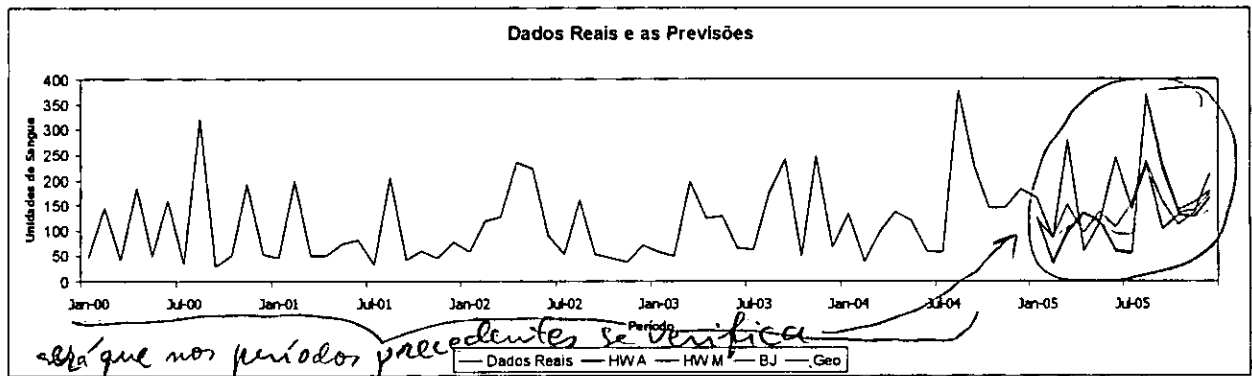


Figura 18. Previsões para as unidades de sangue dos voluntários do BSHGJM

Na ilustração do gráfico figura 18 pode-se notar para os modelos representados que o de geoestatística apresenta melhor ajuste que os outros. Daí conclui-se que para esta série é conveniente, segundo a estatística U, usar-se geoestatística para gerar previsões futuras.

3. Unidades de Sangue dos Dadores Repositores do BSHCM

Seguindo os mesmos procedimentos que das outras séries foi produzida a tabela 20 contendo a informação específica para a comparação dos modelos. Como se pode notar, o modelo de vizinhança 12 correspondente a metodologia de geoestatística é que está no topo da classificação ($U = 0,7294543769$) com uma diferença considerável em comparação a outros modelos.

10 casas decimais?

O mesmo pode-se notar com ajuda do gráfico da figura 19, *a seguir,* que a linha verde é a que melhor se aproxima ^{da} realidade (linha preta), colocando assim mais uma vez a Geoestatística no topo da classificação para as séries em estudo.

Tabela 20. Comparação entre os diferentes modelos aplicados a série da figura 2a)

MODELOS	CRITÉRIOS			
	EAM	EQM	EPAM	U de Theil
Geoestatística	35,306	1733,720	12,461093	0,7294543769
Holt-Winters Aditivo	69,397	5966,994	21,136358	1,116352064
Holt-Winters Multiplicativo	99,917	12288,917	28,533980	1,527897734
Box-Jenkins	112,054	15547,585	34,843598	1,930630108

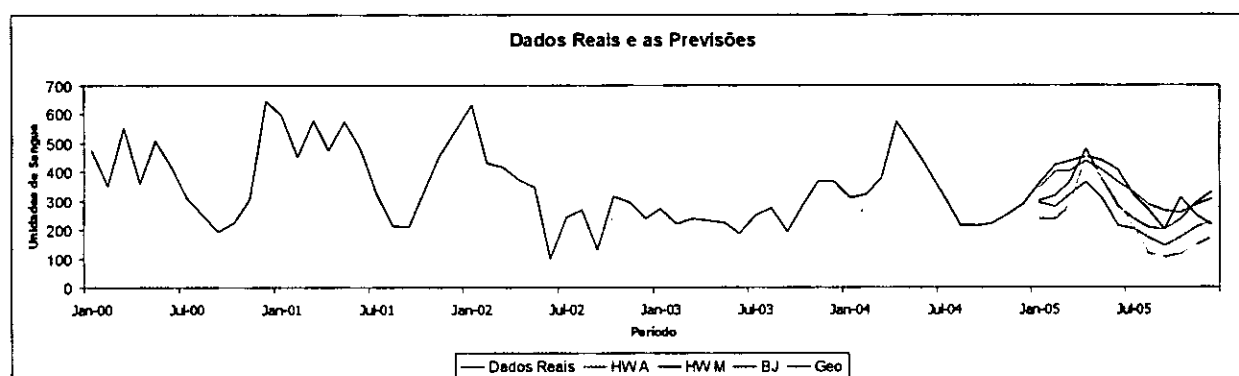


Figura 19. Previsões para as unidades de sangue dos repositores do BSHCM

4. Unidades de Dadores Repositores do BSHGJM

Na tabela 21 está patente a informação referente aos critérios de comparação dos modelos apresentados.

Tabela 21. Comparação entre os diferentes modelos aplicados a série da figura 2b)

MODELOS	CRITÉRIOS			
	EAM	EQM	EPAM	U de Theil
Geoestatística	24,732	865,291	22,616700	0,58783821
Holt-Winters Aditivo	43,000	2778,167	38,600130	1,306439231
Holt-Winters Multiplicativo	45,917	2388,250	38,597138	1,178772222
Box-Jenkins	32,173	1477,786	31,909121	0,83956483

Interpretando os valores apresentados, pode-se seleccionar o modelo de vizinhança 3 também correspondente a metodologia de geoestatística que atingiu o valor de $U = 0,58783821$.

Como se pode observar no gráfico da figura 20, a linha correspondente ao melhor ajuste da série tende a ser rectilínea assemelhando-se a linha que descreve a média dos valores do conjunto teste da série.

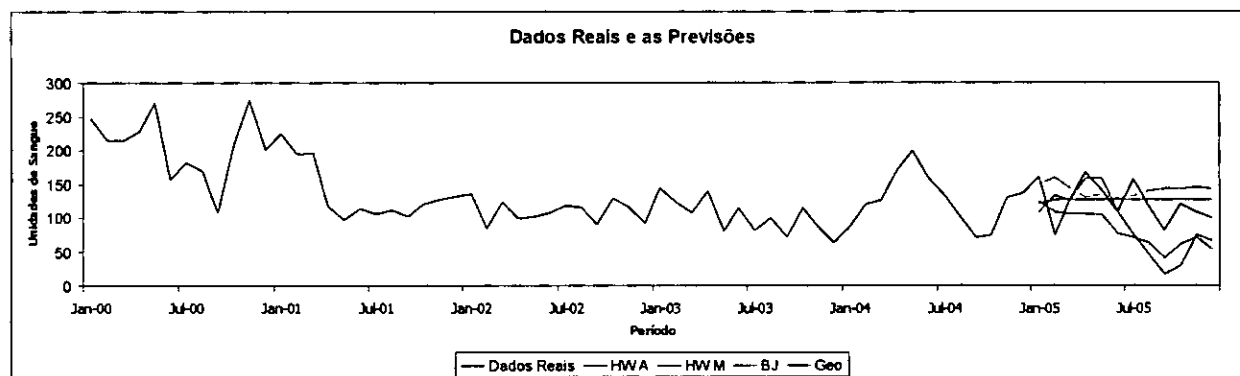


Figura 20. Previsões para as unidades de sangue dos repositores do BSHGJM

5. Número de Unidades de Sangue Colectadas no BSHCM

Para se obter os valores da tabela 22 foi feito o somatório das previsões resultantes em cada modelo para as unidades de sangues dos dadores voluntários e dos repositórios e posteriormente cálculos os valores dos critérios apresentados usando as mesmas fórmulas descritas no item 1 deste capítulo.

O método que mostrou ser o mais excelente em termos de previsões para a série de *número de unidades de sangue colectados no Banco de Sangue do Hospital Central de Maputo* como já era esperado é o de geoestatística. Segundo mostra a tabela 22 este método produz o valor da estatística U de Theil menor ($U = 0,920152029$) que os outros métodos.

Tabela 22. Comparação entre os diferentes modelos aplicados a série de Número de Unidades de Sangue Colectados no BSHCM

MODELOS	CRITÉRIOS			
	EAM	EQM	EPAM	U de Theil
Geoestatística	274,503	105630,059	16,746163	0,920152029
Holt-Winters Aditivo	283,513	101592,955	18,565085	0,964550153
Holt-Winters Multiplicativo	313,083	118974,583	20,257250	1,041449661
Box-Jenkins	387,635	204418,714	23,136444	1,328883878

O gráfico da figura 21 apresenta os valores reais e as previsões resultantes de cada método aplicado. E como se pode observar, a seguir de Geoestatística, o Holt-Winters de sazonalidade aditiva é o que possivelmente pode substituir a geoestatística, visto que as suas previsões provocam uma discrepância medida pela estatística U cujo valor está mais próximo do U de geoestatística.

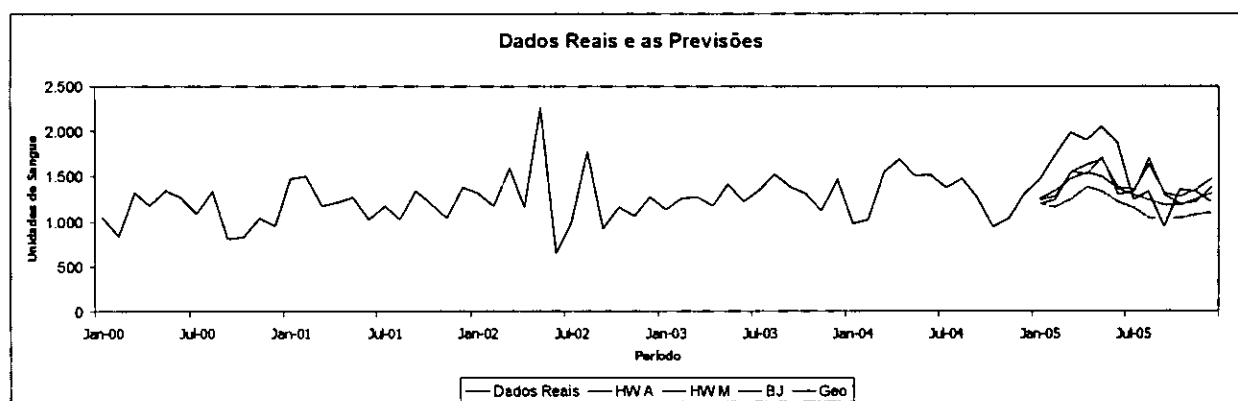


Figura 21. Previsões para as unidades de Sangue Colectadas no BSHCM

6. Número de Unidades de Sangue Colectadas no BSHGJM

Analogamente ao que fizemos para as unidades colectadas no BSHCM segue-se abaixo a tabela contendo a informação relevante para o processo de escolha do modelo para as *unidades de sangue colectadas no BSHGJM* seja realizado.

Tabela 23. Comparação entre os diferentes modelos aplicados a série de Número de Unidades de Sangue Colectadas no BSHGJM

MODELOS	CRITÉRIOS			
	EAM	EQM	EPAM	U de Theil
Geoestatística	52,917	4685,750	19,615193	0,560106892
Holt-Winters Aditivo	80,917	10413,417	26,951714	0,83696813
Holt-Winters Multiplicativo	81,750	11669,250	26,965879	0,889575137
Box-Jenkins	60,833	6081,000	23,322590	0,649215331

Tendo como base o critério em uso (estatística U de Theil), verifica-se que o modelo com maior grau de eficiência dentre os modelos identificados é o modelo de geoestatística com valor de $U = 0,560106892$.

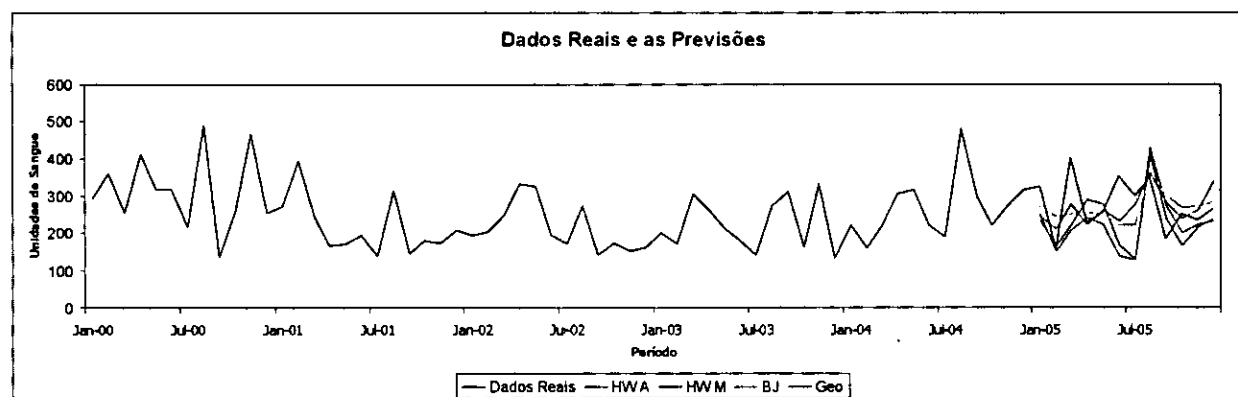


Figura 22. Previsões para as unidades de Sangue Colectadas no BSHGJM

No gráfico da figura 22, pode-se visualizar as diferenças entre os dados reais e os previstos proporcionados pelo modelo vencedor e outros que são referenciados como objectos de estudo deste trabalho. Por meio da análise do gráfico pode-se concluir que dada a suposição de continuidade dos dados existe um potencial de ganho com o método de Geoestatística para a geração de previsões para um horizonte aplicado no trabalho (12 meses).

7. Número de Unidades de Sangue Infectadas pelo Vírus HIV Colectadas no BSHCM.

Na tabela 24 são mostrados os valores dos critérios de selecção dos modelos para esta série.

Tabela 24. Comparação entre os diferentes modelos aplicados a série da figura 3a)

MODELOS	CRITÉRIOS			
	EAM	EQM	EPAM	U de Theil
Geoestatística	14,527	273,024	53,107172	0,540655857
Holt-Winters Aditivo	36,167	1807,333	181,261955	1,373741608
Holt-Winters Multiplicativo	35,750	1765,083	181,055353	1,365685070
Box-Jenkins	28,926	1190,908	139,536140	1,245024336

Como se pode observar na coluna do critério U de Theil da tabela, o modelo de vizinhança 41 correspondente a Geoestatística é que apresenta o menor valor de U de Theil ($U = 0,540655857$).

No gráfico da figura 23 pode-se visualizar as diferenças entre os dados reais e os previstos proporcionados pelos modelos implementados. Tendo em conta a estatística U de Theil pode-se notar que o modelo aditivo de Holt-Winters é o segundo classificado.

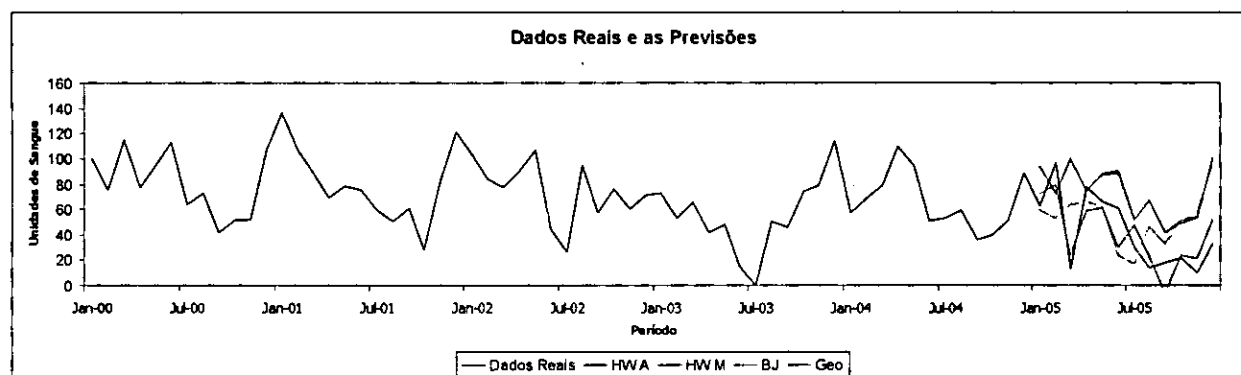


Figura 23. Previsões para as unidades de sangue das unidades de sangue infectados pelo Vírus HIV do BSHCM

8. Número de Unidades de Sangue Infectadas pelo Vírus HIV Colectadas no BSHGJM.

Seguindo os mesmos procedimentos que das outras séries foi produzida a tabela 25 contendo a informação específica para a comparação dos modelos. Como se pode notar o modelo ARIMA (2,1,0) é que está no topo da classificação ($U = 0,743554583$).

Tabela 25. Comparação entre os diferentes modelos aplicados a série da figura 3b)

MODELOS	CRITÉRIOS			
	EAM	EQM	EPAM	U de Theil
Geoestatística	9,099	156,373	35,872354	1,0712140
Holt-Winters Aditivo	10,333	153,667	52,651289	1,062808279
Holt-Winters Multiplicativo	17,667	379,167	98,673658	1,794094999
Box-Jenkins	8,160	150,129	46,169887	0,743554583

O mesmo pode-se notar com ajuda do gráfico da figura 24 que a linha que representa o modelo ARIMA(2,1,0) é a que melhor se aproxima a realidade, colocando assim o método no topo da classificação para esta série.

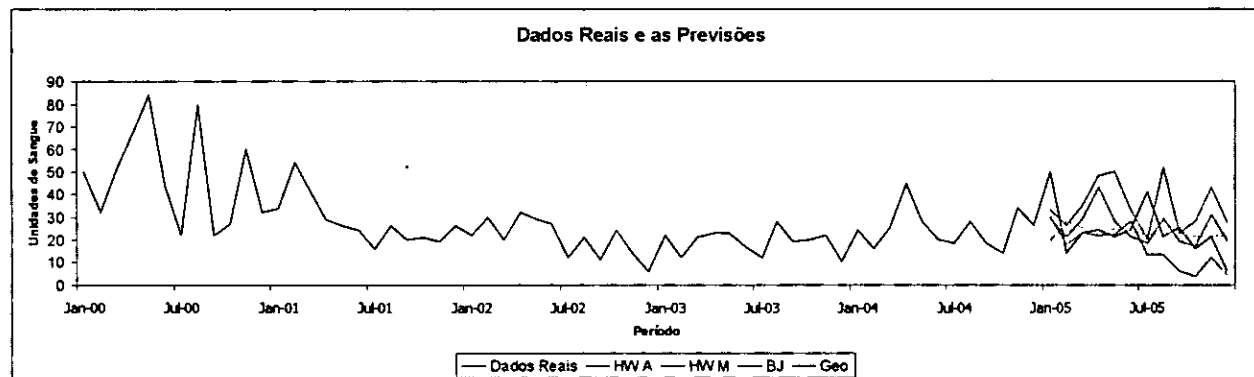


Figura 24. Previsões para as unidades de sangue das unidades de sangue infectados pelo Vírus HIV do BSHGJM

CAPÍTULO VI

Conclusões e Recomendações

1. Conclusões

O objectivo deste trabalho foi de comparar três métodos de previsão de séries temporais, nomeadamente Geoestatística, Holt-Winters e Box-Jenkins aplicando-os, mais precisamente, nas séries de unidades de sangue doadas nos Bancos de Sangue dos Hospitais Central de Maputo e Geral José Macamo, e as unidades de sangue infectadas por HIV nos mesmos Bancos de Sangue. Pelos resultados obtidos conclui-se, de um modo geral, que para este caso de estudo a Geoestatística apresentou resultados satisfatórios.

1. No que diz respeito a potencialidade ^{de} das previsões; ^{das} Das seis séries aplicadas, a geoestatística produziu previsões potencialmente boas em 33.33% ^{delas} das séries; comparando com os outros métodos que não chegaram a apresentar esta qualidade.

2. Em termos de desempenho dos métodos aplicados, ^{Mediante} o critério de decisão assumido (Estatística U de Theil), a geoestatística demonstrou ser o melhor método em relação aos outros (Holt-Winters e Box-Jenkins), gerando previsões com menor margem de erro em 5 (cinco) das 6 (seis) analisadas. Enquanto que o Box-Jenkins mostrou-se vencedor apenas na previsão dos dados da série unidades de sangue infectadas pelo HIV no BSHGJM. Este facto indica que a geoestatística poderia ser utilizada como um método de previsão automática para as séries de doações de sangue nestes hospitais aplicando um horizonte de previsão de 12 meses.

frase incompleta

Respondendo a pergunta de pesquisa, referir que as conclusões acima foram possíveis sob condição de testar os métodos seleccionados e comparar os resultados obtidos. E esta pode ser uma das melhores técnicas de busca de um método fiável para um caso específico.

2. Recomendações

Como já foi referido anteriormente, com este trabalho não se pretende eliminar a dificuldade de procura de um método fiável. Pelo contrário gasta-se muito tempo em testar vários modelos dos métodos escolhidos para a comparação e acaba-se esquecendo alguns detalhes de cada método aplicado.

O intuito principal deste trabalho foi demonstrar que pode-se aplicar vários métodos à mesma série e com ajuda dos critérios de selecção de métodos (neste trabalho, estatística U de Theil) poder escolher-se o melhor que se adequa aos dados em análise.

Usar métodos estatísticos para fazer previsões quantitativas é considerado hoje uma das bases de informação para a tomada de decisão em muitas áreas científicas.

Dentro de contexto, abaixo, são sugeridas ^{as} seguintes recomendações:

- Incluir mais métodos considerados eficientes na previsão de séries temporais como é o caso de modelos dinâmicos. E porquê não os assumidos elementares! Já que, produzir melhores resultados não é sinónimo de uso de métodos sofisticados.
- Explorar a potencialidade de outros estimadores aplicáveis na geoestatística de modo a ser possível detectar o comportamento sazonal das séries temporais.
- Utilização de métodos geoestatísticos apropriados à análise das séries temporais com tendência, como por exemplo, a técnica de *krigagem universal*.

CAPÍTULO VII

Bibliografia

um capítulo
chamado
Bibliográfica
que tal é se reti-
rasse o título
"Capítulo VII" e
deixar simplesmente
Bibliografia

ARMSTRONG, M., DOWD P. A. (1993). Geostatistical Simulations. 255 pp. London: Kluwer Academic Publishers.

BEZERRA, C. A., ARAUJO, M. S (2005). Component-Based Software Engineering: a Case Study Using Holt-Winters Models for Stochastic Decision Support Systems. 33 Jornadas Argentinas de Informática e Investigación Operativa, Universidad Nacional de Córdoba.

BRESEGHELLO, F. N. (2005). Estudo Comparativo de Métodos de Previsão de demanda: Uma Aplicação ao caso dos Aeroportos com Tráfego Aéreo Regular Administrados pelo DAESP. Tese de Mestrado. São José dos Campos, Instituto Tecnológico de Aeronáutica. Recuperado de <http://teses.eps.ufsc.br/defesa/pdf/10422.pdf> em 05 de Março de 2006.

CORTEZ, P. A. R. (1997). Algoritmos Genéricos e Redes Neurais na Previsão de Séries Temporais. (1997) Dissertação de Mestrado. Guimarães: Universidade do Minho.

CORTEZ, P. A. R. (2002). Modelo Inspirado na Natureza para a Previsão de Séries Temporais. Tese de Doutorado. Guimarães: Universidade do Minho.

DAVIS, M. M., AQUILANO, N, J., CHASE, R. B.(2001) Fundamentos da Administração da Produção. 3ª edição. Porto Alegre: Bookman Editora.

ENDERS, W. (1995). Applied Econometric Time Series. John Wiley and Sons.

EHLERS, R. S. (2005). Análise de Séries Temporais. Curitiba: Universidade Federal do Paraná.

GRANGER, C.W.J., NEWBOLD, P. (1986). Forecasting Economic Time Series. New York: Academic Press.

ISSAKS E.H.; SRIVASTAVA R.M.(1989). An Introduction to Applied Geostatistics. New York: Oxford University Press.

- LEWIS, C. D. (1997). Demand Forecasting and Inventory Control. New York: Wiley.
- LOPES, R. D. (2002). Previsão de Autopeças: Estudo de Caso em uma Concessionária de Veículos. Dissertação de Mestrado. Florianópolis: Universidade Federal de Santa Catarina.
- LUZ, A. C.C., SOUZA, J. N., VIANA, M. M., RIOS, J.R.,(2005). Gerência Pró-Ativa de Redes Usando Método de Previsão Baseado em Séries Temporais (Modelo de Box & Jenkins). XXII Simpósio Brasileiro de Telecomunicações SBrT'05, 04-08 de Setembro, Campinas, SP.
- MELO, B. S. V. (2001). Modelo de Previsão para Arrecadação Tributária – VI PRÊMIO Tesouro Nacional, Escola de Administração Fazendária. 97 pp. Recuperado de www.stn.fazenda.gov.br/PremioTN/VIPremio/sistemas/IsiafpVIPTN/MELOBruno.pdf em 14 de Fevereiro de 2006.
- MINGOTI, S. A., GLORIA, F. A. A. (2002). PreGeo – A Geoestatística Aplicada em Estudos de Séries Temporais: Uma Comparação com os Métodos de Holt-Winters e Box e Jenkins. Produtos & Produções, Universidade Federal de Minas Gerais. Vol. 6. n. 1, p. 24-32.
- MINGOTI, S. A., FIDELIS, M. T., (2001). Aplicando a Geoestatística no Controle Estatístico de Processos. Revista Escola de Minas, 5, n. 2, 55-70.
- MINGOTI, S. A., NEVES, O. F. (1999). A Metodologia de Geoestatística Aplicada na Análise de Séries Temporais. Revista Escola de Minas, 52, 3, 182-187.
- MINGOTI, S.A. (1996). As funções de Madograma e Rodograma como Alternativas para Descrever a Variabilidade Espacial dos Dados. Revista Escola de Minas (REM), Ouro Preto. vol 50, 2, p.71-74.
- MORETTIN., p. A.; TOLOI, C. M.C.; (1981) Modelos para Previsão de Séries Temporais. In: Colóquio Brasileiro de Matemática, 13º, 1981, Poços de Caldas-Mg, Rio de Janeiro: Instituto de Matemática Pura e Aplicada.
- MUELLER, A. (1996) Uma Aplicação de Redes Neurais Artificiais na Previsão do Mercado Acionário. Dissertação (programa de Pós-Graduação em Engenharia de Produção) - Departamento de Pós-Graduação em Engenharia de Produção. Florianópolis: Universidade Federal de Santa Catarina.

MURTEIRA, B. J. F., MÜLLER, D.A. e TURKMAN, K.F. (1993). Análise de Sucessões Cronológicas. McGraw Hill, Lisboa.

OLIVEIRA, F. H. (2002). Modelagem de Terreno Utilizando Sistemas Fotogramétricos. pp.183. Dissertação de Doutorado. Florianópolis: Universidade Federal de Santa Catarina.

PELLEGRINI, F. R., FOGLIATTO, F. S. (2000). Estudo Comparativo entre os Modelos de Holt-Winters e de Box-Jenkins para Previsão de Demanda Sazonal. Revista Produto & Produção, v.4, n.2, p.72-85.

REFENES, A. N. (1993) Constructive Learning and its application to currency exchange rate forecasting. In: Trippi, R. R., Turban, E. Neural networks in finance and investing. Chicago: Probus.

SAMOHYL, R.W. (2004). Medidas de Erro e Critérios para Avaliação da Previsão. Recuperado de www.webensino.ufsc.br. em 12/07/2006

SOUZA, G. P. (2005). Previsão do Consumo Industrial de Energia Elétrica no Estado de Santa Catarina: Uma Aplicação da Combinação de Previsões entre Modelos Univariados e de Regressão Dinâmica. 145p. Dissertação de Mestrado. Universidade Federal de Santa Catarina.

SOUZA, R. C. (1989). Modelos Estruturais para Previsão de Séries Temporais. Abordagens Clássica e Bayesiana. In : 17º Colóquio Brasileiro de Matemática. Rio de Janeiro.

SPIEGEL, M. R., (1994). Estatística 3ª Edição, São Paulo.

VASCONCELLOS, M. A. S. e ALVES, D. (2000) Manual de Econometria da Equipe de Professores da USP. São Paulo - SP: Editora Atlas.

WERNER, L., RIBEIRO, J. L. D. (2003). Previsão de Demanda: Uma Aplicação dos Modelos Box-Jenkins na Área de Assistência Técnica de Computadores Pessoais. Revista Produto & Produção v.10, n.1, p.47-67.

WOLD, H. (1998). Study in the analysis of stationary time series. 5ª. ed. Estocolmo: Almqvist & Wiksel.

Anexos

ANEXO A. VALORES REAIS

Tabela 1. Unidades de Sangue dos Dadores Voluntários do BSHCM

(em litros ou em unidades de 500 ml)?

ANO MÊS	2000	2001	2002	2003	2004	2005
Janeiro	573	876	693	863	667	1.114
Fevereiro	485	1.049	749	1.044	704	1.306
Março	770	596	1.171	1.045	1.183	1.557
Abril	813	746	795	946	1.124	1.457
Maio	833	694	1.915	1.191	1.017	1.614
Junho	761	546	553	1.037	1.115	1.469
Julho	669	847	759	1.108	1.068	920
Agosto	1.025	817	1.509	1.254	1.271	1.069
Setembro	553	1.134	790	1.195	1.057	737
Outubro	627	874	845	1.024	727	1.046
Novembro	807	582	767	755	783	1.090
Dezembro	643	831	1.046	1.106	1.013	998

Fonte: Banco de Sangue do Hospital Central de Maputo [Data: 19-07-2006]

Tabela 2. Unidades de Sangue dos Dadores Repositores do BSHCM

(litros ou Sacos de 500ml)?

ANO MÊS	2000	2001	2002	2003	2004	2005
Janeiro	475	598	634	272	313	361
Fevereiro	353	453	429	220	322	425
Março	556	579	417	240	379	438
Abril	360	475	375	234	575	452
Maio	511	575	348	227	493	440
Junho	422	480	97	187	408	407
Julho	311	325	242	250	314	318
Agosto	251	214	269	275	217	270
Setembro	194	210	130	194	214	201
Outubro	226	329	315	289	221	312
Novembro	306	456	295	367	254	249
Dezembro	647	541	240	369	288	218

Fonte: Banco de Sangue do Hospital Central de Maputo [Data: 19-07-2006]

Tabela 3. Unidades de Sangue dos Dadores do BSHCM

ANO MÊS	2000	2001	2002	2003	2004	2005
Janeiro	1.048	1.474	1.327	1.135	980	1.475
Fevereiro	838	1.502	1.178	1.264	1.026	1.731
Março	1.326	1.175	1.588	1.285	1.562	1.995
Abril	1.173	1.221	1.170	1.180	1.699	1.909
Maio	1.344	1.269	2.263	1.418	1.510	2.054
Junho	1.272	1.026	650	1.224	1.523	1.876
Julho	1.091	1.172	1.001	1.358	1.382	1.238
Agosto	1.336	1.031	1.778	1.529	1.488	1.339
Setembro	804	1.344	920	1.389	1.271	938
Outubro	821	1.203	1.160	1.313	948	1.358
Novembro	1.033	1.038	1.062	1.122	1.037	1.339
Dezembro	949	1.372	1.286	1.475	1.301	1.216

Fonte: Banco de Sangue do Hospital Central de Maputo [Data: 19-07-2006]

Tabela 4. Unidades de Sangue Infectadas por HIV do BSHCM

ANO MÊS	2000	2001	2002	2003	2004	2005
Janeiro	100	137	103	73	57	63
Fevereiro	75	107	84	53	69	97
Março	115	88	77	65	79	13
Abril	77	69	90	42	110	77
Maio	95	78	107	48	95	66
Junho	113	75	44	15	51	61
Julho	64	59	26	0	53	30
Agosto	73	50	95	50	59	14
Setembro	42	61	57	46	36	18
Outubro	52	28	76	74	40	22
Novembro	52	82	60	79	51	10
Dezembro	107	121	71	114	89	33

Fonte: Banco de Sangue do Hospital Central de Maputo [Data: 19-07-2006]

Tabela 5. Unidades de Sangue dos Dadores Voluntários do BSHGJM

ANO MÊS	2000	2001	2002	2003	2004	2005
Janeiro	47	46	59	58	133	164
Fevereiro	144	199	119	48	40	84
Março	41	48	125	197	97	277
Abril	184	48	235	124	137	61
Maio	48	73	224	130	120	117
Junho	160	81	90	66	61	245
Julho	34	32	52	62	56	143
Agosto	322	204	159	175	378	233
Setembro	28	42	52	240	227	103
Outubro	50	60	45	48	145	131
Novembro	191	46	37	247	144	128
Dezembro	52	77	71	68	183	165

Fonte: Banco de Sangue do Hospital Geral José Macamo [Data: 07-07-2006]

Tabela 6. Unidades de Sangue dos Dadores Repositores do BSHGJM

ANO MÊS	2000	2001	2002	2003	2004	2005
Janeiro	247	225	135	144	87	161
Fevereiro	216	195	84	123	120	74
Março	214	196	122	107	125	126
Abril	228	117	99	139	171	167
Maio	270	97	102	80	199	141
Junho	157	112	107	114	159	109
Julho	182	105	119	81	134	157
Agosto	169	111	115	100	102	118
Setembro	107	102	90	72	71	82
Outubro	210	120	128	114	75	120
Novembro	275	127	115	87	130	109
Dezembro	202	131	91	63	136	100

Fonte: Banco de Sangue do Hospital Geral José Macamo [Data: 07-07-2006]

Tabela 7. Unidades de Sangue dos Dadores do BSHGJM

ANO MÊS	2000	2001	2002	2003	2004	2005
Janeiro	294	271	194	202	220	325
Fevereiro	360	394	203	171	160	158
Março	255	244	247	304	222	403
Abril	412	165	334	263	308	228
Mai	318	170	326	210	319	258
Junho	317	193	197	180	220	354
Julho	216	137	171	143	190	300
Agosto	491	315	274	275	480	351
Setembro	135	144	142	312	298	185
Outubro	260	180	173	162	220	251
Novembro	466	173	152	334	274	237
Dezembro	254	208	162	131	319	265

Fonte: Banco de Sangue do Hospital Geral José Macamo [Data: 07-07-2006]

Tabela 8. Unidades de Sangue Infectadas por HIV do BSHJGM

ANO MÊS	2000	2001	2002	2003	2004	2005
Janeiro	50	34	22	22	24	50
Fevereiro	32	54	30	12	16	14
Março	51	42	20	21	25	23
Abril	67	29	32	23	45	24
Mai	84	26	29	23	28	21
Junho	44	24	27	17	20	24
Julho	22	16	12	12	18	41
Agosto	80	26	21	28	28	21
Setembro	22	20	11	19	18	25
Outubro	27	21	24	20	14	16
Novembro	60	19	14	22	34	21
Dezembro	32	26	6	10	26	6

Fonte: Banco de Sangue do Hospital Geral José Macamo [Data: 07-07-2006]

ANEXO B: RESULTADOS DA ANÁLISE DA CONSISTÊNCIA INTERNA DOS DADOS

Tabela 1. Séries do BSHGJM

RELIABILITY ANALYSIS - SCALE (ALPHA)

Correlation Matrix						
	DOACHGJM	HIVHGJM				
DOACHGJM	1,0000					
HIVHGJM	,6298	1,0000				
N of Cases =		72,0				
Item Means	Mean	Minimum	Maximum	Range	Max/Min	Variance
	139,9028	27,3194	252,4861	225,1667	9,2420	25350,0139
Item Variances	Mean	Minimum	Maximum	Range	Max/Min	Variance
	3797,5045	230,0514	7364,9576	7134,9061	32,0144	25453442,6
Reliability Coefficients		2 items				
Alpha = ,3551		Standardized item alpha = ,7728				

Tabela 2. Séries do BSHCM

RELIABILITY ANALYSIS - SCALE (ALPHA)

Correlation Matrix						
	DOACHCM	HIVHCM				
DOACHCM	1,0000					
HIVHCM	,2124	1,0000				
N of Cases =		72,0				
Item Means	Mean	Minimum	Maximum	Range	Max/Min	Variance
	679,8194	66,1389	1293,5000	1227,3611	19,5573	753207,649
Item Variances	Mean	Minimum	Maximum	Range	Max/Min	Variance
	45500,9480	883,2480	90118,6479	89235,3998	102,0310	3981478293
Reliability Coefficients		2 items				
Alpha = ,0800		Standardized item alpha = ,3504				

ANEXO C. VALORES PREVISTOS

Tabela 1. Unidades de Sangue dos Dadores Voluntários do BSHCM

MÊS \ MÉTODO	HW A	HW M	BJ	GEO
Jan-05	942	910	965	910
Fev-05	965	958	928	947
Mar-05	1189	1233	970	1078
Abr-05	1154	1169	927	1109
Mai-05	1315	1405	957	1102
Jun-05	1100	1090	929	997
Jul-05	1120	1134	946	966
Ago-05	1429	1530	929	957
Set-05	1118	1145	938	917
Out-05	1049	1024	927	929
Nov-05	1063	1011	932	962
Dez-05	1147	1160	924	1011

Tabela 2. Unidades de Sangue dos Dadores Repositores do BSHCM

MÊS \ MÉTODO	HW A	HW M	BJ	GEO
Jan-05	300	292	236	346
Fev-05	319	280	235	400
Mar-05	370	323	284	406
Abr-05	479	365	456	439
Mai-05	376	309	377	407
Jun-05	281	214	292	366
Jul-05	238	204	209	329
Ago-05	207	172	120	285
Set-05	199	143	105	265
Out-05	238	175	116	261
Nov-05	296	209	149	285
Dez-05	333	230	176	306

Tabela 3. Unidades de Sangue dos Dadores do BSHCM

MÊS \ MÉTODO	HW A	HW M	BJ	GEO
Jan-05	1.242	1.202	1.201	1.256
Fev-05	1.284	1.238	1.163	1.347
Mar-05	1.559	1.556	1.254	1.484
Abr-05	1.633	1.534	1.383	1.548
Mai-05	1.691	1.714	1.334	1.509
Jun-05	1.381	1.304	1.221	1.363
Jul-05	1.358	1.338	1.155	1.295
Ago-05	1.636	1.702	1.049	1.242
Set-05	1.317	1.288	1.043	1.182
Out-05	1.287	1.199	1.043	1.190
Nov-05	1.359	1.220	1.081	1.247
Dez-05	1.480	1.390	1.100	1.317

Tabela 4. Unidades de Sangue Infectadas por HIV do BSHCM

MÉTODOS MÊS	HW A	HW M	BJ	GEO
Jan-05	94	93	60	72
Fev-05	72	73	53	80
Mar-05	101	99	64	24
Abr-05	74	75	66	58
Mai-05	88	87	62	61
Jun-05	90	88	24	30
Jul-05	52	52	17	47
Ago-05	67	67	46	23
Set-05	41	42	33	-6
Out-05	49	51	50	24
Nov-05	53	54	58	21
Dez-05	101	100	94	52

Tabela 5. Unidades de Sangue dos Dadores Voluntários do BSHGJM

MÉTODOS MÊS	HW A	HW M	BJ	GEO
Jan-05	125	125	122	124
Fev-05	34	40	85	86
Mar-05	96	102	108	149
Abr-05	133	136	124	95
Mai-05	118	121	117	137
Jun-05	59	62	93	106
Jul-05	54	57	91	152
Ago-05	367	370	219	239
Set-05	222	233	159	157
Out-05	135	141	127	112
Nov-05	141	155	126	135
Dez-05	172	179	142	212

Tabela 6. Unidades de Sangue dos Dadores Repositores do BSHGJM

MÉTODOS MÊS	HW A	HW M	BJ	GEO
Jan-05	108	124	149	121
Fev-05	132	109	159	125
Mar-05	126	106	144	127
Abr-05	158	106	130	125
Mai-05	158	104	134	125
Jun-05	109	76	128	126
Jul-05	75	70	131	126
Ago-05	46	63	140	126
Set-05	17	41	142	126
Out-05	29	61	142	126
Nov-05	74	70	145	126
Dez-05	66	53	142	126

Tabela 7. Unidades de Sangue dos Dadores do BSHGJM

MÊS \ MÉTODO	HW A	HW M	BJ	GEO
Jan-05	233	249	271	245
Fev-05	166	149	244	211
Mar-05	222	208	252	277
Abr-05	291	242	254	221
Mai-05	276	225	251	263
Jun-05	168	138	221	232
Jul-05	129	127	222	278
Ago-05	413	433	359	364
Set-05	239	274	301	283
Out-05	164	202	269	238
Nov-05	215	225	271	261
Dez-05	238	232	284	337

Tabela 8. Unidades de Sangue Infectadas por HIV do BSHJGM

MÊS \ MÉTODO	HW A	HW M	BJ	GEO
Jan-05	29	33	19	30
Fev-05	21	26	28	18
Mar-05	29	35	25	23
Abr-05	43	48	21	22
Mai-05	28	50	25	22
Jun-05	21	33	24	28
Jul-05	18	20	22	13
Ago-05	29	52	23	13
Set-05	19	23	22	6
Out-05	17	28	21	4
Nov-05	31	43	22	12
Dez-05	19	27	21	4

Anexo D

Dados do BSHCM

Doações por Tipo de Dador

Período : Desde 01-Jan-00 até 31-Dec-00

Ano / Mes	Nº de Doações				Total
	Voluntários		Repositores		
	Nº	%	Nº	%	
2000					
Janeiro	573	54	475	45	1059
Fevereiro	485	57	353	42	847
Marco	770	58	556	42	1336
Abril	813	68	360	30	1190
Maio	833	61	511	37	1365
Junho	761	64	422	35	1195
Julho	669	67	311	31	994
Agosto	1025	79	251	19	1290
Setembro	553	73	194	25	762
Outubro	627	72	226	26	873
Novembro	807	71	306	27	1137
Dezembro	643	49	647	50	1303
Totais do Período	8559	64.11	4612	34.54	13351

Doenças Infecciosas por Faixa Etária

Período : Desde 01-Jan-00 até 31-Dec-00

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
2000													
Janeiro													
Doações	1059	4	0.38	223	21.1	312	29	312	29.5	141	13	63	5.95
Infectadas	142	0	0	19	13.4	67	47	39	27.5	20	14	6	4.23
HIV	100			12	12.	44	44.	27	27.	13	13.	4	4.
VDRL	51			7	13.7	23	45.1	12	23.5	7	13.7	2	3.9
HBS	0												
SGTP	0												
Fevereiro													
Doações	847	1	0.12	193	22.8	284	34	202	23.8	115	14	52	6.14
Infectadas	111	1	0.90	16	14.4	60	54	26	23.4	11	9.9	11	9.91
HIV	75	1	1.3	10	13.3	43	57.3	11	14.7	4	5.3	6	8.
VDRL	50			6	12.	17	34.	15	30.	7	14.	5	10.
HBS	0												
SGTP	0												
Marco													
Doações	1336	4	0.3	441	33.0	368	28	313	23.4	139	10	71	5.31
Infectadas	141	1	0.71	35	24.8	62	44	39	27.7	10	7.1	2	1.42
HIV	115			27	23.5	54	47.	26	22.6	6	5.2	2	1.7
VDRL	34	1	2.9	8	23.5	8	23.5	13	38.2	4	11.8		
HBS	0												
SGTP	0												
Abril													
Doações	1190	4	0.34	390	32.8	318	27	287	24.1	136	11	54	4.54
Infectadas	124	0	0	24	19.4	52	42	40	32.3	14	11	1	0.81
HIV	77			10	13.	35	45.5	27	35.1	5	6.5		
VDRL	54			14	25.9	17	31.5	13	24.1	9	16.7	1	1.9
HBS	0												
SGTP	0												

60-65

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Maio													
Doações	1365	8	0.59	570	41.8	326	24	269	19.7	118	8.6	74	5.42
Infectadas	160	0	0	50	31.3	60	38	37	23.1	17	11	7	4.38
HIV	95			21	22.1	39	41.1	21	22.1	11	11.6	3	3.2
VDRL	74			29	39.2	19	25.7	16	21.6	6	8.1	4	5.4
HBS	1					1	100.						
SGTP	0												
Junho													
Doações	1195	3	0.25	356	29.8	372	31	281	23.5	127	11	55	4.60
Infectadas	175	1	0.57	41	23.4	75	43	49	28	9	5.1	11	6.29
HIV	113			26	23.	49	43.4	30	26.5	5	4.4	3	2.7
VDRL	73	1	1.4	15	20.5	26	35.6	19	26.	4	5.5	8	11.
HBS	0												
SGTP	0												
Julho													
Doações	994	1	0.10	389	39.1	289	29	192	19.3	75	7.5	47	4.73
Infectadas	79	1	1.27	22	27.8	37	47	17	21.5	5	6.3	2	2.53
HIV	64	1	1.6	14	21.9	30	46.9	14	21.9	4	6.3	1	1.6
VDRL	20			8	40.	7	35.	3	15.	1	5.	1	5.
HBS	0												
SGTP	0												
Agosto													
Doações	1290	5	0.39	591	45.8	281	22	257	19.9	109	8.4	47	3.64
Infectadas	112	0	0	43	38.4	37	33	27	24.1	10	8.9	1	0.89
HIV	73			24	32.9	27	37.	16	21.9	6	8.2		
VDRL	45			19	42.2	10	22.2	11	24.4	4	8.9	1	2.2
HBS	0												
SGTP	0												
Setembro													

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Doações	762	5	0.66	268	35.2	193	25	170	22.3	79	10	47	6.17
Infectadas	93	0	0	24	25.8	33	35	31	33.3	8	8.6	4	4.30
HIV	42			6	14.3	20	47.6	12	28.6	3	7.1	1	2.4
VDRL	58			18	31.	13	22.4	19	32.8	5	8.6	3	5.2
HBS	0												
SGTP	0												
Outubro													
Doações	873	1	0.11	307	35.2	210	24	195	22.3	95	11	64	7.33
Infectadas	83	0	0	18	21.7	30	36	22	26.5	12	14	5	6.02
HIV	52			11	21.2	19	36.5	13	25.	8	15.4	1	1.9
VDRL	35			7	20.	11	31.4	9	25.7	4	11.4	4	11.4
HBS	0												
SGTP	0												
Novembro													
Doações	1137	3	0.26	405	35.6	324	28	247	21.7	98	8.6	60	5.28
Infectadas	72	0	0	16	22.2	36	50	19	26.4	2	2.8	3	4.17
HIV	52			11	21.2	21	40.4	16	30.8	1	1.9	3	5.8
VDRL	24			5	20.8	15	62.5	3	12.5	1	4.2		
HBS	0												
SGTP	0												
Dezembro													
Doações	1303	8	0.61	337	25.9	392	30	344	26.4	156	12	66	5.07
Infectadas	136	0	0	29	21.3	55	40	40	29.4	10	7.4	4	2.94
HIV	107			20	18.7	46	43.	33	30.8	6	5.6	2	1.9
VDRL	31			9	29.	9	29.	7	22.6	4	12.9	2	6.5
HBS	0												
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%

Total do período

Doações	13351	47	0.35	4470	33.5	27	3069	23	10	700	5.24		
Infectadas	1428	4	0.28	337	23.6	604	42	386	27.0	128	9	57	3.99
HIV	965	2	.2	192	19.9	427	44.2	246	25.5	72	7.5	26	2.7
VDRL	549	2	.4	145	26.4	175	31.9	140	25.5	56	10.2	31	5.6
HBS	1					1	100.						
SGTP	0												

Doações por Tipo de Dador

Período : Desde 01-Jan-01 até 31-Dec-01

Ano / Mes	Nº de Doações				Total
	Voluntários		Repositores		
	Nº	%	Nº	%	
2001					
Janeiro	876	59	598	40	1489
Fevereiro	1049	69	453	30	1512
Marco	596	50	579	48	1196
Abril	746	60	475	38	1242
Maió	694	54	575	45	1287
Junho	546	52	480	46	1040
Julho	847	71	325	27	1195
Agosto	817	78	214	20	1050
Setembro	1134	84	210	15	1355
Outubro	874	71	329	27	1229
Novembro	582	55	456	43	1057
Dezembro	831	60	541	39	1386
Totais do Período	9592	63.79	5235	34.81	15038

Doenças Infecciosas por Faixa Etária

Período : Desde 01-Jan-01 até 31-Dec-01

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
2001													
Janeiro													
Doações	1489	2	0.13	409	27.5	466	31	365	24.5	173	12	73	4.90
Infectadas	197	0	0	43	21.8	81	41	56	28.4	22	11	7	3.55
HIV	137			26	19.	58	42.3	37	27.	13	9.5	3	2.2
VDRL	72			17	23.6	23	31.9	19	26.4	9	12.5	4	5.6
HBS	0												
SGTP	0												
Fevereiro													
Doações	1512	3	0.2	628	41.5	420	28	290	19.2	123	8.1	48	3.17
Infectadas	134	1	0.75	32	23.9	61	46	31	23.1	8	6	4	2.99
HIV	107	1	.9	24	22.4	51	47.7	22	20.6	6	5.6	3	2.8
VDRL	30			8	26.7	10	33.3	9	30.	2	6.7	1	3.3
HBS	0												
SGTP	0												
Marco													
Doações	1196	4	0.33	385	32.2	299	25	303	25.3	140	12	65	5.43
Infectadas	113	2	1.77	28	24.8	39	35	36	31.9	12	11	3	2.65
HIV	88	2	2.3	18	20.5	30	34.1	28	31.8	9	10.2	1	1.1
VDRL	32			10	31.3	9	28.1	8	25.	3	9.4	2	6.3
HBS	0												
SGTP	0												
Abril													
Doações	1242	4	0.32	507	40.8	334	27	248	20	86	6.9	61	4.91
Infectadas	118	1	0.85	26	22.0	56	47	27	22.9	8	6.8	3	2.54
HIV	69			11	15.9	37	53.6	17	24.6	4	5.8		
VDRL	52	1	1.9	15	28.8	19	36.5	10	19.2	4	7.7	3	5.8
HBS	0												
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%

Maio

Doações	1287	5	0.39	486	37.8	349	27	269	20.9	124	9.6	54	4.2
Infectadas	89	0	0	24	27	35	39	21	23.6	7	7.9	2	2.25
HIV	78			22	28.2	29	37.2	21	26.9	4	5.1	2	2.6
VDRL	11			2	18.2	6	54.5			3	27.3		
HBS	0												
SGTP	0												

Junho

Doações	1040	2	0.19	310	29.8	313	30	253	24.3	111	11	51	4.90
Infectadas	93	0	0	18	19.4	35	38	31	33.3	6	6.5	5	5.38
HIV	75			15	20.	27	36.	26	34.7	3	4.	4	5.3
VDRL	20			3	15.	8	40.	5	25.	3	15.	1	5.
HBS	0												
SGTP	0												

Julho

Doações	1195	3	0.25	650	54.4	234	20	181	15.1	79	6.6	48	4.02
Infectadas	70	0	0	23	32.9	25	36	18	25.7	5	7.1	2	2.86
HIV	59			14	23.7	23	39.	17	28.8	4	6.8	1	1.7
VDRL	14			9	64.3	2	14.3	1	7.1	1	7.1	1	7.1
HBS	0												
SGTP	0												

Agosto

Doações	1050	7	0.67	548	52.2	197	19	172	16.4	80	7.6	46	4.38
Infectadas	54	0	0	21	38.9	23	43	9	16.7	3	5.6	0	0
HIV	50			18	36.	20	40.	9	18.	3	6.		
VDRL	6			3	50.	3	50.						
HBS	0												
SGTP	0												

Setembro

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Doações	1355	5	0.37	728	53.7	268	20	225	16.6	81	6	48	3.54
Infectadas	74	0	0	26	35.1	27	36	16	21.6	3	4.1	2	2.70
HIV	61			20	32.8	24	39.3	14	23.	2	3.3	1	1.6
VDRL	13			6	46.2	3	23.1	2	15.4	1	7.7	1	7.7
HBS	0												
SGTP	0												
Outubro													
Doações	1229	1	0.08	621	50.5	219	18	224	18.2	102	8.3	61	4.96
Infectadas	44	0	0	12	27.3	16	36	12	27.3	3	6.8	1	2.27
HIV	28			8	28.6	11	39.3	6	21.4	3	10.7		
VDRL	16			4	25.	5	31.3	6	37.5			1	6.3
HBS	0												
SGTP	0												
Novembro													
Doações	1057	1	0.09	417	39.5	286	27	216	20.4	86	8.1	51	4.82
Infectadas	113	0	0	37	32.7	39	35	31	27.4	8	7.1	2	1.77
HIV	82			29	35.4	25	30.5	23	28.	5	6.1		
VDRL	35			8	22.9	14	40.	8	22.9	3	8.6	2	5.7
HBS	0												
SGTP	0												
Dezembro													
Doações	1386	7	0.51	537	38.7	352	25	310	22.4	125	9.0	55	3.97
Infectadas	180	3	1.67	62	34.4	63	35	44	24.4	8	4.4	10	5.56
HIV	121	2	1.7	38	31.4	45	37.2	28	23.1	3	2.5	5	4.1
VDRL	69	1	1.4	24	34.8	18	26.1	16	23.2	5	7.2	5	7.2
HBS	0												
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%

Total do período

Doações	15038	44	0.29	6226	41.4	503	25	3056	20.3	503	8.7	661	4.4
Infectadas	1279	7	0.55	352	27.5	500	39	332	26	93	7.3	41	3.21
HIV	955	5	.5	243	25.4	380	39.8	248	26	59	6.2	20	2.1
VDRL	370	2	.5	109	29.5	120	32.4	84	22.7	34	9.2	21	5.7
HBS	0												
SGTP	0												

Doações por Tipo de Dador

Período : Desde 01-Jan-02 até 31-Dec-02

Ano / Mes	Nº de Doações				Total
	Voluntários		Repositores		
	Nº	%	Nº	%	
2002					
Janeiro	693	52	634	47	1339
Fevereiro	749	63	429	36	1187
Marco	1171	73	417	26	1600
Abril	795	66	375	31	1197
Maio	1915	84	348	15	2281
Junho	553	84	97	15	662
Julho	759	76	242	24	1005
Agosto	1509	84	269	15	1788
Setembro	790	84	130	14	940
Outubro	845	72	315	27	1180
Novembro	767	71	295	27	1074
Dezembro	1046	81	240	19	1296
Totais do Período	11592	74.55	3791	24.38	15549

Doenças Infecciosas por Faixa Etária

Período : Desde 01-Jan-02 até 31-Dec-02

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
2002													
Janeiro													
Doações	1339	2	0.15	423	31.6	405	30	324	24.2	127	9.5	58	4.33
Infectadas	114	1	0.88	28	24.6	49	43	28	24.6	9	7.9	2	1.75
HIV	103	1	1.	23	22.3	44	42.7	27	26.2	7	6.8	1	1.
VDRL	14			5	35.7	5	35.7	1	7.1	2	14.3	1	7.1
HBS	0												
SGTP	0												
Fevereiro													
Doações	1187	4	0.34	448	37.7	296	25	271	22.8	109	9.2	58	4.89
Infectadas	96	0	0	30	31.3	24	25	28	29.2	12	13	3	3.13
HIV	84			24	28.6	22	26.2	26	31.	10	11.9	2	2.4
VDRL	13			6	46.2	2	15.4	2	15.4	2	15.4	1	7.7
HBS	0												
SGTP	0												
Marco													
Doações	1600	8	0.5	945	59.1	298	19	221	13.8	93	5.8	35	2.19
Infectadas	83	0	0	24	28.9	29	35	22	26.5	8	9.6	1	1.20
HIV	77			22	28.6	27	35.1	20	26.	7	9.1	1	1.3
VDRL	7			2	28.6	2	28.6	2	28.6	1	14.3		
HBS	0												
SGTP	0												
Abril													
Doações	1197	5	0.42	573	47.9	257	21	233	19.5	86	7.2	42	3.51
Infectadas	97	0	0	33	34.0	35	36	22	22.7	8	8.2	0	0
HIV	90			29	32.2	34	37.8	19	21.1	8	8.9		
VDRL	8			4	50.	1	12.5	3	37.5				
HBS	0												
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		N°	%	N°	%	N°	%	N°	%	N°	%	N°	%
Maio													
Doações	2281	7	0.31	1117	49	508	22	373	16.4	195	8.5	80	3.51
Infectadas	115	0	0	39	33.9	35	30	29	25.2	11	9.6	4	3.48
HIV	107			31	29.	35	32.7	27	25.2	10	9.3	4	3.7
VDRL	11			8	72.7			2	18.2	1	9.1		
HBS	0												
SGTP	0												
Junho													
Doações	662	1	0.15	324	48.9	112	17	117	17.7	76	11	32	4.83
Infectadas	46	0	0	14	30.4	18	39	9	19.6	4	8.7	1	2.17
HIV	44			14	31.8	17	38.6	9	20.5	3	6.8	1	2.3
VDRL	2					1	50.			1	50.		
HBS	0												
SGTP	0												
Julho													
Doações	1005	10	1	633	63	153	15	112	11.1	62	6.2	35	3.48
Infectadas	28	0	0	10	35.7	11	39	6	21.4	0	0	1	3.57
HIV	26			9	34.6	10	38.5	6	23.1			1	3.8
VDRL	2			1	50.	1	50.						
HBS	0												
SGTP	0												
Agosto													
Doações	1788	12	0.67	1047	58.6	330	18	245	13.7	92	5.1	61	3.41
Infectadas	119	1	0.84	53	44.5	45	38	17	14.3	10	8.4	1	0.84
HIV	95	1	1.1	36	37.9	37	38.9	14	14.7	6	6.3	1	1.1
VDRL	32			17	53.1	8	25.	3	9.4	4	12.5		
HBS	0												
SGTP	0												
Setembro													

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Doações	940	8	0.85	587	62.4	147	16	116	12.3	56	6	26	2.77
Infectadas	70	0	0	41	58.6	16	23	10	14.3	5	7.1	0	0
HIV	57			33	57.9	12	21.1	8	14.	4	7.		
VDRL	15			8	53.3	4	26.7	2	13.3	1	6.7		
HBS	0												
SGTP	0												
Outubro													
Doações	1180	12	1.02	695	58.9	196	17	156	13.2	78	6.6	42	3.56
Infectadas	93	0	0	35	37.6	34	37	22	23.7	5	5.4	2	2.15
HIV	76			26	34.2	28	36.8	17	22.4	4	5.3	1	1.3
VDRL	22			9	40.9	6	27.3	5	22.7	1	4.5	1	4.5
HBS	0												
SGTP	0												
Novembro													
Doações	1074	5	0.47	486	45.3	261	24	204	19	79	7.4	38	3.54
Infectadas	84	0	0	28	33.3	35	42	20	23.8	4	4.8	3	3.57
HIV	60			18	30.	28	46.7	10	16.7	3	5.	1	1.7
VDRL	30			10	33.3	7	23.3	10	33.3	1	3.3	2	6.7
HBS	0												
SGTP	0												
Dezembro													
Doações	1296	11	0.85	676	52.2	255	20	226	17.4	90	6.9	37	2.85
Infectadas	97	0	0	43	44.3	30	31	16	16.5	10	10	0	0
HIV	71			27	38.	25	35.2	14	19.7	5	7.		
VDRL	28			16	57.1	5	17.9	2	7.1	5	17.9		
HBS	0												
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%

Total do período

Doações	15549	85	0.55	7954	51.2E+03	21	2598	16.7 E+03	7.4	544	3.5		
Infectadas	1042	2	0.19	378	36.3	361	35	229	22	86	8.3	18	1.73
HIV	890	2	.2	292	32.8	319	35.8	197	22.1	67	7.5	13	1.5
VDRL	184			86	46.7	42	22.8	32	17.4	19	10.3	5	2.7
HBS	0												
SGTP	0												

Doações por Tipo de Dador

Período : Desde 01-Jan-03 até 31-Dec-03

Ano / Mes	Nº de Doações				Total
	Voluntários		Repositores		
	Nº	%	Nº	%	
2003					
Janeiro	863	75	272	24	1149
Fevereiro	1044	82	220	17	1279
Marco	1045	80	240	18	1304
Abril	946	80	234	20	1187
Maió	1191	83	227	16	1431
Junho	1037	84	187	15	1238
Julho	1108	81	250	18	1366
Agosto	1254	81	275	18	1542
Setembro	1195	85	194	14	1408
Outubro	1024	77	289	22	1336
Novembro	755	66	367	32	1141
Dezembro	1106	74	369	25	1489
Totais do Período	12568	79.19	3124	19.68	15870

Doenças Infecciosas por Faixa Etária

Período : Desde 01-Jan-03 até 31-Dec-03

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
2003													
Janeiro													
Doações	1149	15	1.31	487	42.4	303	26	199	17.3	101	8.8	43	3.74
Infectadas	101	1	0.99	29	28.7	50	50	17	16.8	5	5	1	0.99
HIV	73	1	1.4	17	23.3	37	50.7	13	17.8	4	5.5	1	1.4
VDRL	30			12	40.	13	43.3	4	13.3	1	3.3		
HBS	0												
SGTP	0												
Fevereiro													
Doações	1279	85	6.65	800	62.5	156	12	135	10.6	62	4.8	41	3.21
Infectadas	69	0	0	25	36.2	25	36	15	21.7	8	12	2	2.9
HIV	53			15	28.3	19	35.8	12	22.6	6	11.3	1	1.9
VDRL	22			10	45.5	6	27.3	3	13.6	2	9.1	1	4.5
HBS	0												
SGTP	0												
Marco													
Doações	1304	118	9.05	775	59.4	179	14	144	11.0	58	4.4	30	2.30
Infectadas	79	4	5.06	29	36.7	29	37	13	16.5	2	2.5	2	2.53
HIV	65	1	1.5	24	36.9	25	38.5	12	18.5	2	3.1	1	1.5
VDRL	14	3	21.4	5	35.7	4	28.6	1	7.1			1	7.1
HBS	0												
SGTP	0												
Abril													
Doações	1187	62	5.22	663	55.9	213	18	171	14.4	49	4.1	29	2.44
Infectadas	55	0	0	26	47.3	15	27	16	29.1	1	1.8	1	1.82
HIV	42			18	42.9	10	23.8	13	31.			1	2.4
VDRL	17			8	47.1	5	29.4	3	17.6	1	5.9		
HBS	0												
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Maio													
Doações	1431	119	8.32	890	62.2	182	13	147	10.3	64	4.5	29	2.03
Infectadas	64	4	6.25	26	40.6	23	36	12	18.8	1	1.6	0	0
HIV	48	3	6.3	15	31.3	20	41.7	9	18.8	1	2.1		
VDRL	18	1	5.6	11	61.1	3	16.7	3	16.7				
HBS	0												
SGTP	0												
Junho													
Doações	1238	116	9.37	763	61.6	156	13	113	9.13	61	4.9	28	2.26
Infectadas	21	1	4.76	6	28.6	9	43	4	19.0	1	4.8	1	4.76
HIV	15			6	40.	3	20.	4	26.7	1	6.7	1	6.7
VDRL	5	1	20.			4	80.						
HBS	1					1	100.						
SGTP	0												
Julho													
Doações	1366	97	7.10	825	60.4	175	13	164	12.0	73	5.3	32	2.34
Infectadas	22	1	4.55	10	45.5	7	32	2	9.09	1	4.5	1	4.55
HIV	0												
VDRL	22	1	4.5	10	45.5	7	31.8	2	9.1	1	4.5	1	4.5
HBS	0												
SGTP	0												
Agosto													
Doações	1542	92	5.97	905	58.7	277	18	170	11.0	66	4.3	32	2.08
Infectadas	61	4	6.56	27	44.3	23	38	7	11.5	2	3.3	0	0
HIV	50	4	8.	20	40.	18	36.	6	12.	2	4.		
VDRL	13			7	53.8	5	38.5	1	7.7				
HBS	0												
SGTP	0												
Setembro													

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Doações	1408	173	12.3	821	58.3	203	14	123	8.74	57	4.0	30	2.13
Infectadas	51	2	3.92	18	35.3	22	43	8	15.7	3	5.9	0	0
HIV	46	1	2.2	16	34.8	18	39.1	8	17.4	3	6.5		
VDRL	7	1	14.3	2	28.6	4	57.1						
HBS	0												
SGTP	0												
Outubro													
Doações	1336	89	6.66	705	52.8	256	19	179	13.4	61	4.6	46	3.44
Infectadas	86	2	2.33	39	45.3	29	34	14	16.3	5	5.8	2	2.33
HIV	74	1	1.4	30	40.5	24	32.4	13	17.6	4	5.4	2	2.7
VDRL	17	1	5.9	9	52.9	5	29.4	1	5.9	1	5.9		
HBS	0												
SGTP	0												
Novembro													
Doações	1141	25	2.19	532	46.6	253	22	196	17.2	95	8.3	40	3.51
Infectadas	93	0	0	39	41.9	36	39	15	16.1	7	7.5	1	1.08
HIV	79			32	40.5	29	36.7	13	16.5	4	5.1	1	1.3
VDRL	19			7	36.8	7	36.8	2	10.5	3	15.8		
HBS	0												
SGTP	0												
Dezembro													
Doações	1489	52	3.49	765	51.4	326	22	200	13.4	112	7.5	34	2.28
Infectadas	141	2	1.42	55	39.0	53	38	30	21.3	8	5.7	2	1.42
HIV	114	2	1.8	40	35.1	42	36.8	25	21.9	4	3.5	1	.9
VDRL	36			15	41.7	11	30.6	5	13.9	4	11.1	1	2.8
HBS	0												
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%

Total do período

Doações	15870	1043	6.57	8931	56.3E+03	17	1941	12.2	859	5.4	414	2.61	
Infectadas	843	21	2.49	329	39.0	321	38	153	18.1	44	5.2	13	1.54
HIV	659	13	2.	233	35.4	245	37.2	128	19.4	31	4.7	9	1.4
VDRL	220	8	3.6	96	43.6	74	33.6	25	11.4	13	5.9	4	1.8
HBS	1						1	100.					
SGTP	0												

Doações por Tipo de Dador

Período : Desde 01-Jan-04 até 31-Dec-04

Ano / Mes	Nº de Doações				Total
	Voluntários		Repositores		
	Nº	%	Nº	%	
2004					
Janeiro	667	67	313	32	991
Fevereiro	704	68	322	31	1037
Marco	1183	75	379	24	1577
Abril	1124	66	575	34	1714
Maio	1017	67	493	32	1520
Junho	1115	73	408	27	1537
Julho	1068	77	314	23	1389
Agosto	1271	85	217	14	1501
Setembro	1057	83	214	17	1275
Outubro	727	75	221	23	966
Novembro	783	75	254	24	1043
Dezembro	1013	77	288	22	1312
Totais do Período	11729	73.94	3998	25.20	15862



Doenças Infecciosas por Faixa Etária

Período : Desde 01-Jan-04 até 31-Dec-04

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
2004													
Janeiro													
Doações	991	35	3.53	436	44	219	22	200	20.2	73	7.4	28	2.83
Infectadas	74	2	2.70	23	31.1	27	36	19	25.7	7	9.5	2	2.70
HIV	57	1	1.8	18	31.6	20	35.1	13	22.8	5	8.8		
VDRL	23	1	4.3	5	21.7	7	30.4	6	26.1	2	8.7	2	8.7
HBS	0												
SGTP	0												
Fevereiro													
Doações	1037	105	10.1	538	51.9	164	16	155	14.9	47	4.5	28	2.70
Infectadas	95	3	3.16	37	38.9	29	31	22	23.2	5	5.3	3	3.16
HIV	69	1	1.4	26	37.7	25	36.2	14	20.3	2	2.9	1	1.4
VDRL	30	2	6.7	11	36.7	4	13.3	8	26.7	3	10.	2	6.7
HBS	0												
SGTP	0												
Marco													
Doações	1577	232	14.7	881	55.9	198	13	165	10.5	70	4.4	31	1.97
Infectadas	93	4	4.30	45	48.4	30	32	14	15.1	4	4.3	0	0
HIV	79	4	5.1	35	44.3	23	29.1	13	16.5	4	5.1		
VDRL	18			10	55.6	7	38.9	1	5.6				
HBS	0												
SGTP	0												
Abril													
Doações	1714	170	9.92	868	50.6	322	19	212	12.4	100	5.8	42	2.45
Infectadas	127	6	4.72	44	34.6	49	39	25	19.7	5	3.9	1	0.79
HIV	110	6	5.5	38	34.5	41	37.3	21	19.1	4	3.6		
VDRL	20			6	30.	8	40.	4	20.	1	5.	1	5.
HBS	0												
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Maio													
Doações	1520	228	15	729	48	248	16	191	12.6	92	6.1	32	2.11
Infectadas	102	6	5.88	41	40.2	35	34	19	18.6	5	4.9	1	0.98
HIV	95	4	4.2	34	35.8	34	35.8	19	20.	3	3.2	1	1.1
VDRL	12	2	16.7	7	58.3	1	8.3			2	16.7		
HBS	0												
SGTP	0												
Junho													
Doações	1537	203	13.2	785	51.1	259	17	168	10.9	83	5.4	39	2.54
Infectadas	56	3	5.36	23	41.1	17	30	10	17.9	2	3.6	1	1.79
HIV	51	3	5.9	20	39.2	16	31.4	9	17.6	2	3.9	1	2.
VDRL	5			3	60.	1	20.	1	20.				
HBS	0												
SGTP	0												
Julho													
Doações	1389	272	19.6	690	49.7	207	15	128	9.22	77	5.5	15	1.08
Infectadas	55	2	3.64	21	38.2	29	53	5	9.09	3	5.5	0	0
HIV	53	2	3.8	19	35.8	24	45.3	5	9.4	3	5.7		
VDRL	7			2	28.6	5	71.4						
HBS	0												
SGTP	0												
Agosto													
Doações	1501	251	16.7	782	52.1	205	14	161	10.7	69	4.6	33	2.2
Infectadas	61	1	1.64	22	36.1	24	39	9	14.8	3	4.9	4	6.56
HIV	59	1	1.7	22	37.3	22	37.3	8	13.6	3	5.1	3	5.1
VDRL	4					2	50.	1	25.			1	25.
HBS	0												
SGTP	0												
Setembro													

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Doações	1275	285	22.4	603	47.3	182	14	129	10.1	54	4.2	22	1.73
Infectadas	40	5	12.5	22	55	10	25	3	7.5	3	7.5	0	0
HIV	36	3	8.3	18	50.	9	25.	3	8.3	3	8.3		
VDRL	7	2	28.6	4	57.1	1	14.3						
HBS	0												
SGTP	0												
Outubro													
Doações	966	110	11.4	484	50.1	159	16	133	13.8	57	5.9	23	2.38
Infectadas	48	3	6.25	19	39.6	18	38	4	8.33	5	10	3	6.25
HIV	40	2	5.	14	35.	15	37.5	3	7.5	4	10.	2	5.
VDRL	12	1	8.3	5	41.7	3	25.	1	8.3	1	8.3	1	8.3
HBS	0												
SGTP	0												
Novembro													
Doações	1043	52	4.99	479	45.9	226	22	166	15.9	92	8.8	28	2.68
Infectadas	60	0	0	22	36.7	26	43	8	13.3	3	5	1	1.67
HIV	51			19	37.3	22	43.1	7	13.7	2	3.9	1	2.
VDRL	9			3	33.3	4	44.4	1	11.1	1	11.1		
HBS	0												
SGTP	0												
Dezembro													
Doações	1312	72	5.49	615	46.9	255	19	236	18	97	7.4	37	2.82
Infectadas	102	3	2.94	37	36.3	38	37	17	16.7	4	3.9	3	2.94
HIV	89	3	3.4	33	37.1	33	37.1	17	19.1	3	3.4		
VDRL	13			4	30.8	5	38.5			1	7.7	3	23.1
HBS	0												
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%

Total do período

Doações	15862	2015	12.7	7890	49.7E+03	17	2044	12.9	911	5.7	358	2.26	
Infectadas	913	38	4.16	356	39	332	36	155	17	49	5.4	19	2.08
HIV	789	30	3.8	296	37.5	284	36.	132	16.7	38	4.8	9	1.1
VDRL	160	8	5.	60	37.5	48	30.	23	14.4	11	6.9	10	6.3
HBS	0												
SGTP	0												

Doações por Tipo de Dador

Período : Desde 01-Jan-05 até 31-Dec-05

Ano / Mes	Nº de Doações				Total
	Voluntários		Repositores		
	Nº	%	Nº	%	
2005					
Janeiro	1114	75	361	24	1492
Fevereiro	1306	75	425	24	1745
Marco	1557	78	438	22	2007
Abril	1457	76	452	24	1915
Maió	1614	78	440	21	2059
Junho	1469	78	407	22	1883
Julho	920	74	318	26	1245
Agosto	1069	79	270	20	1347
Setembro	737	78	201	21	946
Outubro	1046	77	312	23	1367
Novembro	1090	81	249	19	1345
Dezembro	998	82	218	18	1219
Totais do Período	14377	77.42	4091	22.03	18570

Doenças Infecciosas por Faixa Etária

Período : Desde 01-Jan-05 até 31-Dec-05

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
2005													
Janeiro													
Doações	1492	133	8.91	660	44.2	313	21	230	15.4	116	7.8	40	2.68
Infectadas	71	2	2.82	31	43.7	23	32	10	14.1	5	7.0	1	1.41
HIV	63	1	1.6	27	42.9	21	33.3	9	14.3	4	6.3	1	1.6
VDRL	3			2	66.7					1	33.3		
HBS	5	1	20.	2	40.	1	20.	1	20.				
SGTP	0												
Fevereiro													
Doações	1745	319	18.3	832	47.7	273	16	220	12.6	77	4.4	24	1.38
Infectadas	103	8	7.77	43	41.7	31	30	19	18.4	2	1.9	0	0
HIV	97	7	7.2	40	41.2	30	30.9	18	18.6	2	2.1		
VDRL	4	1	25.	1	25.	1	25.	1	25.				
HBS	2			2	100.								
SGTP	0												
Marco													
Doações	2007	494	24.6	921	45.9	278	14	180	8.97	100	5	34	1.69
Infectadas	39	8	20.5	25	64.1	9	23	1	2.56	1	2.6	0	0
HIV	13			8	61.5	4	30.8			1	7.7		
VDRL	3			2	66.7	1	33.3						
HBS	26	8	30.8	15	57.7	2	7.7	1	3.8				
SGTP	0												
Abril													
Doações	1915	316	16.5	916	47.8	306	16	244	12.7	103	5.4	30	1.57
Infectadas	155	16	10.3	85	54.8	43	28	18	11.6	6	3.9	2	1.29
HIV	77	5	6.5	38	49.4	19	24.7	11	14.3	3	3.9	1	1.3
VDRL	4			3	75.			1	25.				
HBS	77	11	14.3	44	57.1	12	15.6	6	7.8	3	3.9	1	1.3
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Maio													
Doações	2059	564	27.4	868	42.2	280	14	226	11	90	4.4	31	1.51
Infectadas	146	29	19.9	73	50	36	25	12	8.22	6	4.1	0	0
HIV	66	6	9.1	29	43.9	18	27.3	9	13.6	4	6.1		
VDRL	2					2	100.						
HBS	80	23	28.8	44	55.	8	10.	3	3.8	2	2.5		
SGTP	0												
Junho													
Doações	1883	559	29.7	787	41.8	228	12	206	10.9	75	4	28	1.49
Infectadas	139	26	18.7	62	44.6	41	29	18	12.9	2	1.4	2	1.44
HIV	61	6	9.8	25	41.	17	27.9	10	16.4	2	3.3	1	1.6
VDRL	0												
HBS	78	20	25.6	37	47.4	12	15.4	8	10.3			1	1.3
SGTP	0												
Julho													
Doações	1245	303	24.3	544	43.7	206	17	108	8.67	65	5.2	19	1.53
Infectadas	65	16	24.6	28	43.1	17	26	8	12.3	3	4.6	0	0
HIV	30	1	3.3	13	43.3	11	36.7	4	13.3	1	3.3		
VDRL	2							1	50.	1	50.		
HBS	37	15	40.5	15	40.5	3	8.1	3	8.1	1	2.7		
SGTP	0												
Agosto													
Doações	1347	292	21.7	618	45.9	193	14	152	11.3	77	5.7	15	1.11
Infectadas	31	3	9.68	17	54.8	11	35	3	9.68	1	3.2	0	0
HIV	14	1	7.1	8	57.1	3	21.4	1	7.1	1	7.1		
VDRL	0												
HBS	17	2	11.8	9	52.9	4	23.5	2	11.8				
SGTP	0												
Setembro													

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		N°	%	N°	%	N°	%	N°	%	N°	%	N°	%
Doações	946	291	30.8	357	37.7	138	15	95	10.0	50	5.3	15	1.59
Infectadas	57	26	45.6	18	31.6	8	14	5	8.77	2	3.5	0	0
HIV	18	2	11.1	7	38.9	4	22.2	4	22.2	1	5.6		
VDRL	0												
HBS	39	24	61.5	11	28.2	2	5.1	1	2.6	1	2.6		
SGTP	0												
Outubro													
Doações	1367	229	16.8	666	48.7	244	18	145	10.6	63	4.6	20	1.46
Infectadas	27	5	18.5	12	44.4	8	30	4	14.8	0	0	1	3.70
HIV	22	3	13.6	9	40.9	6	27.3	3	13.6			1	4.5
VDRL	4	1	25.	1	25.	2	50.						
HBS	4	1	25.	2	50.			1	25.				
SGTP	0												
Novembro													
Doações	1345	129	9.59	719	53.5	240	18	162	12.0	75	5.6	20	1.49
Infectadas	23	3	13.0	13	56.5	3	13	2	8.7	2	8.7	0	0
HIV	10	1	10.	5	50.	3	30.	1	10.				
VDRL	0												
HBS	13	2	15.4	8	61.5			1	7.7	2	15.4		
SGTP	0												
Dezembro													
Doações	1219	162	13.3	538	44.1	227	19	179	14.7	91	7.5	22	1.80
Infectadas	42	4	9.52	15	35.7	12	29	8	19.0	2	4.8	1	2.38
HIV	33			11	33.3	12	36.4	7	21.2	2	6.1	1	3.
VDRL	0												
HBS	9	4	44.4	4	44.4			1	11.1				
SGTP	0												

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%

Total do período

Doações	18570	3791	20.4	8426	45.4E+03	16	2147	11.6	982	5.3	298	1.60	
Infectadas	898	146	16.3	422	47	242	27	108	12.0	32	3.6	7	0.78
HIV	504	33	6.5	220	43.7	148	29.4	77	15.3	21	4.2	5	1.
VDRL	22	2	9.1	9	40.9	6	27.3	3	13.6	2	9.1		
HBS	387	111	28.7	193	49.9	44	11.4	28	7.2	9	2.3	2	.5
SGTP	0												

Dados do BSHGJM

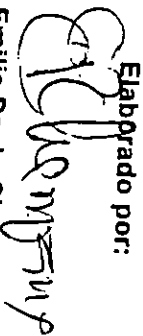
HOSPITAL GERAL JOSÉ MACAMO
BANCO DE SANGUE

DADOS ESTATÍSTICOS DO ANO 2000

Mês	Benévolo	Repositores	Palestra	Total	C. Glob.	PFC	RPR	HIV	Inutilizado
Janeiro	47	247	0	294	676	6	29	50	50
Fevereiro	45	216	99	360	641	13	19	32	32
Março	41	214	0	255	840	23	31	51	51
Abril	28	228	156	412	902	30	28	67	67
Mai	48	270	0	317	856	34	29	84	84
Junho	53	157	107	317	551	22	10	44	44
Julho	34	182	0	216	688	38	16	22	22
Agosto	41	169	281	491	610	10	47	80	80
Setembro	28	107	0	135	533	40	15	22	22
Outubro	50	210	0	260	555	24	17	27	27
Novembro	42	275	149	499	977	33	31	60	60
Dezembro	52	202	0	254	634	40	24	32	32
TOTAL	509	2477	792	3810	8463	313	296	571	571

Maputo, 29 de Janeiro de 2001

Elaborado por:



Emilia Paulo Chemane

(Técnica de Laboratorio "C" Principal)

HOSPITAL GERAL JOSÉ MACCAMO
BANCO DE SANGUE

DADOS ESTATÍSTICOS DO ANO 2001

Mês	Benévolo	Repositores	Palestras	Total	C. Globs	PFC	RPR	HIV	Inutilizado
Janeiro	46	225	0	271	665	22	31	34	34
Fevereiro	25	195	174	394	665	22	18	54	54
Março	48	196	0	244	591	31	19	42	42
Abril	48	117	0	165	343	29	16	29	29
Maio	42	97	31	170	360	46	13	26	26
Junho	34	112	47	193	348	34	14	24	24
Julho	32	105	0	137	305	42	7	16	16
Agosto	45	111	159	315	253	60	20	26	26
Setembro	42	102	0	144	439	61	9	20	20
Outubro	60	120	0	180	464	63	10	21	21
Novembro	46	127	0	173	422	43	15	19	19
Dezembro	77	131	0	208	511	39	16	26	26
Total	545	1638	411	2594	5366	492	188	337	337

MINISTÉRIO DA SAÚDE
DIRECÇÃO DE SAÚDE DA CIDADE DE MAPUTO
PROGRAMA NACIONAL DE TRANSFUSÃO DE SANGUE

HOSPITAL: GERAL JOSÉ MACAMO

MOVIMENTO ESTATÍSTICO

ANO 2002

Mes	DOAÇÕES DE SANGUE				SELEÇÃO DE DADORES			PRODUÇÃO			PEDIDO DE SANGUE			SEROPositivIDADE				OBSERVAÇÕES
	BENEVOLO	REPOSIT.	PALESTRA	TOTAL	CANDIDAT. A DADORES	AUTO EXCLUSÃO	DADORES INAPTIOS	SANGUE TOTAL	CONCENT. GLOB. VERM.	PLASMA F.C	PED. NÃO ATENDIDO	PEDIDO ATENDIDO	TOTAL	POSIT. RPR	POSITIVO DETERMINE	POSIT UNIGOLD	SANGUE INUTILIZ	
Jan	59	135	-	194	200	1	6	194	471	57	10	528	538	17	22	22	23	
Fev.	50	84	69	203	209	-	6	203	326	49	11	375	386	19	30	30	32	
Mar	67	122	58	247	250	-	3	247	552	59	7	611	618	17	20	20	26	
Abril	61	99	174	334	344	-	10	334	538	65	12	603	615	25	32	32	34	
Mai	97	102	127	326	333	-	7	326	605	88	8	693	701	16	29	29	31	
Junho	36	107	54	197	208	-	11	197	565	63	9	628	637	5	27	27	33	
Julho	52	119	-	171	181	-	10	171	354	64	13	418	431	13	12	12	14	
Agosto	89	115	70	274	280	-	6	274	310	51	10	361	371	14	23	21	26	
Set.	52	90	-	142	150	-	8	142	402	82	5	484	489	18	13	11	13	
Out.	45	128	-	173	180	-	7	173	358	56	5	414	419	16	25	24	25	
Nov.	37	115	-	152	159	-	7	152	356	50	6	406	412	11	14	14	17	
Dez.	71	91	-	162	165	-	3	162	339	34	8	373	381	5	6	6	7	
Total	716	1307	552	2575	2659	1	84	2575	5176	718	104	5894	5998	176	253	248	281	

Maputo, 25 de Fevereiro de 2003

Assinatura do Responsável
Luiz Fernando
 Data: 22/1/2002



MINISTÉRIO DA SAÚDE
 DIREÇÃO NACIONAL DE SAÚDE
 PROGRAMA NACIONAL DE TRANSFUÇÃO DE SANGUE
 MOVIMENTO ESTATÍSTICO

HOSPITAL: *General José Flacciano*

ANO: *2002*
 MÊS: *Janeiro*

MÊS	DOAÇÕES DE SANGUE			SELEÇÃO DE DADORES			PRODUÇÃO						PERDIDOS DE SANGUE			SEROPositividade								
	BENEVOLO	RECEIT.	PALESTRA	TOTAL	CAMPAINHA D. DADORES	AUTO-EXCLUIDO	DADORES HÁPTICOS	SANGUE TOTAL	CONCENT. DE VERM.	PLASMA F.C.	PLASMA R.F.	CONC. PLASMA	OUTRAS	PERD. MÚLTIPLAS	PERDIDOS ATEND.	TOTAL	POST. RPR	POST. GENESEC.	POSITIVO DETERMINE	POSIT. UNOCL.	POSIT. BIAO	SANGUE NATIVE	OBSERVAÇÕES	
JAN.	58	144	-	202	206	-	4	306	432	34				5	466		19		22	22		23		
FEV.	40	123	-	163	175	-	4	178	361	31				4	396		8		21	22		18		
MAR.	52	101	-	145	164	-	6	310	475	35				2	510		14		23	21		23		
ABRIL	60	131	64	263	266	4	3	260	601	56				10	651		14		25	23		25		
MAIO	64	80	66	210	215	1	4	215	464	81				9	521		14		25	23		25		
JUNHO	55	114	21	180	196	3	13	196	438	44				5	474		6		20	11		20		
JULHO	62	01	-	143	156	2	7	150	325	46				10	370		10		16	12		16		
AGOS.	76	100	99	275	260	1	4	260	278	44				9	322		35		30	28		28		
SET.	59	22	161	242	220	2	6	220	255	43				-	438		60		21	19		20		
OUT.	49	114	-	163	165	2	5	165	475	33				10	508		15		21	20		20		
NOV.	27	81	220	324	325	2	10	325	379	29				-	408		13		29	22		29		
DEZ.	68	83	-	131	136	2	3	136	447	69				5	506		4		10	10		10		
TOTAL	677	1224	786	2881	2746	14	69	2746	5060	577				69	5581		214		259	229		249		

Assinatura do Rep. Responsável
Wilson F. P. Barbosa
 Data: 28/12/10

MINISTÉRIO DA SAÚDE
 DIREÇÃO NACIONAL DE SAÚDE
 PROGRAMA NACIONAL DE TRANSFUSÃO DE SANGUE
 MOVIMENTO ESTATÍSTICO Anual

HOSPITAL: *General José Facama*

ANO *2004*
 Mês *Jan*

MÊS	DOAÇÕES DE SANGUE			SELEÇÃO DE DADORES			PRODUÇÃO						PERDAS DE SANGUE			SEROPOSITIVIDADE								
	BENÉVOLA	REQUISI	PALESTRA	TOTAL	CAMP. 1	AUTO-EXCLUSÃO	DADORES RAPTOS	SANGUE TOTAL	CONCENT. GL. DE VENA	PLASMA F.C.	PLASMA R.P.	CONC. PLAS.	OUTRAS	PER. MIO ATENDIDO	PERDAS ATEND.	TOTAL	PORT. RPR	PORT. GENEIC.	POSITIVO DETERMINE	PORT. LABOR. IBI	PORT. IBI	SANGUE MULTIC.	OBSERVAÇÕES	
JAN.	47	87	86	220	230	-	10	220	220	220				05	396	401	13	-	25	24	-	-	25	
FEB.	40	120	-	160	165	-	05	160	160	160				06	409	415	9	-	16	16	-	-	17	
MAR.	81	125	16	222	225	-	03	222	222	222				05	555	560	18	-	30	25	-	-	31	
ABRIL	54	171	83	308	318	-	10	308	308	308				04	565	569	14	-	45	45	-	-	45	
MAIO	55	199	65	319	320	-	01	319	319	319				04	537	541	20	-	31	28	-	-	30	
JUNHO	61	159	-	220	225	-	05	220	220	220				05	320	325	14	26	12	20	-	-	26	
JULHO	56	134	-	190	196	01	05	190	190	190				04	286	290	09	25	12	18	-	-	25	
AGOS.	51	102	327	480	490	01	09	480	480	480				-	512	512	11	44	12	28	-	-	44	
SET.	62	71	165	298	300	00	02	298	298	298				-	407	407	08	58	-	18	-	-	59	
OUT.	36	75	109	220	225	-	05	220	220	220				02	430	432	05	35	-	14	-	-	37	
NOV.	47	130	97	274	276	-	02	274	274	274				04	335	339	25	119	-	34	-	-	122	
DEZ.	59	136	124	319	325	-	06	319	319	319				-	323	323	24	61	-	26	-	-	62	
TOTAL	649	1508	1072	3290	3295	02	63	3230	3230	3230				39	5125	5164	188	336	171	296	-	-	313	

HOSPITAL GERAL JOSÉ MACAMO
BANCO DE SANGUE
DADOS ESTATÍSTICOS DO ANO 2005

Mês	Benévolo	Repositore	Palestras	Total	P. RPR	P. Genesc.	P. Det.	P. Unigold	Inutilizado
Janeiro	65	161	99	325	14	79		50	
Fevereiro	35	74	49	158	17	18		14	
Março	51	126	226	403	11	40		23	
Abril	61	167	0	228	17	20	14	24	
Maior	54	141	63	258	21	48		21	
Junho	56	109	189	354	17	40		24	
Julho	46	157	97	300	47	80	2	41	
Agosto	44	118	189	351	25		30	21	
Setembro	50	82	53	185	5	16	30	25	
Outubro	57	120	74	251	14	10	18	16	
Novembro	56	109	72	237	15	37		21	
Dezembro	68	100	97	265	6	10		6	

Maputo 07 de Julho de 2006

A responsável do banco de sangue