



**UNIVERSIDADE EDUARDO MONDLANE**

**Faculdade de Ciências**

**Departamento de Matemática e Informática**

**Curso de Licenciatura em Informática**

**TRABALHO DE LICENCIATURA**

**APLICAÇÃO DE SÉRIES TEMPORAIS EM REDES  
NEURONAIS**

**Um Estudo de Caso do Banco de Sangue do Hospital Central de Maputo**

**Supervisor:** Dr. Milagre David Mula

**Co - Supervisor:** Dr. Alberto Chicafo Mulenga

**Estudante:** Sitói Marcelino Monteiro

**ET-19**

**Maputo, Outubro de 2007**



**UNIVERSIDADE EDUARDO MONDLANE**

**Faculdade de Ciências**

**Departamento de Matemática e Informática**

**Curso de Licenciatura em Informática**

**TRABALHO DE LICENCIATURA**

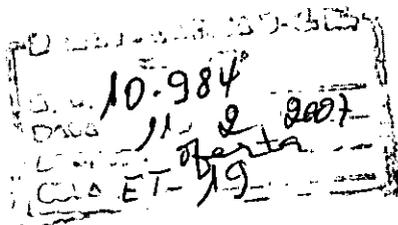
**APLICAÇÃO DE SÉRIES TEMPORAIS EM REDES NEURONAIS**

**Um Estudo de Caso do Banco de Sangue do Hospital Central de Maputo**

**Supervisor:** Dr. Milagre David Mula

**Co - Supervisor:** Dr. Alberto Chicafo Mulenga

**Estudante:** Sitói Marcelino Monteiro



**Maputo, Outubro de 2007**



### **Dedicatória**

Dedico este trabalho a família Monteiro pelo carinho que me deram durante a minha formação académica e em memória dos meus entequeridos avos, irmãos, tios e primos.



## **Agradecimentos**

Primeiramente agradeço à Deus, que me concedeu a vida, proporcionando-me momentos maravilhosos em minha existência

Aos meus pais Marcelino Monteiro e Luísa Pequenino, por me terem feito homem e estarem sempre ao meu lado nos momentos que mais precisei e preciso de apoio.

Ao meu Supervisor dr. Milagre D. Mula e ao meu Co – Supervisor dr. Alberto C. Mulenga, pelas suas valiosas críticas, sugestões, paciência e por me terem suportado durante o tempo da realização deste Trabalho. Também vai o meu especial agradecimento ao Eng. Macia pelas suas análises criteriosas e encorajadoras.

Agradeço aos meus tios José Joaquim Soares e sua esposa Lurdes Lucas pelo incansável apoio que me deram e me tem dado até a este momento.

Aos meus irmãos, Horácio, Lizete, Sílvio, Quitéria, Nelson, Sérgio, Celina e Maria do Céu que aceitaram de forma indulgente a minha ausência nestes últimos anos, tão difíceis em suas vidas e a minha prima Stela que soube me dar força quando triste estivesse.

Agradeço ainda ao Banco de Sangue do Hospital Central de Maputo por terem concedido os dados desejados para a realização deste trabalho, e em especial ao Dr. Joel.

Aos meus amigos e colegas Alberto Capece, Fernando Marraves, Nelson Magalhães, Ilda Sandra, Sandra Amigos, Nelson Cazonda, João J. Alguineiro, Hermes Guluve e Tadiana Couana, Óscar Chachiuva, Roberto Nhamunue, que souberam ponderar em algumas vezes as minhas absurdas contribuições durante a minha formação académica.

A todos os professores e funcionários do Departamento de Matemática e Informática que contribuíram directamente para a minha formação académica.

E, finalmente, a todas as pessoas que, directa ou indirectamente contribuíram para a realização deste Trabalho.



### Declaração de Honra

Declaro que esta monografia científica é resultado da minha investigação pessoal e das orientações dos meus supervisores, o seu conteúdo é original e todas as fontes consultadas estão devidamente mencionadas no texto, nas notas e na bibliografia.

Declaro ainda que este trabalho não foi submetido em nenhuma outra instituição para obtenção de qualquer grau académico.

Maputo, Outubro de 2007

-----  
*Sitói Marcelino Monteiro*  
-----

(*Sitói Marcelino Monteiro*)



*Epigrafo*

*“Veja, não penses que a vitória  
está perdida, se é de batalhas que  
se vive a vida tente outra vez...”*

Raul Seixas



## Resumo

De acordo com os objectivos e o caso de estudo deste trabalho, foram colhidos dados correspondentes a sangue infectado pelo vírus de HIV-SIDA no Banco de Sangue (BS) do Hospital Central de Maputo (HCM).

Os dados que ao longo da descrição deste trabalho são chamados de séries temporais foram extraídos da base de dados em uso no BS do HCM e são referentes ao sangue infectado pelo vírus de HIV – SIDA, registado desde Janeiro de 2000 à Dezembro de 2005, perfazendo um total de 72 observações. Segundo Luz et al (2000) citado em Box et al (1994) considera que um mínimo de 50 observações é satisfatório para o seu uso em modelos de previsão, sendo 72 observações adequadas para o seu uso de acordo com os objectivos deste trabalho.

Com a série obtida, foi necessário e primeiro realizar a ACP, principalmente a tendência e a sazonalidade; em segundo, identificar a característica estacionária da série visto que as RNs exigem que a série seja estacionária e que não possua factores aleatórios para que possa fornecer resultados satisfatórios quando o objectivo é usar as STS em RNs para que estas (RN) possam prever a série.

A identificação da tendência, foi possível pela comparação dos erros EPAM, EMQ e o DAM, gerados pelo software *Minitab for Windows*, sendo a tendência escolhida a tendência quadrática, por possuir menor erro de previsão. A transformação da ST em estacionária foi possível por meio da aplicação do OD que teve o seu auxílio o software *MicroSoft Excel 2003*. Para a confirmação da estacionariedade após o uso do OD o teste ADF foi usado, o que de facto veio a confirmar a estacionariedade da série.

Após a análise da série, esta foi modelada pelo Software *DTREG*, que proporcionou a melhor arquitectura (topologia) desta, quanto ao número de camadas da rede a ser usada, treinamento, número de neurónios para cada camada bem como as funções de activação para cada camada. O software *NeuralWare* foi usado para a obtenção dos resultados de previsão da série em estudo bem como os seus erros de previsão.

**Palavras chaves:** *Estacionariedade, Tendência, Previsão, Rede Neuronal, Série Temporal*



## Índice de Figuras

Figura 1. Modelo genérico de um neurónio artificial .....	24
Figura 2. Estrutura de uma RNAD ou RN Feedforward .....	30
Figura 3. Estrutura de uma RNR .....	31
Figura 4. Activação de um neurónio.....	32
Figura 5. Exemplo de uma rede de Previsão em STs .....	34



## Índice de Tabelas

Tabela 1. Potencialidades de Previsão do EPAM.....	18
Tabela 2. Funções de activação .....	32
Tabela 3. Valores dos erros das Tendências .....	39
Tabela 4. Teste ADF para confirmação da Estacionariedade .....	41
Tabela 5. Parâmetros da RN .....	43
Tabela 6. Optimização do número de neurónios na camada escondida .....	44
Tabela 7. Dados obtidos pelo software NeuralWare .....	45
Tabela 8. Valores de previsão gerados pelo s/w NeuralWare .....	46
Tabela 9. Dados usados na arquitectura da RN .....	50



## Índice de Gráficos

Gráfico 1. Representação da ST bruta .....	38
Gráfico 2. Representação da tendência Quadrática .....	39
Gráfico 3 Serie não estacionária (Modelada à partir da ST real).....	40
Gráfico 4 Série estacionária (Modelada à partir da diferenciação da primeira ordem ).....	41
Gráfico 5. Representação da ST após a estacionariedade.....	42
Gráfico 6. Comparação da ST bruta com a da Previsão .....	48
Gráfico 7. Comparação da ST bruta com a da Transformação de Previsão .....	48



## Abreviaturas

ACP	Análise das Componentes Principais
ADF	Dickey Fuller Aumentado
BS	Banco de Sangue
CE	Camada de Entrada
CI	Camada Intermediária
CS	Camada de Saída
DAM	Desvio Absoluto Médio
EM	Erro Médio
EMQ	Erro Médio Quadrático
EPAM	Erro Percentual Absoluto Médio
FEr	Função de Erro
HCM	Hospital Central de Maputo
HIV	Vírus de Imunodeficiência Humana
LS	Ligações Sinápticas
MB	Modo Batch
MMQ	Método dos Mínimos Quadráticos
MP	Modo Padrão
NE	Neurónios de Entrada
NF	Neurónio Formal
NI	Neurónios Intermediários
NP	Neurónios Periféricos
NS	Neurónios de Saída
OD	Operador Diferença
PS	Pesos Sinápticos
RD	Rede Delta
RL	Rede Linear
RN(s)	Rede(s) Neuronal(is)
RNAD	Redes Neurais de Alimentação Directa
RNC	Redes Neurais por Camada
RNF	Rede Neuronal Feedforward
RNP <sub>s</sub>	Redes Neurais Planares
RNR	Redes Neurais Recorrentes
RP	Rede Perceptron
RPMC	Rede Perceptron Multi – Camada
SIDA	Síndrome de Imunodeficiência Adquirida
ST(s)	Série(s) Temporal(ais)
S/W	Software
TE	Tendência Exponencial
TL	Tendência Linear
TQ	Tendência Quadrática



## ÍNDICE

Dedicatória .....	i
Agradecimentos.....	ii
Declaração de Honra .....	iii
Resumo.....	v
Índice de Figuras.....	vi
Índice de Tabelas .....	vii
Índice de Gráficos .....	viii
Abreviaturas.....	ix
<b>Capítulo I: INTRODUÇÃO E OBJECTIVOS.....</b>	<b>1</b>
1.1 Introdução .....	1
1.2 Definição do Problema .....	3
1.3 Objectivos .....	4
1.3.1 Objectivo geral .....	4
1.3.2 Objectivos específicos.....	4
<b>Capítulo II: REVISÃO DA LITERATURA.....</b>	<b>5</b>
2.1 Séries Temporais .....	5
2.2 Objectivo do Estudo das Séries Temporais .....	6
2.3 Movimentos Característicos das Séries Temporais .....	6
2.3.1 Tendência.....	7
2.3.1.1 Estimação da Tendência .....	8
2.3.1.1.1 Médias Móveis .....	8
2.3.1.1.2 Operador diferença (OD).....	10
2.3.1.1.3 Método a Sentimento.....	12
2.3.1.1.4 Método das Semimédias .....	12
2.3.2 Sazonalidade .....	12
2.3.3 Estacionariedade .....	13
2.3.3.1 Avaliação das Variações Estacionais. Índice de Estacionariedade.....	13
2.3.3.1.1 O Método da Percentagem Média.....	14
2.3.3.1.2 Método da Tendência ou Relação Percentual .....	14
2.3.3.1.3 Método da Média Móvel Percentual ou da Relação entre as Médias .....	15
2.3.3.1.4 O Método dos Elos Relativos .....	15
2.3.4 Destacionalização dos Dados .....	16
2.3.5 Estimação das Variações Cíclicas .....	16
2.3.6 Estimação das Variações Irregulares ou Aleatórias .....	16
2.3.7 Comparabilidade dos Dados.....	16
2.4 Análise do erro na Previsão de Séries Temporais.....	17
2.5 Raiz Unitária e Teste ADF .....	19
2.6 Autocorrelação .....	20
2.6.1 Correlação de Curto-prazo .....	22
2.6.2 Correlação negativa.....	22



<b>2.7 Generalidades de redes neuronais .....</b>	<b>22</b>
2.7.1 TIPOS DE REDES NEURONAIS.....	25
2.7.1.1 Redes Perceptron (RP) com uma Camada.....	25
2.7.1.2 Redes Perceptron Multi-camadas (RPMC) .....	25
2.7.2 Caracterização das redes neuronais.....	26
2.7.2.1 Classificação quanto à arquitectura .....	26
2.7.2.1.1 Redes Neurais Planares (RNPs).....	26
2.7.2.1.2 Redes Neurais por Camadas (RNC) .....	26
2.7.2.2 Classificação quanto ao processo de aprendizagem .....	27
2.7.3 Correção de erros.....	29
2.7.4 Topologia das RNs.....	29
2.7.5 Activação e Função de Activação .....	30
2.7.6 Implementação da Rede Neuronal .....	32
2.7.7 Aplicação das RNs .....	33
2.7.7.1 RNs Aplicadas para Previsões em STs.....	33
 <b>Capítulo III: MATERIAL E MÉTODOS.....</b>	<b>36</b>
3.1 Material .....	35
3.2 Métodos.....	35
 <b>Capítulo IV: RESULTADOS E DISCUSSÃO.....</b>	<b>38</b>
4.1 Resultados.....	37
4.1.1 Modo de análise dos Dados usadas no Hospital Central de Maputo (HCM) .....	37
4.1.2 Modelação da Série Temporal Recolhida.....	38
4.1.2.1 Análise da Série Temporal Recolhida .....	38
4.1.2.1.1 Análise da Tendência.....	38
4.1.2.1.2 Análise da Sazonalidade .....	39
4.1.2.1.3 Ajustamento da ST à Estacionária.....	40
4.1.3 Modelação via RNs.....	42
4.1.3.1 Parâmetros Usados para Modelação.....	43
4.1.3.2 Arquitectura da Rede Utilizada para a Modelação .....	43
4.1.3.3 Implementação da Rede Neuronal.....	44
4.2 Discussão dos Resultados .....	49
4.2.1 Série Temporal recolhida .....	49
4.2.2 Parâmetros da RN .....	50
4.2.3 Resultados obtidos pela RN .....	50
 <b>Capítulo V: CONCLUSÕES E RECOMENDAÇÕES.....</b>	<b>53</b>
5.1 Conclusões .....	52
5.2 Recomendações.....	52
 <b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>53</b>
 <b>ANEXOS.....</b>	<b>56</b>



# **CAPÍTULO I**

## **Introdução e Objectivos**

---

---

### **1.1 Introdução**

As redes neuronais (RNs), segundo Garcia (1995), são estruturas de inspirações biológicas capazes de processar informação, sendo a sua unidade fundamental o neurónio formal (NF). Estas redes estão propondo soluções interessantes a problemas em várias áreas de administração como finanças, saúde, marketing, vendas e compras ou mesmo recursos humanos (Almeida, 1995).

O grande interesse no uso de RNs desenvolveu-se a partir dos meados da década de oitenta, quando se re-inventou o algoritmo de treino de backpropagation (Russel et al; 1995). Desde então as RNs são objecto de atenção por parte da comunidade científica e usadas na prática. Este ressurgimento também se deve ao facto das RNs se armarem como ferramentas poderosas no reconhecimento de padrões, ultrapassando outras técnicas, quer em fiabilidade quer em velocidade de processamento (Azoff, 1995).

As redes neuronais ao invés de trabalhar com regras explícitas como os sistemas especialistas, elas utilizam critérios mais complexos e implícitos baseados em aprendizagem a partir de exemplos. Não há no caso das RNs, uma codificação de programas a fim de introduzir o conhecimento sobre um problema.

Através de um processo iterativo chamado processo de aprendizagem, as redes neuronais lêem os exemplos fornecidos sobre um problema e criam assim um modelo de resolução do problema. As RNs são bem adaptadas a dois tipos de tarefas: reconhecimento de formas e generalização (Dutta et al, 1992).

Segundo Moreira (1997) as RNs têm sido utilizadas na modelação de memória associativa, reconhecimento de padrões, representação de funções booleanas, representação de funções contínuas, previsão de séries temporais e optimização.



De acordo com a enciclopédia livre (Wikipédia, 2007), as redes neuronais se saem particularmente bem nas seguintes tarefas:

- Aproximação de funções;
- Previsão de séries temporais (STs);
- Classificações;
- Reconhecimento de Padrões.

Apesar de se saírem bem nas tarefas acima, elas actualmente são mais exploradas na previsão de STs e na avaliação de Riscos.

Actualmente assiste-se a uma mundialização da economia e por conseguinte a um acréscimo de concorrência entre as organizações cujas estruturas, processos e tecnologias, se optimizados, lhes darão vantagens competitivas (Correia, 1995). Contudo, inovar num ambiente de incerteza pode ser desastroso, pelo que é natural pressupor que as organizações estão interessadas em obter previsões fundamentadas sobre o seu futuro ( Hanke et al, 1989).

Este desejo de compreender o passado e prever o futuro impulsiona a procura de critérios que expliquem o comportamento de dados, fenómenos ou acontecimentos. Se são conhecidas equações determinísticas que os explicam, então é possível efectuar previsões. No entanto, na ausência de regras que definam o comportamento de um sistema, procura-se determinar o seu comportamento futuro a partir de observações do passado (Weigend et al, 1994).

A análise de uma série temporal (ST), pode ajudar na previsão de resultados desejados já que uma das características importante dos dados numa ST é o facto das suas observações vizinhas serem dependentes e nesse caso é necessário analisar e modelar esta dependência, daí, a necessidade de serem usadas na análise dos resultados obtidos aplicando as RNs.

Os autores Latorre e Cardoso (2001) afirmam que ao se analisar uma ST, primeiramente deseja-se modelar o fenómeno estudado para, a partir daí, descrever o comportamento da série, fazer estimativas e, por último, avaliar quais os factores que influenciaram o comportamento da série, buscando definir relações de causa e efeito entre duas ou mais séries. Para tal, há um conjunto de técnicas estatísticas disponíveis que dependem do modelo definido (ou estimado para a série), bem como do tipo de série analisada e do objectivo do trabalho.



## **1.2 Definição do Problema**

A busca de resultados a curto e longo prazo a partir de observações do comportamento de uma ST, tem estimulado os investigadores no sentido de encontrar métodos capazes de dar respostas a partir de processos matemáticos, que apresentem um grau de aderência à serie de dados compilados no tempo, de forma a satisfazer as necessidades de obter valores confiáveis para o futuro (Martin, 2000).

Como concluíram Eastman e Mckendry (1991), existe uma grande falta de técnicas estatísticas que possam lidar com eventos espaço temporais e que apresentam uma ligação simultânea no espaço e no tempo. Porém, segundo eles, a Análise de Componentes Principais (ACP) apresenta possibilidades perturbantes, tais como, a tendência e a sazonalidade.

Com a ajuda das RNs por utilizarem metodologias que permitem, a partir de valores observados no passado de uma variável em estudo e com uma arquitectura na forma de neurónios, a exemplo do funcionamento do cérebro humano, estabelece uma resposta adequada para valores que extrapolem o tempo presente.

Segundo Martin (2000), os problemas de previsão de valores futuros e inconsistentes em STs, podem ser resolvidos empregando técnicas de RNs aplicadas a STs, por estas apresentarem resultados bastante satisfatórios para perspectivas de curto e longo prazo, permitindo, a partir de um processo de aprendizagem, ensinar a rede a utilizar os valores históricos para dar as respostas de previsões futuras de uma determinada série.

É nesta perspectiva que surge o presente trabalho com o tema “ **Aplicação de Séries Temporais em Redes Neurais** e que tem como caso de estudo o **Banco de Sangue do Hospital Central de Maputo** ”, com uma atenção virada apenas a previsão da ST recolhida por meio de RNs.

Este tema tem como relevância a chamada de atenção aos agentes do BS do HCM no que diz respeito a tomada de atenção nas precauções a tomar com o stock de sangue, visto que a falta deste precioso liquido pode causar sérios constrangimentos no trabalho clínico.



### **1.3 Objectivos**

#### **1.3.1 Objectivo geral**

- Aplicar as RNs à ST recolhida no BS do HCM de modo a prever as unidades de sangue inutilizadas pelo vírus de HIV – SIDA.

#### **1.3.2 Objectivos específicos**

- Identificar e descrever as técnicas de previsão em séries temporais usadas no HCM;
- Verificar a estacionaridade da ST recolhida e identificar os principais movimentos desta.
- Identificar a melhor topologia das RNs usadas para a previsão de STs, seu treinamento, funções de activação e a forma da normalização dos dados.
- Aplicar a ST recolhida em RNs para a avaliação do desempenho das RNs.



# CAPÍTULO II

## Revisão da Literatura

---

### 2.1 Séries Temporais

Uma série temporal (ST), também denominada série histórica, é uma sequência de dados obtidos em intervalos regulares de tempo durante um período específico. Este conjunto pode ser obtido através de observações periódicas do evento de interesse como, por exemplo, o valor máximo diário da concentração do ozono no ar num país, ou através de processos de como o total mensal de óbitos por cancro numa unidade hospitalar. Se a série temporal for denominada como  $X$ , o valor da série no momento " $t$ " pode ser escrito como  $X_t$  ( $t=1,2,\dots,n$ ) (Morettin, 1985).

Denomina-se trajectória de um processo, a curva obtida no gráfico da série temporal e o conjunto de todas as possíveis trajectórias é denominado como um processo estocástico (Lattore et al, 2001). Considera-se que uma ST é uma amostra deste processo.

Lattore e Cardoso (2001), afirmam que para analisar uma ST, primeiramente deseja-se modelar o fenómeno estudado para, a partir daí, descrever o comportamento da série, fazer estimativas e, por último, avaliar quais os factores que influenciam o comportamento da série, buscando definir relações de causa e efeito entre duas ou mais séries. As séries temporais podem exibir pelo menos até três características diferentes no seu comportamento, nomeadamente, a tendência, sazonalidade e a variação aleatória (Breseghello, 2005).

#### Terminologias :

- Séries temporais **contínuas**: Medidas no tempo contínuo (Ex: medidas de temperatura a partir de um termógrafo).
- Séries temporais **discretas**: Observações tomadas em intervalos de tempo específico e igualmente espaçadas ( Ex: temperatura média mensal).



## 2.2 Objectivo do Estudo das Séries Temporais

Os objectivos principais do estudo das séries temporais são: descrição, modelação, previsão e controlo.

A **descrição** de uma série temporal é um objectivo modesto<sup>1</sup>, mas quase sempre presente, no estudo de uma série e envolve: calculo de estatísticas sumárias da série; representação gráfica dos dados da série e de função ( empírica) dos mesmos; e o ajustamento de características simples da série (Exemplo: a função do valor esperado  $\mu = E[X, ]$ )

A **modelação** consiste na explicação de uma série como tendo sido gerada a partir de um certo modelo e envolve a estimação de parâmetro e a avaliação da qualidade do diagnóstico, em termos da qualidade estatística e do ajustamento do modelo(s) proposto(s)

A **previsão** consiste na estimação de valores futuros de uma serie temporal, o que facilita a realização de planos a curto, médio e longo prazo bem como a tomada de decisões apropriadas.

A sua importância advém do facto de ser habitualmente vantajoso, nomeadamente em conhecer a evolução provável de uma série no futuro.

O **Controlo** consiste na monitorização dos valores da série tendo em vista a detenção de alterações das características da mesma, as quais podem indicar que: o modelo proposto não é valido; o modelo que estava a ser usado deixou de ser valido (exemplo: a série correspondente a um índice pode sofrer uma roptura por mudança de modo de calculo do índice); e alguns dos parâmetro do modelo mudaram de valor.

Quando os resultados obtidos neta fase não se ajustam ao conjunto dos valores observados (brutos), deve-se voltar a fase da modelação dos dados (Pires, 2001).

## 2.3 Movimentos Característicos das Séries Temporais

Segundo Cortez (1997) os modelos dos métodos de previsão das STs são criados baseando-se na tendência e/ou na sazonalidade por serem as principais características que influenciam as observações de uma ST.

---

<sup>1</sup> Refere-se a moderado.



É interessante imaginar que o gráfico de uma ST, é descrito por um ponto que se move com o decorrer do tempo, de alguma forma análoga a trajectória de uma partícula material que se desloca sob a influência de forças físicas. Entretanto, o movimento pode ser provocados por uma combinação de forças económicas, sociológicas, psicológicas e outras.

Experiências realizadas com muitos exemplos de STs revelaram certos movimentos ou variações características, alguns dos quais, ou todos, estão presentes em graus diversos (Spiegel, 1994). A análise desses movimentos é de grande valor em vários casos, um dos quais é o problema de previsão de movimentos futuros. Em consequência, não deve constituir surpresa pelo facto de muitas indústrias e sectores governamentais estarem profundamente interessado nesse importante assunto.

### 2.3.1 Tendência

Tendência pode ser definida como uma mudança de longo-termo no nível médio ou seja é o efeito de longo prazo na média (Chatfield, 1996). As tendências mais simples conhecidas como *tendência linear mais ruído*, para qual a observação no tempo  $t$  é uma variável aleatória  $X_t$  dada por:

$$X_t = \alpha + \beta t + \varepsilon_t \quad (2.1)$$

Onde  $\alpha$ ,  $\beta$  são constantes e  $\varepsilon_t$  representa o erro aleatório com média igual a zero (conhecido também como ruído branco). O nível médio no tempo  $t$  é dado por  $m_t = (\alpha + \beta t)$ , o que é algumas vezes chamado de *termo de tendência*. Alguns autores preferem descrever o coeficiente angular  $\beta_t$  como a tendência. Em outras palavras, a tendência é a mudança no nível médio por unidade de tempo.

A tendência na equação (2.1) é uma função determinística do tempo e em algumas vezes é chamada de *tendência global*. É geralmente não realística, e por isso existe agora uma maior ênfase em *tendências locais* onde os parâmetros  $\alpha$  e  $\beta$  variam no tempo. As tendências também podem ser não-lineares. Um crescimento exponencial ou quadrático são alguns exemplos.

Assim deve-se ter na mente que a análise de uma ST que exhibe tendência depende se o pesquisador quer exactamente medir essa tendência ou se quer remover de forma a analisar



flutuações locais. Quando os dados apresentam sazonalidade, é bom começar calculando médias anuais sucessivas porque estas podem fornecer uma descrição simples das tendências implícitas.

### 2.3.1.1 Estimação da Tendência

A estimação da tendência pode ser obtida de várias maneiras possíveis.

#### 2.3.1.1.1 Médias Móveis

Dado um conjunto de números  $y_1, y_2, y_3, \dots$  define-se uma média móvel de ordem  $N$ , a que é obtida pela sequência das médias aritméticas

$$\frac{y_1 + y_2 + \dots + y_N}{N}, \frac{y_2 + y_3 + \dots + y_{N+1}}{N}, \frac{y_3 + y_4 + \dots + y_{N+2}}{N}, \dots \quad (2.2)$$

A soma do numerador da expressão (2.2) é denominado *Totais móveis de ordem N*. Costuma-se localizar cada número da média móvel em sua posição apropriada em relação aos dados originais.

Se os dados são anualmente ou mensalmente, as médias móveis de ordem  $N$  são denominados, respectivamente, *média móvel de N anos* ou *N meses*. É claro que pode ser usada qualquer outra unidade de tempo.

As médias móveis têm a propriedade de tenderem a reduzir o total da variação que se apresenta em um conjunto de dados. No caso das STs, essa propriedade é frequentemente usada para eliminar flutuações indesejadas e o processo é denominado *alisamento da ST*.

Se na expressão (2.2) forem usadas as médias aritméticas ponderadas, sendo os pesos especificados antecipadamente, a sequência, resultado é denominada *média móvel ponderada de ordem N*.

Uma outra maneira de calcular as médias móveis é considerar que instantes sucessivos de observação da série estão igualmente espaçadas e, sem perda de generalidades, designar-se a série em análise por  $X = \{X_t; t \in Z\}$ .



Uma média móvel de  $X$  é uma série  $Y = \{Y_t; t \in Z\}$  da forma

$$Y_t = \sum_{j=-s}^p w_j X_{t+j}, t \in Z \quad (2.3)$$

para certos pesos  $w_j$ , com  $s, p \geq 0$ . a função  $m_t = Y_t$  é uma boa alternativa para ajustar a tendência da série  $X$ , especialmente se os valores da série não oscilarem muito bruscamente.

Considerando o caso em que  $p=s>0$ , diz-se que  $Y$  é uma média móvel de ordem  $(2p+1)$  de  $X$ . Se obtivermos uma realização  $\{X_t; t=1,2,\dots,n\}$  de  $X$ , uma correspondente realização de  $Y$  dada por:

$$y_t = \sum_{j=\max(-p,1-t)}^{\min(p,n-t)} w_j x_{t+j}, t=1,2,\dots,n. \quad (2.4)$$

Muitas vezes os pesos são escolhidos de forma que  $w_j = w_{-j}$ , para todo o  $j$ , e  $w_0 \geq w_1 \geq \dots \geq w_p \geq 0$ . Por exemplo, se a unidade de tempo for o mês  $X$  para uma série  $\{X_t; t \in Z\}$  com a componente sazonal de período 12 (anual), é comum considerar a seguinte média móvel de 12:

$$Y_t = \frac{0.5X_{t-6} + X_{t-5} + X_{t-4} + \dots + X_{t+4} + X_{t+5} + 0.5X_{t+6}}{12} \quad (2.5)$$

Por sucessivas aplicações de operadores de média móveis obtêm-se médias móveis de ordem superior.

Mediante o emprego de médias móveis de ordens apropriadas, podem ser eliminadas as variações cíclicas, estacionárias e irregulares, conservando-se, dessa forma, apenas o movimento de tendência.

Uma desvantagem desse método é que desaparecem os dados do começo e do fim da série. outra desvantagem é que as médias móveis podem gerar movimentos cíclicos, ou de outra natureza, que não existem nos dados originais. Uma terceira desvantagem é que as médias móveis são fortemente afectados pelos valores extremos. Para superar de certo modo essa desvantagem, usa-



se as vezes uma média móvel ponderada com pesos apropriados. Nesses casos, ao item central (ou itens) é atribuído o maior peso e aos valores extremos os menores.

### 2.3.1.1.2 Operador diferença (OD)

Quando se faz 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 intervalo  $t$  de tempo.

Um tipo especial de filtro, muito útil para ajustar a componente de tendência neste caso a sazonal, consiste em diferenciar a série até que ela se torne estacionária. Para dados não sazonais, a primeira diferença é em geral suficiente para induzir estacionariedade aproximada. Diferenciação de primeira ordem é a mais utilizada enquanto que pode-se usar a diferenciação de segunda ordem.

Os métodos que tem vindo a ser apresentados são métodos que servem para ajustar a tendência de uma série, sendo nesse caso o ruído da série a diferença entre a série original e a tendência ajustada. Contrariamente a esses métodos, o uso do operador diferença visa o ajustamento do ruído de uma série, com o consequente ajuste estacionário da série.

O operador diferença,  $D$ , é tal que:

$$DX_t = X_t - X_{t-1}, t \in Z \quad (2.6)$$

Diferença de mais alta ordem são obtidas por aplicações sucessivas do operador diferença, nomeadamente:  $D^k X_t = D(D^{k-1} X_t)$  para  $k \geq 2$ . Assim, por exemplo,

$$D^2 X_t = D(DX_t) = DX_t - DX_{t-1} = X_t - 2X_{t-1} + X_{t-2}$$

e, em geral:

$$D^k X_t = \sum_{j=0}^k \binom{k}{j} (-1)^j X_{t-j}, k \geq 0 \quad (2.7)$$



Seguem-se algumas observações relativas as propriedades do operador diferença e ao seu uso:

- P1. Se  $\{X_t\}$  é uma série cuja tendência é um polinómio de ordem  $k$ , então  $\{D^k X_t\}$  não possui a componente da tendência;
- P2. Construir  $\{D^k X_t\}$  é equivalente a remover da série  $\{X_t\}$  uma média móvel de ordem  $k+1$  da própria série;
- P3. Se  $\{X_t\}$  é um passeio aleatório<sup>2</sup>, então  $\{DX_t\}$  não possui uma componente da tendência.
- P4. Para remover a tendência de uma série  $\{X_t\}$  usando as suas diferenças, devem construir-se sucessivamente as séries das primeiras diferenças, segundas diferenças, etc. ( $\{DX_t\}, \{D^2 X_t\}, etc.$ ), até a inspeção visual indicar que a série das  $k$ -ésimas diferenças,  $\{D^k X_t\}$ , não possui tendência – geralmente  $k=1$  ou  $k=2$ .

A última observação está ligada ao facto de ao remover-se a componente da tendência de uma ST poder introduzir-se correlação na série do ruído.

Segundo Ender (1995), se a segunda diferença não ajustar a série numa estacionariedade deve-se calcular o logaritmo natural de cada dado observável ( $(\ln(x_t), t=1, \dots, n)$ ) e depois usar-se a diferenciação. Para este autor o número de *lags*<sup>3</sup> deve ser calculado usando a fórmula abaixo:

$$\frac{n}{4} \quad (2.8)$$

onde  $n$  corresponde ao número total de observações.

<sup>2</sup>  $\{X_t\}$  é passeio aleatório se  $X_t = X_{t-1} + \varepsilon_t$ , com  $E[\varepsilon_t] = 0$  e  $Cov(\varepsilon_t, \varepsilon_s) = 0$  para  $s \neq t$ .

<sup>3</sup> Corresponde a cada pauzinho representado no gráfico 3 ou 4, do capítulo IV.



### 2.3.1.1.3 Método a Sentimento

Este método consiste no ajustamento de uma recta ou curva de tendência, mediante a simples inspecção do gráfico, pode ser adoptado para a estimação de tendência. Entretanto, ele apresenta a desvantagem evidente de depender consideravelmente do critério individual.

### 2.3.1.1.4 Método das Semimédias

Consiste em separar os dados em duas partes ( de preferência iguais) e determinar a média de cada uma, obtendo-se, dessa forma, dois pontos do gráfico das STs. É desenhada, então, uma recta de tendência entre esses dois pontos e os valores da tendência podem ser determinados directamente, sem o emprego de um gráfico.

Embora este método seja de aplicação simples, ele pode conduzir a resultados medíocres, quando usado indiscriminadamente. Este método é aplicável quando a tendência é linear ou aproximadamente linear, embora possa ser estendido aos casos nos quais os dados são fragmentados em várias partes, em cada uma das quais a tendência é linear.

## 2.3.2 Sazonalidade

Um fenómeno sazonal é aquele que ocorre regularmente em períodos fixos de tempo e esta componente é difícil de ser estimada (Lattore e Cardoso, 2001).

Se houver uma sazonalidade dita determinista podem ser utilizados modelos de regressão que incorporem funções do tipo *seno* ou *co-seno* à variável tempo. Para se descobrir se existe sazonalidade na série de valores e verificar qual o ritmo, é importante realizar uma análise espectral (Lattore e Cardoso, 2001). Segundo eles, com este tipo de análise é possível identificar um padrão sazonal, mesmo dentro de uma variabilidade aleatória. A análise espectral utiliza um conjunto de funções que contêm seno e co-seno e tenta ajustá-las à variância observada em uma série de observações no tempo, levando em conta a amplitude das ondas, o período em que elas se repetem e a fase em que se iniciam. Para se retirar o efeito da sazonalidade de uma série, pode se usar a média móvel centrada no número de períodos que compõem uma repetição (por



exemplo, para sazonalidade anual, seria utilizada a média móvel de 12 meses) ou, então, pode-se trabalhar com a diferença entre a série original ( $X_t$ ) e o polinómio estimado para a sazonalidade.

### 2.3.3 Estacionariedade

Uma série estacionária é aquela que esta em **equilíbrio estatístico**<sup>4</sup>, enquanto que uma série não estacionária é aquela cujas propriedades mudam com o tempo (Jenkins, 1968; Wilks, 1995). Na prática, as séries são usualmente de 3 tipos:

- Aquelas que exibem *propriedades de estacionariedade em longo período*;
- As que possuem *uma razoável Estacionariedade em períodos curtos*;
- Finalmente, séries que são *obviamente não estacionárias*, no sentido que suas propriedades estão continuamente mudando com o tempo. A não-estacionariedade de uma série implica que:

- a) há inclinação da recta de tendência e ela não permanece ao redor de uma linha horizontal ao longo do tempo e/ou
- b) a variação dos dados não permanece essencialmente constantes sobre o tempo, isto é, as flutuações aumentam ou diminuem com o passar do tempo, indicando que a variância está se alterando.

Para detectar a não-estacionariedade de uma série, o comportamento temporal pode ser analisado graficamente, buscando padrões (a) e (b).

#### 2.3.3.1 Avaliação das Variações Estacionais. Índice de Estacionariedade

Para determinar o factor estacionário, deve-se estimar de que maneira os dados de uma ST variam de mês a mês, através de um ano típico. Um conjunto de números que mostre os valores relativos de uma variável durante os meses do ano é denominado *índice de estacionariedade* da variável. Se, por exemplo, se sabe que as infecções de HIV-SIDA, durante o meses de Janeiro, Fevereiro, Março, etc., foram 50, 120, 90, ...por cento da média mensal de todo o ano, os números 50, 120, 90, ...proporcionam um índice anual por estação e são, às vezes, designados por

<sup>4</sup> diz-se que uma ST possui um equilíbrio estatístico quando ela não contém tendência.



números índices de estacionariedade. O índice médio por estação (a média) correspondente ao ano todo seria de 100%, isto é, a soma dos números índices seria 1.200%.

Dispõe-se de vários métodos para o cálculo do índice de estacionariedade.

#### **2.3.3.1.1 O Método da Percentagem Média**

Neste método, os dados de cada mês são expressos em percentagens da média anual. As percentagens dos meses correspondentes dos diferentes anos são balanceadas mediante o emprego de uma média ou mediana. Se for adoptada a média, é melhor que se evitem os valores extremos que possam ocorrer.

As 12 percentagens resultantes dão os índices de estacionariedade. Se uma média não for de 100% (isto é, se a soma não for 1.200%), elas devem ser ajustadas, mediante sua multiplicação por factores convenientes.

#### **2.3.3.1.2 Método da Tendência ou Relação Percentual**

Neste método, os dados de cada mês são expressos em percentagens dos valores da tendência mensal. Uma média adequada das percentagens dos meses correspondentes dá, então, o índice desejado. Como no método anterior, elas são ajustadas quando a média não for de 100%.

Note-se que a divisão de cada valor mensal,  $Y$ , pelo correspondente da tendência  $T$ , produz o valor:

$$\frac{Y}{T} \quad (2.9)$$

O resultado desta divisão, corresponde ao índice de estacionariedade contendo movimentos cíclicos, estacionais e irregulares, sendo que a média subsequente destes valores produz os índices de estacionariedade que podem incluir variações cíclicas e irregulares, especialmente quando elas são grandes. Isto pode representar uma desvantagem importante deste método.



### 2.3.3.1.3 Método da Média Móvel Percentual ou da Relação entre as Médias

Neste método, calcula-se uma média móvel de 12 meses. Como os resultados assim obtidos caem entre meses sucessivos, em vez de no meio de um deles, como ocorre com os dados originais, calcula-se a média móvel de 2 meses em relação a média móvel de 12 meses já calculada. O resultado é frequentemente denominado *média móvel centrada de 12 meses*.

Depois disso, os dados originais de cada mês são expressos em percentagens da média móvel centrada de 12 meses correspondente. Calcula-se, então, a média das percentagens dos meses correspondentes, que dá o índice desejado. Como anteriormente, eles serão ajustados quando não apresentarem a média 100%.

Note-se que o raciocínio lógico em que se baseia este método, provém da equação  $Y = T \times C \times S \times I = TCSI$ .

Uma média móvel centrada de 12 meses dos valores de  $Y$  presta-se para eliminar os movimentos estacionais e irregulares,  $S$  e  $I$ , e é, portanto, equivalente aos valores dados por  $TC$ . Então, a divisão dos dados originais por  $TC$  produz os valores de  $IS$ . As médias subsequentes para os meses correspondentes prestam-se para eliminar a irregularidade  $I$  e, dessa forma, conduz a um índice  $S$  conveniente.

### 2.3.3.1.4 O Método dos Elos Relativos

Neste método, os dados de cada mês são expressos em percentagens nos dados do mês anterior. Essas percentagens são denominadas elos relativos porque encadeiam cada mês ao precedente. Toma-se, então, uma média adequada dos elos relativos, referentes aos meses correspondentes.

Desses 12 elos relativos médios podem ser obtidas as percentagens relativas de cada mês, em relação à Janeiro, que é considerada de 100%. Depois disso ter sido feito, verifica-se, usualmente, que o próximo Janeiro tem uma percentagem acumulada superior ou inferior a 100%, dependendo de ter havido, ou não, acréscimo ou decréscimo da tendência. Ao empregar este método, as várias percentagens obtidas podem, então, ser ajustadas para essa tendência. As percentagens finais, ajustadas de modo a apresentarem a média de 100%, proporcionam o índice estacionário desejado.



### 2.3.4 Destacionalização dos Dados

Se os dados mensais originais são divididos pelos índices estacionais correspondentes, diz-se que os dados resultantes estão *desestacionalizados* ou *ajustados à variação estacionária*. Esses dados incluem ainda movimentos de tendência, cíclicos e irregulares.

### 2.3.5 Estimação das Variações Cíclicas

Depois dos dados serem desestacionalizados, eles podem também ser ajustados à tendência, mediante sua simples divisão pelos valores de tendência correspondentes. De acordo com a equação  $Y = T \times C \times S \times I = TCSI$ , o processo de ajustamento à variação por estação e à tendência correspondem à divisão de  $Y$  para  $S T$ , o que dá  $C I$ , isto é, as variações cíclicas e irregulares. Uma média móvel apropriada, da duração de uns poucos meses (3, 5 ou 7 meses, por exemplo, de modo que a centralização subsequente não seja necessária) serve, então, para atenuar as variações irregulares  $I$  e deixar apenas as cíclicas. Uma vez isoladas, elas podem ser estudadas detalhadamente. Se ocorrer a periodicidade (ou a periodicidade aproximada) dos ciclos, podem ser idealizados *índices cíclicos*, de modo muito semelhante aos índices de estacionariedade.

### 2.3.6 Estimação das Variações Irregulares ou Aleatórias

A estimação das variações irregulares ou aleatórias pode ser realizada mediante o ajustamento dos dados à variações de tendência, estacionais e cíclicas. Isso importa em dividir os dados originais  $Y$  por  $T$ ,  $S$  e  $C$ , o que, de acordo com a equação  $Y = T \times C \times S \times I = TCSI$  produz  $I$ . Na prática verifica-se que os movimentos irregulares tendem a ser de pequena amplitude e que eles, frequentemente, tendem a seguir o padrão de uma distribuição normal, isto é, aquela na qual os pequenos desvios ocorrem com grande frequência e os grandes com pequena frequência.

### 2.3.7 Comparabilidade dos Dados

Quando se deseja comparar dados de diferentes anos, há que ter em conta as variações do número de dias de um mês em relação ao mesmo mês em anos diferentes.

Por exemplo, ao comparar os dados de Março com os de Fevereiro, deve-se recordar que Março tem 31 dias, enquanto Fevereiro tem 28 ou 29 dias. De modo semelhante, ao comparar os meses de Fevereiro de diferentes anos, deve-se lembrar que, nos anos bissextos, Fevereiro tem 29 em



vez de 28 dias. Os números de dias de trabalho de vários meses do mesmo ano, ou de anos diferentes, podem, também, ser diferentes por causa de férias, greves, licenças, etc.

Na prática, nenhuma regra definida é seguida para efectuar os ajustamentos devidos a essas variações. A necessidade desse ajustamento é deixada ao arbítrio do investigador.

#### 2.4 Análise do erro na Previsão de Séries Temporais

Quando são avaliados os diversos métodos para a previsão de STs, torna-se extremamente necessário o estabelecimento de procedimentos de avaliação de desempenho da metodologia empregada, visando determinar aquele que melhor possa vir a representar os valores obtidos no passado, e conseqüentemente melhor resultado poderá apresentar para valores que extrapolem o tempo presente (Martin, 2000).

Assim pode-se definir o erro no período  $t$ , como a diferença entre o valor actual  $X_t$  e o valor ajustado  $S_t$ ,

$$e_t = X_t - S_t \quad (2.10)$$

A soma dos erros em 2.11,

$$\sum_{t=1}^N e_t = \sum_{t=1}^N (X_t - S_t) \quad (2.11)$$

não é uma medida válida da efectividade do método de previsão, mas pode ser utilizada como medida de tendência do ajuste.

Normalmente são utilizados os valores indicados nas expressões 2.12, 2.13 e 2.14.

$$\text{Desvio absoluto médio : } DAM = \frac{\sum_{t=1}^N (X_t - S_t)}{N} \quad (2.12)$$

$$\text{Erro médio quadrático : } EMQ = \frac{\sum_{t=1}^N (X_t - S_t)^2}{N} \quad (2.13)$$



$$\text{Erro percentual absoluto médio : } EPAM = \frac{100}{N} \sum_{i=1}^N \left[ \left| \frac{X_i - S_i}{X_i} \right| \right] \quad (2.14)$$

O DAM é considerado como a média dos erros absolutos. Esse procedimento supera a característica de cancelamento dos erros positivos e negativos no erro médio<sup>5</sup> (EM). O erro é tratado como uma distância (entre a previsão  $S_i$  e o valor observado  $X_i$ ) 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 DAM pequeno.

EMQ 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 EMQ é utilizado para determinar o modelo de previsão mais preciso, isto é, considera-se o melhor modelo, aquele que possuir o menor EMQ.

O EPAM 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. De acordo com Lewis (1997), os valores do EPAM podem ser relacionados com a potencialidade das previsões segundo o quadro abaixo:

Tabela 1. Potencialidades de Previsão do EPAM

EPAM	POTENCIALIDADE DE PREVISÃO
Menor que 10%	Previsão é potencialmente muito boa
10% – 20%	Previsão é potencialmente boa
20% – 30%	Previsão é potencialmente razoável
Maior que 30%	Previsão é potencialmente inexata

Fonte: Lewis (1997)

<sup>5</sup> Erro médio (EM) 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.



## 2.5 Raiz Unitária e Teste ADF

A presença de tendência numa ST compromete a aplicação de algumas metodologias de previsão de séries temporais. Os modelos de RNs só têm suas propriedades asseguradas se os valores da variável em análise forem estacionários.

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

$$(1 - B)^d z_t = \varepsilon_t, \quad t = 1, 2, \dots, n \quad (2.15)$$

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

Os testes de raízes unitárias são capazes de detectar se a série foi suficientemente diferenciada para se tornar estacionária, sendo:

$H_0$ : 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 em um 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  $\varepsilon_t$  não é ruído branco, utiliza-se uma correcção. No caso, será utilizado o teste com a correcção paramétrica, conhecido como teste de Dickey-Fuller Aumentado (ADF) (Said e Dickey, 1984), que sugere aumentar a regressão adicionando termos suficientes em  $\Delta z_{t-1}$  para "branquear" os resíduos. Então o teste ADF é a estatística-t usada para testar o coeficiente de  $z_{t-1}$  em:

$$\Delta z_t = (\rho - 1)z_{t-1} + \sum \theta_i \Delta z_{t-i} + \varepsilon_t \quad (2.16)$$



que tem a mesma distribuição assintótica de Dickey-Fuller e não depende de  $\theta$ . No entanto, é necessário truncar a ordem da autorregressão por um método automático como o de AIC 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.

## 2.6 Autocorrelação

Uma ferramenta importante para se identificar as propriedades de uma ST é uma série de quantidades chamadas coeficientes de autocorrelação amostral (Ehlers, 2005). A ideia é similar ao coeficiente de correlação usual, i.e., para  $n$  pares de observações das variáveis  $x$  e  $y$ , o coeficiente de correlação amostral é dado por:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.17)$$

Aqui no entanto quer-se medir a correlação entre as observações de uma mesma variável em diferentes horizontes de tempo, isto é, correlações entre observações defasadas 1, 2, ... períodos de tempo. Assim, dadas  $n$  observações  $x_1, \dots, x_n$  de uma ST discreta podemos formar os pares  $(x_1; x_2), \dots, (x_{n-1}; x_n)$ . considerando  $x_1, \dots, x_{n-1}$  e  $x_2, \dots, x_n$  como duas variáveis o coeficiente de correlações entre  $x_t$  e  $x_{t+1}$  é dado por

$$r_1 = \frac{\sum_{t=1}^{n-1} (x_t - \bar{x}_1)(x_{t+1} - \bar{x}_2)}{\sqrt{\sum_{t=1}^{n-1} (x_t - \bar{x}_1)^2 \sum_{t=1}^{n-1} (x_{t+1} - \bar{x}_2)^2}} \quad (2.18)$$

Onde as médias amostrais são:

$$\bar{x}_1 = \sum_{t=1}^{n-1} \frac{x_t}{(n-1)} \quad \text{e} \quad \bar{x}_2 = \sum_{t=2}^n \frac{x_t}{(n-1)}$$



Como o coeficiente  $r_1$  mede as correlações entre observações sucessivas ele é chamado de coeficiente de autocorrelação ou coeficiente de correlação serial.

É usual simplificar a equação 2.18 utilizando-se a média de todas as observações, isto é,

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} \text{ já que } \bar{x}_1 \approx \bar{x}_2, \text{ e assumindo variância constante.}$$

Assim, a versão simplificada de (2.16) fica:

$$r_1 = \frac{\sum_{i=1}^{n-1} (x_i - \bar{x})(x_{i+1} - \bar{x})}{(n-1) \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n}} \quad (2.19)$$

Sendo que alguns autores ainda retiram o termo  $n/(n-1)$  que é próximo de 1 para  $n$  não muito pequeno. Esta última forma simplificada, sem o termo  $n/(n-1)$  será utilizada neste trabalho.

A equação 2.19 pode ser generalizada para calcular a correlação entre observações defasadas de  $k$  períodos de tempo, isto é,

$$r_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.20)$$

A equação 2.20 fornece o coeficiente de correlação de ordem  $k$ . Assim como o coeficiente de correlação usual, as autocorrelações são a dimensionais e  $-1 < r_k < 1$ .

Na prática é mais usual calcular primeiro o coeficiente de autocovariância  $\{c_k\}$ , definidos por analogia com a fórmula usual de covariância, isto é,

$$c_k = \sum_{i=1}^{n-k} \frac{(x_i - \bar{x})(x_{i+k} - \bar{x})}{n} \quad (2.21)$$



Os coeficientes de autocorrelação são então obtidos como  $r_k = \frac{c_k}{c_0}$ .

### 2.6.1 Correlação de Curto-prazo

É uma ST na qual uma observação está acima da média “tende” a ser seguida por uma ou mais observações acima da média, similarmente para observações abaixo da média. Um correlograma<sup>6</sup> desta série deverá exibir um valor relativamente grande de  $r_1$  seguido por valores que tendem a ficar sucessivamente menores. A partir de uma certa defasagem  $k$ , os valores de  $r_k$  tendem a ser aproximadamente zero.

### 2.6.2 Correlação negativa

Se os valores de uma série temporal tendem a se alternar acima e abaixo de um valor médio, o correlograma desta série também tende a se alternar. O valor de  $r_1$  será negativo enquanto o valor de  $r_2$  será positivo já que as observações desfasadas de 2 períodos tendem a estar do mesmo lado da média.

## 2.7 Generalidades de redes neuronais

As redes neuronais (RNs) têm sido utilizadas na modelação de memória associativa, reconhecimento de padrões, representação de funções booleanas, representação de funções contínuas, previsão de séries temporais, optimização etc.

Segundo Garcia (2002), as RN são estruturas de inspirações biológicas capazes de processar informações e a sua unidade fundamental é o neurónio formal. A rede é constituída por várias dessas unidades, dispostas em camadas sendo, a primeira camada constituída por neurónios sensores<sup>7</sup>.

Uma rede neural, segundo Haykin (2001), pode ser definida como um processador maciço e paralelamente distribuído, constituído de unidades de processamento simples, que têm a

---

<sup>6</sup> Um correlograma é um gráfico com os  $k$  primeiros coeficientes de autocorrelação como função de  $k$  (EHLERS, 2005). Esta pode ser uma ferramenta poderosa para identificar características da série temporal. Porém isto requer uma interpretação adequada do correlograma, i.é., devem ser associados certos padrões do correlograma como determinadas características de uma série temporal e esta nem sempre é uma tarefa simples

<sup>7</sup> Neurónios sensores são aqueles que se limitam a passar o estímulo a outros neurónios.



propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela assemelha-se ao cérebro em dois aspectos: (1) o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem; (2) forças de conexão entre neurónios (os pesos sinápticos) são utilizadas para armazenar o conhecimento adquirido.

As redes neuronais são formadas por neurónios e conexões entre eles. O neurónio (Figura 1) representa uma região onde informações são processadas. Seus três elementos básicos são: os pesos sinápticos, a função de soma e a função de transferência. O conjunto das componentes individuais de estímulo que atingem a camada de entrada da rede constitui o vector de estímulo da rede enquanto que os neurónios das outras camadas que efectuam algum tipo de processamento sobre os estímulos que recebem são chamados neurónios processadores (Garcia, 2002).

Cada neurónio pode estar ligado a outros neurónios mediante a utilização de ligações chamadas *ligações sinápticas (LS)* (ou sinapse) onde cada ligação (sinapse) permite a activação do neurónio seguinte. Esta sinapse é caracterizada por um peso que se chama *peso sináptico (PS)* e a comunicação entre os neurónios é realizada unidireccionalmente.

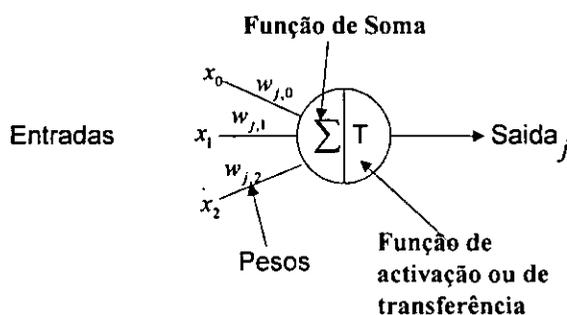
O primeiro modelo lógico matemático de um neurónio foi desenvolvido por McCulloch and Pitts em 1943 (McCulloch, 1943) tendo sido designado por TLU (Threshold Logic Unit) e este consiste numa unidade computacional que opera com sinais binários (0/1) de acordo com as seguintes regras:

1. Recebe as entradas  $x_1, x_2, x_3, \dots, x_n$  em  $n$  sinapses de excitação e as entradas  $y_1, y_2, y_3, \dots, y_m$  em  $m$  sinapses inibidoras;
2. Se pelo menos uma das entradas inibidoras,  $y_1, y_2, y_3, \dots, y_m$  é 1 então a unidade computacional é inibida apresentando como saída o valor 0;
3. Se todas as entradas inibidoras,  $y_1, y_2, y_3, \dots, y_m$  são 0 então a unidade computacional apresenta o sinal de saída  $f(x_1 + x_2 + \dots + x_n - b)$  em que  $f$  é a função degrau unitário<sup>8</sup> e a constante  $b$  representa um desvio (bias).

<sup>8</sup> função degrau unitária é o tipo mais simples de função de activação. Sua resposta pode



Figura 1. Modelo genérico de um neurónio artificial



Fonte: Miguel Ângelo Moreira (1997)

Os PS são responsáveis pelo armazenamento das informações. Além disso, elas definem o efeito que a saída de um neurónio exerce sobre a entrada do neurónio seguinte. Os PS são de grande importância para uma RN, pois determinam toda a manipulação de valores da rede (Alves, 2001). A função de soma processa os estímulos ponderados pelos respectivos pesos, ou seja:

$$x_j = \sum_i w_{i,j} y_i \quad (2.22)$$

onde  $y_i$  é a saída gerada por cada neurónio da camada anterior.

Já a função de transferência, também chamada de função de activação, limita a amplitude do intervalo do sinal de saída do neurónio para algum valor finito, geralmente no intervalo normalizado  $[0,1]$  ou  $[-1,1]$ .

$$y_j = f(x_j) \quad (2.23)$$

---

assumir dois valores: 0 ou 1, como demonstrado a seguir:  $f(x) = \begin{cases} 1, se & x \geq 0 \\ 0, se & x < 0 \end{cases}$



## 2.7.1 TIPOS DE REDES NEURONAIS

O exemplo mais antigo de redes neuronais é a rede *perceptron* (RP) com uma camada de nós de saída, conectados às entradas por conjuntos de PS. Esta topologia pode ser considerada a forma mais simples de rede em avanço.

### 2.7.1.1 Redes Perceptron (RP) com uma Camada

RP com uma camada são o tipo mais antigo de redes neuronais, as quais são formadas por uma camada única de NS, os quais estão conectados por pesos às entradas. Este é o exemplo mais simples de **redes em avanço**. A soma do produto entre pesos de entrada alimentam cada neurónio de saída, e se o resultado desta operação exceder um certo limite, o neurónio coloca o valor 1 como sendo o valor de saída; se o resultado for inferior ao limite, o neurónio coloca o valor -1. Como foi dito acima, neurónios com esse comportamento são chamados de *neurónios de McCulloch-Pitts* ou podem ser chamados de RN com pesos sinápticos binários como será visto mais adiante. Na literatura técnica o termo *perceptron* diz respeito a redes com apenas um desses neurónios.

As RP podem ser treinadas por um algoritmo de aprendizagem simples, chamado geralmente de regra-delta (RD). Este algoritmo calcula os erros entre a saída dos dados calculados e a saída desejada, e utiliza isso para ajustar os pesos.

### 2.7.1.2 Redes Perceptron Multi-camadas (RPMC)

Esta classe de rede consiste em múltiplas camadas com unidades computacionais, geralmente inter-conectadas em uma forma **feedforward**<sup>9</sup>. Em muitas aplicações as unidades dessas redes aplicam uma função sigmoid (em forma de S) como a função de activação<sup>10</sup>.

RPMC podem usar um grande número de técnicas de aprendizagem, sendo que a mais popular é a propagação reversa. Neste caso os valores de saída são comparados com a resposta correcta para comparar com o valor predefinido através de uma função de erro (FER) predefinida. Por

<sup>9</sup> Também conhecidas por **retropropagação** e isso quer dizer que cada neurónio em uma camada tem sinapses directas a neurónios da próxima camada e em algumas literaturas as redes que apresentam sinapses deste tipo são chamadas de RN feedforward (RNF) ou RN de Retropropagação.

<sup>10</sup> Função de activação encontra-se definida em 2.7.5 e em 2.7 deste capítulo.



alguma técnica o erro é então alimentado de volta na rede. Usando essa informação, o algoritmo ajusta os pesos de cada conexão para reduzir o valor do FEr, após repetir este processo por um número suficiente de vezes.

## **2.7.2 Caracterização das redes neuronais**

### **2.7.2.1 Classificação quanto à arquitectura**

#### **2.7.2.1.1 Redes Neurais Planares (RNPs)**

As RNPs são constituídas por neurónios intermediários (NI) e por neurónios periféricos (NP), todos eles dispostos numa superfície bidimensional (Moreira, 1997). Estes últimos podem subdividir-se em neurónios de entrada (NE) e de saída (NS).

A disposição dos diferentes neurónios pode ou não estar organizada por camadas. Sempre que a resposta dos neurónios de uma rede é binária (0/1 ou -1/+1, por exemplo) esta pode designar-se binária e quando esta rede é constituída por neurónios binários e pesos sinápticos binários são conhecidas por redes McCulloch-Pitts.

#### **2.7.2.1.2 Redes Neurais por Camadas (RNC)**

As RNs podem ser constituídas por diferentes camadas dispostas paralelamente. A primeira é designada por camada de entrada (CE) e a última por camada de saída (CS). As camadas intermédias (CI) são designadas por camadas escondidas. As redes constituídas por camadas constituem casos particulares das RNP.

Na contagem do número de camadas de uma rede, alguns autores não entram em consideração com a camada de entrada. No entanto, esta camada é considerada sempre na operação da sua enumeração. Assim, uma rede com duas camadas será constituída por uma CE e por uma CS.

Na numeração das camadas atribui-se o número 0 à camada de entrada e os números crescentes consecutivos às camadas seguintes.



Quando a informação de cada camada não constitui informação de entrada de nenhuma camada anterior a rede designa-se por redes neuronais de alimentação directa (RNAD<sup>11</sup>; figura 2) e estas não possuem realimentação. Caso contrário designa-se por redes recorrentes (RNR<sup>12</sup>; figura 3).

As RNAD em que a informação de saída é constituída por padrões binários são designadas por alguns autores por redes de *perceptrões* (RP). Mais rigorosamente um perceptrão é considerado uma unidade que calcula o sinal dum combinação linear das variáveis de entrada.

Em algumas situações certas camadas possuem neurónios cujas funções de transferência são lineares. Neste caso estas camadas designam-se lineares. Uma rede apenas com camadas lineares designa-se rede linear (RL). As camadas não lineares designam-se normalmente pelo nome da função de transferência comum aos diferentes neurónios utilizados nessas camadas. As redes com uma ou mais camadas não lineares podem designar-se redes não lineares.

Redes com camadas sigmoidais ou Gaussianas constituem exemplos de redes não lineares. Redes com camadas deste último tipo podem designar-se por redes de base radial.

#### **2.7.2.2 Classificação quanto ao processo de aprendizagem**

Segundo (\*), a propriedade mais importante das RNs é a habilidade de aprender de seu ambiente e com isso melhorar seu desempenho. Isso é feito através de um processo iterativo de ajustes aplicado a seus pesos, o treinamento. A aprendizagem ocorre quando a RN atinge uma solução generalizada para uma classe de problemas.

Denomina-se algoritmo de aprendizagem a um conjunto de regras bem definidas para a solução de um problema de aprendizagem (\*). Existem muitos tipos de algoritmos de aprendizagem específicos para determinados modelos de RNs, estes algoritmos diferem entre si principalmente pelo modo como os pesos são modificados.

---

<sup>11</sup> O outro nome atribuído as RNAD é feedforward e são as redes mais vulgarizadas e mais utilizadas.

<sup>12</sup> Estas redes apresentam sinapses entre quaisquer neurónios

\* <http://www.icmc.usp.br/~andre/research/neural/desenv.htm>



Outro factor importante é a maneira pela qual uma RN se relaciona com o ambiente. Nesse contexto existem os seguintes paradigmas de aprendizagem:

- **Aprendizagem Supervisionado**, quando é utilizado um agente externo que indica à rede a resposta desejada para o padrão de entrada. Este tipo de aprendizagem, quando um padrão é inicialmente apresentado à rede, ela produz uma saída. Após medir a distância entre a resposta actual e a desejada, são realizados os ajustes apropriados nos pesos das conexões de modo a reduzir esta distância. Este procedimento é conhecido como **Regra Delta** (RD). O esquema de treinamento é o seguinte:

- Iniciar todas as conexões com pesos aleatórios;
- Repita até que o erro  $E$  seja satisfatoriamente pequeno ( $E = e$ )
- Para cada par de treinamento  $(X, d)$ , faça:
- Calcular a resposta obtida  $O$ ;
- Se o erro não for satisfatoriamente pequeno  $E > e$ , então:
- Atualizar Pesos:  $\text{PesoNovo} := \text{Peso anterior} + \text{Neta } E X$

Onde:

- ✓ O par de treinamento  $(X, d)$  corresponde ao padrão de entrada e a sua respectiva resposta desejada;
  - ✓ O erro  $E$  é definido como:  $\text{Resposta Desejada} - \text{Resposta Obtida} (d - O)$ ;
  - ✓ A taxa de aprendizagem  $\text{Neta}$  é uma constante positiva, que corresponde à velocidade de aprendizagem.
- **Aprendizagem Não Supervisionado** (auto-organização), quando não existe uma agente externo indicando a resposta desejada para os padrões de entrada;
  - **Reforço**, quando um crítico externo avalia a resposta fornecida pela rede.

Denomina-se ciclo uma apresentação de todos os  $N$  pares (entrada e saída) do conjunto de treinamento no processo de aprendizagem. A correcção dos pesos num ciclo pode ser executado de dois modos:



1) **Modo Padrão (MP)**: A correcção dos pesos acontece a cada apresentação à rede de um exemplo do conjunto de treinamento. Cada correcção de pesos baseia-se somente no erro do exemplo apresentado naquela iteração. Assim, em cada ciclo ocorrem N correcções.

2) **Modo Batch (MB)**: Apenas uma correcção é feita por ciclo. Todos os exemplos do conjunto de treinamento são apresentados à rede, seu erro médio é calculado a partir deste erro e fazem-se as correcções dos pesos.

### 2.7.3 Correcção de erros

A correcção de erros neste trabalho é feita por meio do algoritmo backpropagation. Segundo Jacinto et al (2004), este algoritmo adapta as correcções dos erros procurando minimizar a soma dos erros quadráticos das saídas. ( Equação 2.24)

$$E(w) = \frac{1}{2} \sum_{i=1}^k (d_i - y_i(w))^2 \quad (2.24)$$

onde:

$k$  é o número de neurónios de saída da rede;

$d_i$  é a saída desejada para o neurónio  $i$ ;

$y_i$  é a saída corrente da rede.

### 2.7.4 Topologia das RNs

A definição da topologia da rede é o parâmetro que restringe o tipo de problema a ser tratado (Jacinto et al, 2004).

Fazem parte da definição da topologia das RNs, os seguintes parâmetros:

- a) Número de camadas;
- b) Número de nós em cada camada;
- c) Tipo de conexão entre os nós da rede.

Uma rede com uma CI pode implementar qualquer função contínua (Jacinto et al, 2004 ). Segundo Jacinto et al (2004) citado em Haykin (2001) a utilização de duas CI permite a



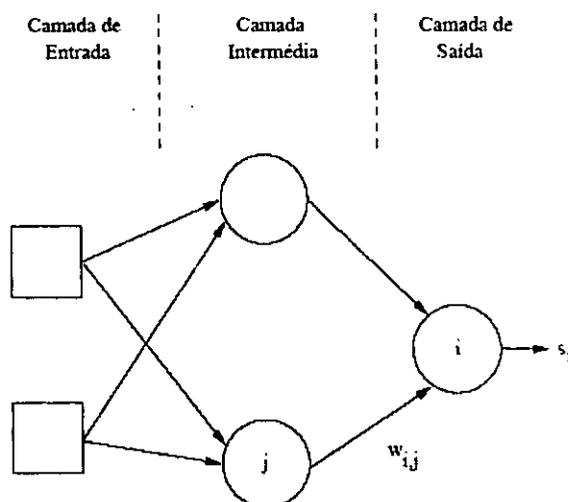
aproximação de qualquer função, mas dependendo da distribuição dos dados, a rede pode convergir para um mínimo local ou pode demorar para encontrar a saída desejada.

Neste trabalho, escolher-se-á RN do tipo Perceptron multicamadas (RPMC), que apresentam uma CI. As redes de múltiplas camadas apresentam um poder maior do que as redes sem camadas intermediárias, pois podem implementar qualquer função, seja ela linearmente separável ou não. A precisão obtida e a implementação da função objectiva depende do número de neurónios utilizados na CI.

### 2.7.5 Activação e Função de Activação

Seja dada um exemplo de uma RN feedforward, teremos:

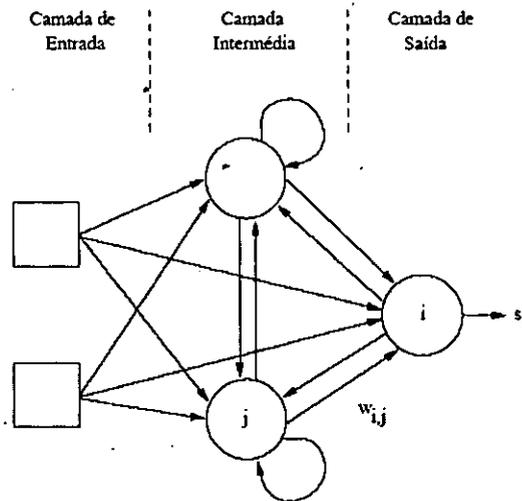
Figura 2. Estrutura de uma RNAD ou RN Feedforward



Fonte: Paulo Alexandre Ribeiro Cortez (Braga, 1997)



Figura 3. Estrutura de uma RNR



Fonte: Paulo Alexandre Ribeiro Cortez (Braga, 1997)

Cada neurónio calcula a sua activação<sup>13</sup> a partir dos valores das saídas dos neurónios que lhe são ligados e dos pesos (PS) associados as respectivas sinapses. Representando a energia de activação do neurónio  $i$  por  $s_i$  e do neurónio  $j$  por  $s_j$ , como se pode ver na figura 4, sendo:

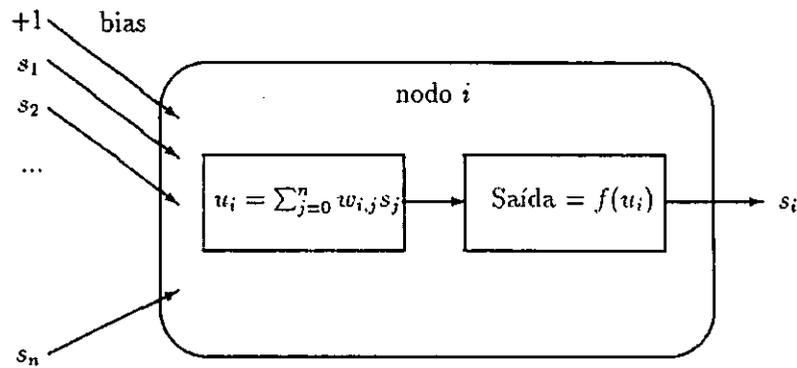
$$u_i = \sum_{j=0}^n w_{i,j} \times s_j \quad (2.25)$$
$$s_i = f(u_i)$$

onde  $w_{i,j}$  representa o peso associado a ligação entre o neurónio  $j$  e o neurónio  $i$ . A função  $f$  é chamada de função de activação. Os valores da energia de activação podem ser discretos (tomando normalmente os valores 1; 0 ou -1) ou contínuos. No caso dos valores da energia de activação serem contínuos,  $f$  é normalmente uma função não linear limitada e diferenciável por troços.

<sup>13</sup> Que também pode ser chamado de energia de activação.



Figura 4. Activação de um neurónio



Fonte: Paulo Alexandre Ribeiro Cortez (Braga, 1997)

As principais funções de activação encontram-se na tabela 2.

Tabela 2. Funções de activação

NOME	FUNÇÃO F(X)	CONTRADOMÍNIO
Linear	$x$	$]-\infty, +\infty[$
Sigmoid	$\frac{1}{1 + e^{-x}}$	$[0,1]$
Sigmoid1	$\frac{2}{1 + e^{-x}} - 1$	$[-1,1]$
Sigmoid2	$\frac{x}{1 +  x }$	$[-1,1]$
Tanh	$\tanh(x)$	$[-1,1]$
Sin	$\text{Sin}(x \text{ mod } 2\pi)$	$[-1,1]$
Cos	$\text{Cos}(x \text{ mod } 2\pi)$	$[-1,1]$
Gaussian	$e\left(\frac{-x^2}{2}\right)$	$[-1,1]$

Fonte: Paulo Alexandre Ribeiro (Braga, 1997)

### 2.7.6 Implementação da Rede Neuronal

A implementação da RN começa pela identificação e colecta dos dados, dados estes que devem ser históricos. O passo seguinte é a preparação e adequação dos dados ao formato requerido pela rede, que será seguida pela separação dos dados em dois conjuntos: o conjunto de treinamento o



qual é gerado a partir de dados históricos e o conjunto de teste com o qual o funcionamento da rede é testado. O passo seguinte é realizar a escolha do modelo neural a ser adoptado e definir a topologia da rede. Em seguida realiza-se o desenvolvimento, treinamento e optimização do modelo, seguidos pela validação do mesmo. Na etapa de validação, faz-se uma comparação do resultado obtido pela rede com o resultado desejado.

### **2.7.7 Aplicação das RNs**

As redes neuronais podem ser utilizadas em:

- **Classificação** – É aquela que permite escolher entre duas classes que o elemento pode pertencer a classe A, a classe B ou a nenhuma,
- **Auto-associação** – Isto é, quando uma RN tem exactamente tantos neurónios de entrada quanto aos de saída, ela é treinada para reproduzir na saída dados apresentados na entrada,
- **Previsão de STs** – Servem para prever resultados futuros de uma ST;
- **Aproximação de funções** – Consiste na aproximação de funções que têm como domínio e contra domínio o conjunto dos vectores de valores reais.

De acordo com o objectivo deste trabalho falar-se-á das RNs que são usadas para a previsão de STs, como se pode ver a seguir.

#### **2.7.7.1 RNs Aplicadas para Previsões em STs**

A Previsão de valores futuros de uma ST é um problema extremamente comum. Qualquer RN que seja capaz de aceitar vectores com valores reais como entrada e produzir valores reais como saída pode ser usado para esse tipo de Previsão (Martin, 2000).

Segundo Moreira (1997), na previsão de séries temporais associadas a fenómenos não lineares têm sido utilizadas redes não lineares e que possuem camadas com alimentação directa.

A figura abaixo mostra o uso de  $n$  pontos numa série temporal para prever o valor associado ao ponto  $(n+1)$ . Para tal, emprega-se uma série de treinamento de  $m$  pontos, onde  $m > n$  e, a partir

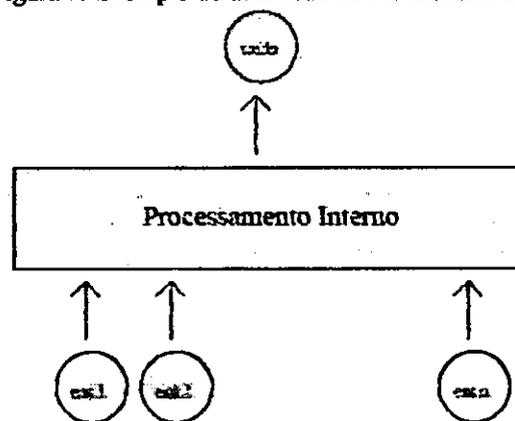


dessa, geram-se um grande número de amostras de  $(n + 1)$  elementos. Cada amostra para o modelo consiste de um valor actual e  $(n-1)$  valores históricos.

De acordo com Martin (2000), a rede é ajustada para que o valor de saída coincida com o valor  $(n+1)$ . Um aspecto vital no processamento de STs consiste na necessidade de se eliminar tendências e variações sazonais.

Eliminando tendências que são facilmente previsíveis e podem ser adicionadas posteriormente, liberta-se a rede para um processamento em maior profundidade. A sazonalidade pode provocar uma concentração em aspectos lineares quando o que realmente se pretende é explorar a característica não linear das RNs.

Figura 5. Exemplo de uma rede de Previsão em STs



Fonte: Claudio Martin (Florianópolis, 2000)



# CAPÍTULO III

## Material e Métodos

---

### 3.1 Material

O material usado para esta pesquisa foi:

- A ST extraída do BS do HCM;
- O Software estatístico *Minitab for Windows*;
- O Software *NeuralWare*<sup>14</sup> de gestão das RNs;
- Software *DTREG*<sup>15</sup> - Programa de modelação e previsão;
- O pacote [WEB:REG] UNIT ROOT TEST (ADF-TEST)<sup>16</sup>;
- O Microsoft *Excel 2003*;
- Microsoft *Word 2003*;
- Um processador Pentium IV, 2.0 de velocidade de processamento, 512 de memória RAM e 80 de disco duro.
- Impressora deskjet 610C;

### 3.2 Métodos

A ST usada para este trabalho foi extraída da Base de dados do Banco de Sangue (BS) do Hospital Central de Maputo (HCM) e esta corresponde ao sangue infectado pelo vírus de HIV-SIDA.

A modelação desta ST graficamente permitiu a identificação dos principais movimentos desta, enquanto que, para a verificação da não estacionaridade bem como a estacionaridade da série, foi por meio da representação gráfica da função de autocorrelação parcial. Para o caso em que a série não era estacionária, o OD de ordem  $n$  foi usado até que este tornasse a série estacionária. O teste ADF também foi usado para a confirmação da estacionariedade da série.

---

<sup>14</sup> Encontrado em <http://ftp://ftp.neuralware.com/PredictDemo.exe>

<sup>15</sup> Encontrado em <http://www.dtreg.com>

<sup>16</sup> Encontra-se em <http://www.web-reg.de/>



Uma entrevista semi-estruturada (Anexo H) feita aos agentes responsáveis do BS do HCH mediante a apresentação de uma credencial cedida pelo Departamento de Matemática e Informática (DMI) que comprovou ser estudante e desejar a ST meramente para a elaboração do trabalho de licenciatura, permitiu identificar e descrever as técnicas de previsão usadas no BS do HCM.

Para uma análise mais sustentável da eficiência das RNs quando usadas na previsão de STs foi feita com base na análise dos erros de previsão. Neste caso, o EPAM foi usado para a avaliação deste modelo segundo os padrões estabelecidos por Lewis (2001).

A revisão da literatura permitiu identificar a topologia da rede bem como o número de CI a usar. Quanto ao número de neurónios a usar na CI, estes foram otimizados pelo s/w *DTREG* (anexo E) por meio da introdução de dois (2) a trinta (30) neurónios para esta camada (anexo F). A selecção das suas funções de activação, isto é, na CI e na CS também foram seleccionados por este s/w.

A apresentação dos resultados provenientes da modelação da rede bem como a sua aplicação, foram feitas por meio do software *NeuralWare*, que também permitiu a modelação da RN bem como a aplicação da ST em estudo.



# CAPÍTULO IV

## Resultados e Discussão

---

---

### 4.1 Resultados

#### 4.1.1 Modo de análise dos Dados usadas no Hospital Central de Maputo (HCM)

Os dados fornecidos pelo BS do HCM correspondentes as unidades de sangue infectadas pelo vírus de HIV-SIDA, obtidas pela Base de Dados usada pela instituição são de duas naturezas, nomeadamente:

- Doadores repositores: correspondente as pessoas que doam sangue por possuírem familiares que estejam necessitando deste, para o seu acompanhamento médico.
- Doadores voluntários: como o nome diz, estes são os que doam por vontade própria e algumas vezes, esses são motivados pelas palestras feitas pelos agentes do BS.

Após a doação do sangue este é filtrado de modo que se encontre a parte útil<sup>17</sup> e a parte inútil<sup>18</sup>. De seguida, separa-se o útil do inútil, sendo o sangue útil correspondente aos doadores voluntários usado para o stock do BS para questões de emergência.

As técnicas de previsão do número de doadores, bem como a parte do sangue inutilizado por diversos motivos como é o caso do vírus de HIV e outras doenças encontradas no anexo "I", baseiam-se em opiniões dos agentes responsáveis desta unidade hospitalar. Estes analisam situações similares, e em conjunto com os dados existentes, procuram prever os seus valores futuros.

---

<sup>17</sup> Parte útil corresponde ao sangue sem impurezas e que pode ser doado as pessoas necessitadas.

<sup>18</sup> Parte inútil corresponde ao sangue que possui impurezas, isto é, infectado pelo vírus do HIV ou mesmo possui uma outra anomalia como por exemplo, sangue com bolhas parecendo coágulos, sangue proveniente de um dador que possui meningite, etc.

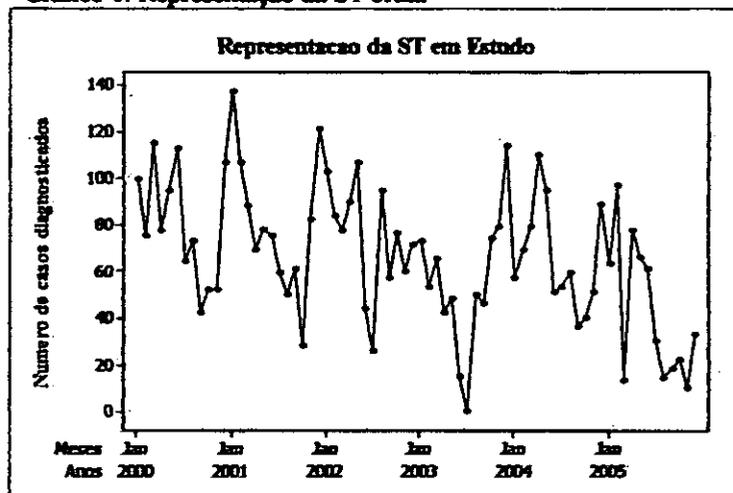


## 4.1.2 Modelação da Série Temporal Recolhida

### 4.1.2.1 Análise da Série Temporal Recolhida

A ST de sangue infectado pelo vírus de HIV recolhida no BS do HCM, graficamente representa a seguinte estrutura:

Gráfico 1. Representação da ST bruta



Fonte: Adaptação do Autor

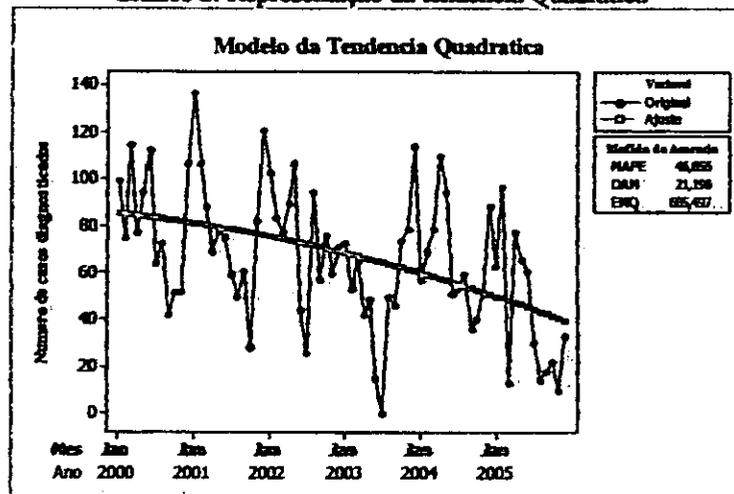
Olhando pela representação do gráfico, é de notar que fica pouco difícil confirmar a existência das principais componentes duma ST, como é o caso da tendência, sazonalidade bem como a estacionariedade. Para a sua identificação, uma análise cuidadosa destes componentes será necessário.

#### 4.1.2.1.1 Análise da Tendência

A Tendência dá-nos uma medida do crescimento ou decréscimo dos valores de uma série. Uma série que possui este tipo de comportamento é chamada não estacionária. No gráfico 2 está representada a tendência dos dados de HIV-Sida, correspondentes a ST em estudo.



Gráfico 2. Representação da tendência Quadrática



Fonte: Adaptação do Autor

A partir do gráfico anterior (gráfico 2) é de notar que esta série possui uma tendência quadrática (TQ) que foi identificada a partir de uma comparação entre a tendência linear (TL), quadrática e exponencial (TE), sendo o critério de comparação a escolha da tendência que menor valor do EMQ possui. Olhando para este erro na tabela 3 é possível notar que o EMQ para a tendência quadrática é o mais adequado para esta série, isto é, por ser o menor comparativamente as outras tendências (TL e TE)..

Tabela 3. Valores dos erros das Tendências

Tipo da Tendência	DAM	EMQ	EPAM
TL	21.366	689.140	48.051
TQ	21.196	685.497	46.856
TE	22.3.5	768.131	111.613

Fonte: Adaptação do Autor

Nota: Para a identificação da TE foi necessário alterar o valor da observação para o mês de Julho de 2003, do valor zero (0) para um (1). Esta alteração foi necessária porque para a modelação da TE os dados observados não devem possuir valores iguais a zero (0).

#### 4.1.2.1.2 Análise da Sazonalidade

Analisando o factor sazonal ou cíclico, nota-se que a série em estudo possui sazonalidade visto que, pela modelação gráfica dos dados, dados estes encontrados no anexo A, pode-se notar que estes possuem sazonalidade como se pode notar no gráfico 1. Esta sazonalidade é notória pelo comportamento constante no gráfico durante os mesmos períodos de tempo, nomeadamente nos finais de cada ano.

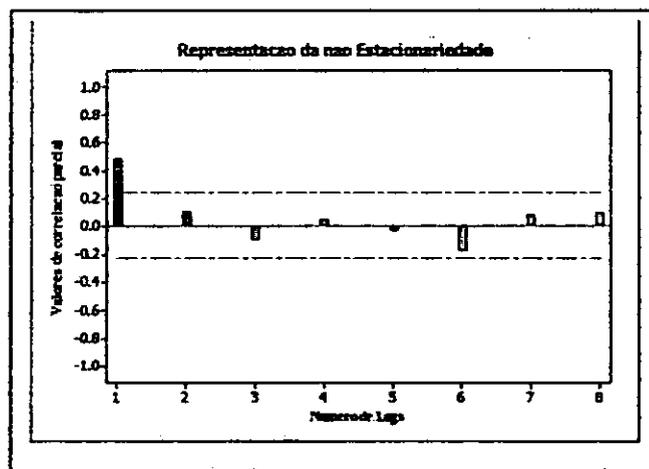


#### 4.1.2.1.3 Ajustamento da ST à Estacionária

Analisando-se a tendência da série notou-se que esta possui uma TQ, e quando uma série possui tendência diz-se que ela não é estacionária<sup>19</sup> segundo o primeiro parágrafo em 4.1.2.1.1 deste capítulo.

Procurando tornar a série em estacionária, foi determinada a primeira diferença<sup>20</sup> em relação aos dados brutos por meio do uso do operador diferença ou seja, operador diferença da primeira ordem o que permitiu a transformação desta série que não era estacionária em série estacionária. O gráfico 3 e gráfico 4 ilustram esta situação.

Gráfico 3 Serie não estacionária (Modelada à partir da ST real)



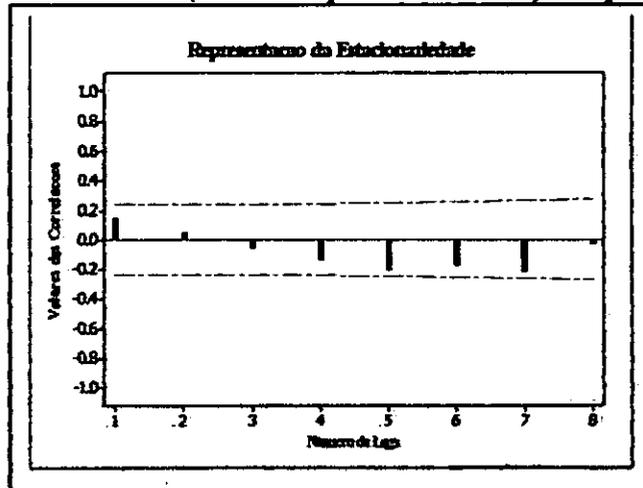
Fonte: Adaptação do Autor

<sup>19</sup> O gráfico 3 vem a reforçar esta situação, visto que existem lags (definido em 2.3.1.1.2) que não se encontram dentro dos limites de autocorrelação.

<sup>20</sup> A tabela dos dados correspondentes a primeira diferença encontram-se no anexo B.



Gráfico 4 Série estacionária (Modelada à partir da diferenciação da primeira ordem )



Fonte: Adaptação do Autor

O número de lags presentes na figura acima corresponde a 8, foi calculado a partir da fórmula 2.8 do capítulo II, enquanto que os valores [0.2; -0.2] são os limites correspondentes aos valores de autocorrelação, e estes são válido de acordo com o ponto 2.6 da revisão da literatura.

verificando a variação dos lags do gráfico 3 é de notar que nem todos lags se encontram dentro dos coeficiente de autocorrelação que corresponde a 0.2 em módulo, daí a razão de dizer que esta série é não estacionária. Olhando para o gráfico 4, a situação é diferente, visto que é de notar que os lags encontram-se dentro dos limites determinado pelo coeficiente de autocorrelação, o que faz com que esta série seja estacionária. Para o ajuste desta série foi necessário usar um limite de significância de 5%.

Como não sendo suficiente, o uso do teste ADF de raiz unitária vem a reforçar o argumento, sendo que a tabela abaixo mostra o resultado do teste.

Tabela 4. Teste ADF para confirmação da Estacionariedade

Augmented Dickey-Fuller (ADF) test statistic	Valor Crítico do teste:  <b>-4.916249</b>	1% level	-4.118488
		5% level	-3.486516
		10% level	-3.171563

Fonte: Adaptação do Autor

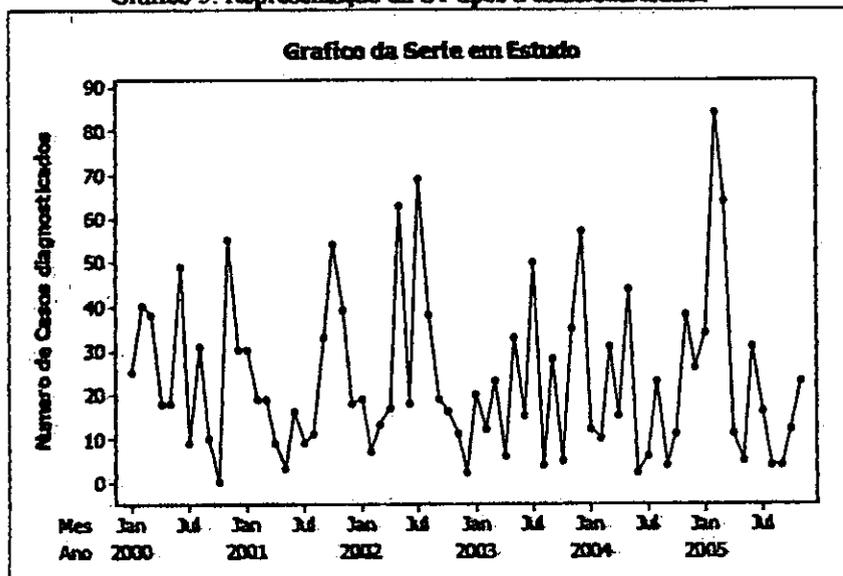
A tabela 4 apresenta o valor crítico do teste em módulo. Este valor é maior em relação aos outros valores também em módulo, isto para os níveis de significância igual a 1%, 5% bem como a 10%, o que leva de facto a confirmar que a série é realmente estacionária. É de referir que a série



usada para a análise do teste ADF de raiz unitária corresponde a série anteriormente considerada estacionária (correspondente ao gráfico 4).

Representando os dados após o ajustamento teremos:

Gráfico 5. Representação da ST após a estacionariedade



Fonte: Adaptação do Autor

Pela leitura do gráfico 5, o factor estacionário pode não ser claro, daí que houve a necessidade da confirmação da estacionariedade desta série, confirmação esta feita pelo teste de raiz unitária de Dickey-Fuller ( teste ADF) e cujo resultado encontra-se na tabela 4.

#### 4.1.3 Modelação via RNs

Para a modelação da série via RNs foram usadas as STs que originaram o gráfico 3 e gráfico 4. O sustento do uso deve-se a utilização da aprendizagem supervisionada de acordo com o ponto 2.7.2.2 da revisão da literatura, sendo a ST do gráfico 3 a que representa o padrão de entrada e a ST do gráfico 4 a que representa a resposta desejada. O outro factor da utilização da ST correspondente ao gráfico 4 é pelo facto de ao eliminar as tendências que são facilmente previsíveis e que podem ser adicionadas posteriormente, liberta-se a rede para um processamento com maior profundidade. A sazonalidade pode provocar uma concentração em aspectos lineares quando o que realmente se pretende é explorar a característica não-linear das RNs.



#### 4.1.3.1 Parâmetros Usados para Modelação

Uma das decisões mais importantes na construção de uma RN é a escolha das variáveis nas quais a rede irá se basear para a realização do seu treinamento.

Tabela 5. Parâmetros da RN

Parâmetro	Especificação
Número de variáveis de previsão	1
Modelo Usado	Múltiplas camadas com alimentação directa
Número de camadas	3
Número de neurónios na camada escondida	10
Função de activação	Logística
Aprendizagem	Supervisionado
Categoria dos pesos	Distribuição normal
Número de Dados	72
Dados sem o valor de previsão	1

Fonte: Adaptação do Autor

A determinação das variáveis ou parâmetros para a modelação dos dados em estudo, parâmetros estes encontrados na tabela 5, foram obtidas por meio do software *DTREG* que é um dos softwares de modelação e previsão usados em RNs.

#### 4.1.3.2 Arquitectura da Rede Utilizada para a Modelação

A arquitectura da rede usada para a modelação foi gerada pelo software *DTREG*. A natureza da introdução dos parâmetros ou seja, dados para a optimização da arquitectura podem ser vistos no anexo E. A arquitectura usada possui uma alimentação directa entre os neurónios da camada de entrada, escondida e saída. Este tipo de arquitectura pode ser designada por RNAD ou redes de retropropagação como se pode ver em 2.7.2.1.2 da revisão da literatura, o que quer dizer que os dados de saída dos neurónios da camada de entrada funcionam como dados de entrada para os neurónios da camada escondida e por sua vez, a saídas deste são as entradas da camada de saída. As saídas obtidas na camada de saída correspondem os dados de previsão obtidos pela rede.

Todas as redes neuronais possuem uma CE e CS, enquanto que as CI podem ser várias. Para esta arquitectura foi usado apenas uma CI e esta possuirá 10 neurónios de acordo com o software



DTREG (Programa de modelação e previsão) que minimizou entre dois a trinta neurónios introduzidos com um salto de dois neurónios na sua escolha como é de notar no anexo E. A optimização dos neurónios para a CI encontra-se a seguir bem como pode ser visto graficamente no anexo F.

Tabela 6. Optimização do número de neurónios na camada escondida

Neurónios da CI	% de Mini-classificação
2	100.00000
4	97.06745
6	99.70674
8	100.00000
<b>10</b>	<b>97.06745</b>
11	100
12	98.53372
14	98.24047
16	100
18	100

Tamanho óptimo ⇨

Fonte: Adaptação do Autor

Além da optimização do número de neurónios na CI também foi possível identificar os PS (Anexo G) usados nesta camada bem como na CS, uma vez que a CE não possui PS por ser a camada que activa o funcionamento da rede.

#### 4.1.3.3 Implementação da Rede Neuronal

Para a implementação da RN, foram usadas duas STs, nomeadamente: a do gráfico 3 e a do gráfico 4. A ST do gráfico 3 correspondente a ST bruta foi usada como a ST de entrada na rede ou seja, como ST de alimentação da rede e a ST do gráfico 4 correspondente a primeira diferença como a ST ajustadora<sup>21</sup>, uma vez que estes (correspondentes a primeira diferença ou simplesmente gráfico 4) não possuem factores que possam influenciar na sua análise visto que já estão analisados quanto a estacionariedade, sazonalidade e tendência, que são factores importantíssimos quando o objectivo é implementar via RNs. O outro factor da utilização da série

<sup>21</sup> refere-se a ST adequada para Previsão.



correspondente ao gráfico 4 deve-se ao facto da utilização da aprendizagem supervisionada da rede na sua implementação.

A tabela 7 obtida por meio do software NeuralWare pela introdução da ST bruta (correspondente ao gráfico 3) e a correspondente a ST da primeira diferença (gráfico 4), permite a validação deste modelo por meio dos erros de previsão, o que de acordo com Lewis (1997) este modelo pode ser considerado potencialmente bom uma vez que o EPAM calculado na entrada, treino e na fase de teste da rede encontram-se na escala de um modelo considerado potencialmente bom como se pode ver na tabela 1 correspondente as potencialidade de previsão do EPAM.

Tabela 7. Dados obtidos pelo software NeuralWare

Valores de ajuste	Entrada	Treino	Teste
Autocorrelação	0.2342075	0.240271	0.2200172
EPAM	13.03302	12.82058	13.50617
Máximo absoluto	68.59068	68.59068	43.1016
EMQ	18.08336	18.27982	17.63793
Precisão (20%)	0.8028169	0.8163265	0.7727273
Intervalo de confiança (95%)	35.88759	36.6188	36.68519
Dados	71	49	22

Fonte: Adaptação do Autor

Com base nos dados usados para a modelação, o software *NeuralWare* gera valores de previsão adequados, como se pode ver na tabela 8.



Tabela 8. Valores de previsão gerados pelo s/w NeuralWare

Mes/Ano	Dados Brutos de HIV	Primeira Diferença	Valores previstos pelo S/W	Transformação Da ST de Previsão
Jan-00	100	-----	26	44
Fev-00	75	25	18	46
Mar-00	115	40	28	47
Abr-00	77	38	19	45
Mai-00	95	18	26	54
Jun-00	113	18	28	43
Jul-00	64	49	15	31
Ago-00	73	9	16	34
Set-00	42	31	18	35
Out-00	52	10	17	35
Nov-00	52	0	17	44
Dez-00	107	55	27	59
01-Jan	137	30	31	59
01-Fev	107	30	27	52
01-Mar	88	19	25	40
01-Abr	69	19	15	35
01-Mai	78	9	20	37
01-Jun	75	3	18	34
01-Jul	59	16	16	34
01-Ago	50	9	18	33
01-Set	61	11	16	32
01-Out	28	33	17	39
01-Nov	82	54	22	51
01-Dez	121	39	29	56
02-Jan	103	18	27	50
02-Fev	84	19	23	43
02-Mar	77	7	19	44
02-Abr	90	13	25	52
02-Mai	107	17	27	45
02-Jun	44	63	18	34
02-Jul	26	18	17	43
02-Ago	95	69	26	42
02-Set	57	38	16	35
02-Out	76	19	18	34
02-Nov	60	16	16	31
02-Dez	71	11	15	32
03-Jan	73	2	16	33
03-Fev	53	20	17	32
03-Mar	65	12	15	33
03-Abr	42	23	18	35
03-Mai	48	6	18	33
03-Jun	15	33	16	30
03-Jul	0	15	14	32
03-Ago	50	50	18	35
03-Set	46	4	18	35
03-Out	74	28	17	37



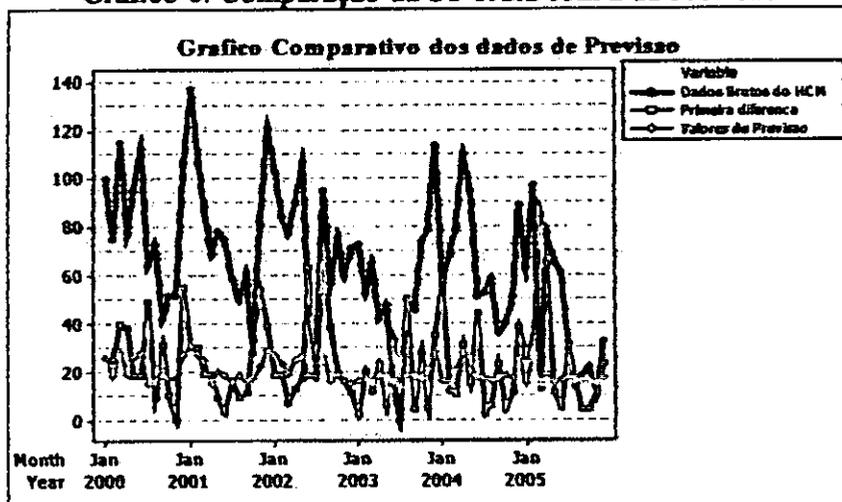
03-Nov	79	5	21	49
03-Dez	114	35	28	45
04-Jan	57	57	16	31
04-Fev	69	12	15	35
04-Mar	79	10	21	48
04-Abr	110	31	28	53
04-Mai	95	15	26	43
04-Jun	51	44	17	35
04-Jul	53	2	17	33
04-Ago	59	6	16	33
04-Set	36	23	17	35
04-Out	40	4	18	35
04-Nov	51	11	17	42
04-Dez	89	38	25	40
05-Jan	63	26	15	41
05-Fev	97	34	26	42
05-Mar	13	84	15	34
05-Abr	77	64	19	34
05-Mai	66	11	15	30
05-Jun	61	5	16	33
05-Jul	30	31	17	32
05-Ago	14	16	16	31
05-Set	18	4	16	32
05-Out	22	4	16	31
05-Nov	10	12	15	32
05-Dez	33	23	17	

Fonte: BS do HCM e Adaptação do Autor

Os dados da terceira coluna da tabela acima foram gerados pelo software *NeuralWare* e estes representam a melhor previsão. Para o melhor sustento da afirmação, este facto pode ser visto no gráfico a seguir que representa o comportamento da segunda, terceira e quarta coluna da tabela.



Gráfico 6. Comparação da ST bruta com a da Previsão

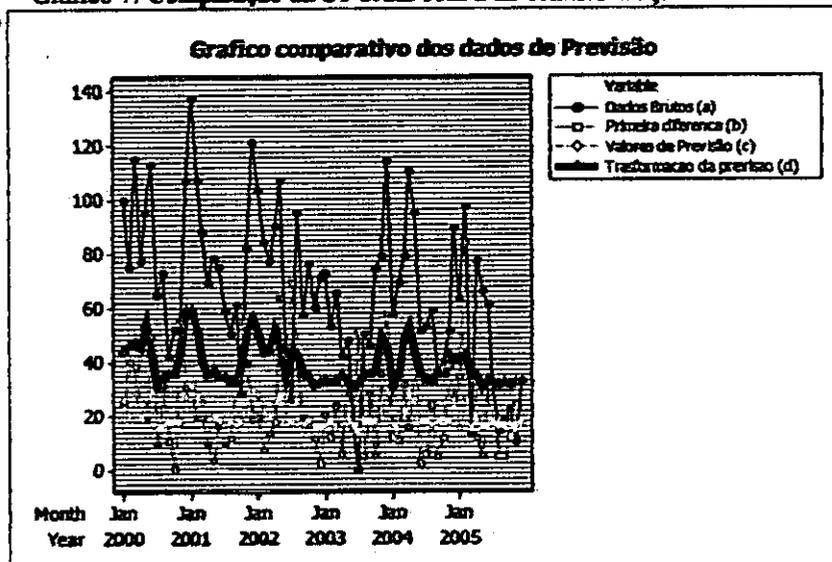


Fonte: Adaptação do Autor

Analisando o gráfico 6 nota-se que existe uma diferença muito maior entre a ST correspondente aos dados brutos, primeira diferença em relação a ST gerada pelo software NeuralWare. Esta diferença deve-se pelo facto dos dados obtidos pelo software terem se baseado nos da primeira diferença. Para a correcção desta maior diferença foi necessária a transformação da série de previsão gerada pelo software de modo que se adeque aos da primeira diferença, por estes corresponderem numa forma directa a ST proveniente da ST original e esta encontra-se na última coluna da tabela 8.

Representando graficamente as STs da tabela 8 para melhor visualização, obtém-se o seguinte:

Gráfico 7. Comparação da ST bruta com a da Transformação de Previsão



Fonte: Adaptação do Autor



Observando o gráfico 7 é de notar que os dados correspondentes a transformação da série de previsão aproximam-se aos da primeira diferença. Estes dados são os que correspondem a previsão da série em estudo pela modelação via RN.

## **4.2 Discussão dos Resultados**

### **4.2.1 Série Temporal recolhida**

A ST recolhida no BS do HCM referentes aos sangue infectado pelo vírus de HIV e que foi facultado pela Base de Dados em uso nesta unidade hospitalar, esta representa uma TQ como se pode ver na tabela 3 deste capítulo. A identificação da tendência da série recolhida foi de acordo com análise dos erros de previsão nomeadamente, DAM, EMQ e EPAM sendo a TQ a que apresenta o menor erro de previsão. Pelo facto da TQ apresentar menor erro de previsão para DAM, EMQ e EPAM., o principal erro em uso nesta trabalho é o EMQ segundo Lewis (1997). Para a determinação do DAM, EMQ e EPAM foram usadas as equações 2.12, 2.13, 2.14 que se encontram em 2.4 do capítulo II da revisão da literatura.

Quanto a análise da sazonalidade desta série, dizer que, pela observação do gráfico 1, é de notar que a série possui um movimento sazonal, visto que esta possui regularmente e em períodos fixos de tempo (principalmente nos meses de Novembro à Janeiro de cada ano) comportamentos quase iguais.

Discutindo a estacionariedade desta série, chega-se a conclusão que esta é não estacionária, e isto pode ser visto pela comparação da representação do gráfico 3 em relação ao gráfico 4, sendo o gráfico 3 o que representa a não estacionariedade da série e o gráfico 4 o que representa a estacionariedade da série.

Para reforçar o parágrafo acima veja os resultados do teste ADF encontrado na tabela 4 que vem a confirmar a estacionariedade da série correspondente ao gráfico 4. Com base na tabela 4, esta confirma de facto que a série é estacionária visto que os valores do teste da série a um nível de significância iguais a 0.01, 0.05 e 0.1 são todos eles inferiores ao valor crítico do teste ADF correspondente a 4.916249 em módulo.



#### 4.2.2 Parâmetros da RN

Os parâmetros usados para a RN encontram-se na tabela 5, sendo o número de variáveis de previsão correspondente a uma que tem a ver com a única série em estudo (HIV) com um número de observações correspondentes a 72; o modelo usado para a rede possui múltiplas camadas com uma alimentação directa, isto é, cada neurónio alimenta apenas os neurónios de frente e não os que se encontram atrás, sendo a rede composta por três camadas, a de entrada, que permite a entrada da série em estudo; a camada escondida (intermediária) que funciona como uma caixa negra e esta possui dez (10) neurónios como se pode ver na tabela 6 após uma criação progressiva da RN consoante o exemplo representado no anexo D; finalmente a camada de saída que ilustra os resultados da modelação que se encontram na última coluna da tabela 8.

As funções de activação usadas na camada escondida e na camada de saída são não lineares como ilustra a tabela 9 e o sustento desta escolha encontra-se no segundo parágrafo do ponto 2.7.5 da revisão da literatura; e a correcção dos PS é feita pelo MB, sendo os seus valores máximos e mínimos mostrados na tabela 9, bem como o número de neurónios para cada camada. Quanto ao processo de treinamento esta foi treinada através de um treinamento supervisionado ou aprendizagem supervisionada<sup>22</sup>, visto que para além da introdução da série a prever (dados bruto) também lhe foi atribuído a série desejada para a previsão (correspondentes a primeira diferença) ao qual a rede usou para a sua aprendizagem.

Tabela 9. Dados usados na arquitectura da RN

Camada	Neurónios	Função de activação	Peso mínimo	Peso máximo
Escondida	10	Logística	-3.028E+001	2.946E+001
Saída	71	Logística	-3.343E+001	2.125E+001

Fonte: Adaptação do Autor

#### 4.2.3 Resultados obtidos pela RN

Os resultados obtidos pela RN com a arquitectura mencionada em 4.1.3.2 da análise dos resultados encontram-se na tabela 7, sendo apresentados os valores de ajuste antes, durante o treino e após o teste. Olhando para o valor de autocorrelação, este possui os valores de autocorrelação antes, durante o treino (supervisionado) e durante o teste, nomeadamente,

<sup>22</sup> Pode ver a sua definição no capítulo da revisão da literatura no ponto 2.7.2.2



0.2342075, 0.240271 e 0.2200172 que são adequados mediante o intervalo aconselhado por este de acordo com o ponto 2.6 da revisão da literatura.

Quanto aos erros EPAM, DAM e EMQ, estes possuem erros relativamente menores comparados com os mesmos ilustrados na tabela 3. Avaliando estes erros pela modelação via RN, o erro EMQ apresenta o menor erro de previsão durante o teste comparando com o EMQ durante o treinamento e este (EMQ durante o treinamento) por ser menor que o EMQ antes da modelação, enquanto que o EPAM apresenta-se melhor durante o processo de treinamento comparando com o mesmo (EPAM) na entrada e durante o teste, o que para a sua inteira validade deveria ser menor durante o teste comparando com a entrada e treinamento. Este facto leva a validação do EMQ como sendo o erro que apresenta melhor empenho em RNs apesar do DAM comportar-se relativamente bem pelo facto deste possuir o menor erro durante o teste comparativamente a entrada e treino, este possui um erro de teste relativamente maior em relação ao EMQ, o que mais uma vez vem a provar que o EMQ comporta-me melhor em RN.

Pelo facto do EMQ ser o menor comparando-se com EPAM e DAM e segundo Lewis (1997), no ponto 2.4 da revisão da literatura, pode-se dizer que as RNs se comportam eficientemente para a previsão de STs.

O gráfico 7 representa a comparação da previsão da ST de HIV recolhida com a ST prevista pela a RN, sendo:

- A série em (a) o que representa a originalidade da série, isto é, a série recolhida sem tratamento algum; a série em (b) o que representa a série após um tratamento que consistiu na estacionariedade da série, visto que esta não era estacionária. Esta série por ser estacionária foi usada na altura da aprendizagem supervisionada de acordo com o primeiro parágrafo do ponto 4.1.3.3
- A série representada em (d) correspondente a transformação inversa dos dados de previsão, representa a previsão adequada para a série em estudo. Olhando para gráfico 7, é de notar que esta série, correspondente a previsão via RNs, comporta-se melhor que a série bruta e a da primeira diferença. Os valores de previsão desta série encontram-se na tabela 8 deste capítulo.



# **CAPÍTULO V**

## **Conclusões e Recomendações**

---

### **5.1 Conclusões**

Com base nos resultados apresentados, conclui-se que:

- ✓ Os métodos de previsão usadas no BS do HCM enquadram-se no grupo dos métodos de previsão qualitativos.
- ✓ A ST em estudo possui movimento sazonal, tendência exponencial e uma característica não estacionária.
- ✓ A topologia das RNs com uma CI apresenta melhor resultado quando usada para prever STs.
- ✓ As RNs constituem um bom modelo para a previsão de ST, visto que estas apresentam resultados bastante satisfatórios.

### **5.2 Recomendações**

Recomenda-se que:

- ✓ A comparação de métodos que utilizam redes neuronais com aqueles que usam metodologias convencionais, continua sendo investigada, considerando que ambos podem apresentar resultados comparáveis, principalmente para séries com base histórica elevada.
- ✓ Trabalhos que usen RNs continuem sendo desenvolvidos nas mais diversas áreas de investigação científica.



## REFERÊNCIAS BIBLIOGRÁFICAS

- Almeida, Fernando C. (1995).** *Desvendando o Uso de Redes Neurais em Problemas de Administração de Empresas*, jan./fev.
- ALVES, C. A. M (2001)** . *Uma Ferramenta de Extração de Regras de Redes Neurais*. Dissertação (Mestrado em Ciências em Engenharia Civil) – COPPE, UFRJ, Rio de Janeiro.
- Azoff, E. (1995).** *Neural Network Time Series Forecasting of Financial Markets*, John Wiley & Sons, New York, USA.
- 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, Instituto Tecnológico de Aeronáutica, São José dos Campos. Recuperado de <http://teses.eps.ufsc.br/defesa/pdf/10422.pdf> em 05 de Março de 2006.
- Chatfield C., 1996.** *The analysis of Time Series : An introduction*. Chapman & Hall, fifth edition, NY.283 pp. .
- Correia, M. (1995).** *Tecnologias Avançadas de Produção e Gestão das pessoas. Comportamento Organizacional e Gestão*. Vol 1, nº 2, pg. – 265-284.
- Cortez, Paulo Alexandre Ribeiro (1997).** *Algoritmos Genéricos e Redes Neurais na Previsão de Séries Temporais*. Dissertação de Mestrado. Guimarães: Universidade de Minho, Braga , Dezembro de 1997.
- Cortez, P.A. R (2002).** *Modelo Inspirado na Natureza para a Previsão de Séries Temporais*. Tese de Doutorado. Guimarães: Universidade de Minho.
- Dutta S., S. SHEKHAR et W.Y. WONG (1992).** *Decision support in non-conservative domains : generalization with neural networks*. Paris: INSEAD, 1992. (Working Paper nº 92- 31)
- Eastman & Mckendry (1991 )**. *Explorations in Geographic Information Systems Technology*, Volume 1, Change and time series analysis. Geneva. UNITAR European Office.



- Ehlers, R. S. (2005). *Análise de Séries Temporais*. Departamento de Estatística, UFPR. Disponível em <http://www.est.ufpr.br/~ehlers/notas>. Acessado em : 21 de Outubro de 2006.
- Ender, W. (1995). *Applied Econometric Times Series*. John Wiley and Sons.
- Garcia, João (2002). *Aplicação de Redes Neurais de Competição à Desmodulação de Sinal em Modulação Digital*. Portugal, Lisboa, Novembro de 2002.
- Hanke, J. and Reitsch, A. (1989). *Business Forecasting*, 3ª Edição, Allyn and Bacon, Massachussetts, USA.
- Hayken, S. (2001). *Redes Neurais – Princípios e práticas*, Ed. Bookman, Porto Alegre.
- Jacinto, A. M; Moreira, T. C. S; Pessanha, V. G. (2001). *Optimização da demanda Futura: Uma aplicação prática das redes neurais artificiais e dos algoritmos genéticos*, Brazil, Rio de Janeiro, Junho – 2004.
- Jenkins, G. M. and Watts, D. G., 1968. *Spectral Analysis and its Applications*. Holden – Day, SF, 525 pp.
- Latorre e Cardoso (2001). *Time Series analysis in epidemiology: an introduction to methodological aspects*. Rev. Bras. Epidemiol. Vol.4, Nº 3.
- LEWIS, C. D. (1997), *Demand Forecasting and Inventory Control*. New York: Wiley.
- Martin, Claudio (2000). *Aplicação de redes neurais para prognóstico com base em séries temporais*, Florianópolis, Novembro-2000.
- McCulloch, W. W. and Pitts, W., *A logical calculus of the ideas imminent in nervous activity*, Bull. Math. Biophys., no. 5, pp. 115-133, 1943.
- Moreira, Miguel Ângelo (1997). *Introdução às Redes Neurais Artificiais*, Outubro de 1997.
- Morettin PA, Toli CMC. *Previsão de series temporais*. 2ª Ed. São Paulo : Actual Editora;1985.
- Russell. S and Norvig, P. (1995). *Artificial Intelligence: A Modern Approach Prentice*. Prentice-Hall, New Jersey, USA.



Spiegel, Murray R. *Coleção Schaum*, Estatística, 3ª Edição, São Paulo (SP), Brasil, 1994.

Weigend, A. and Gershenfeld, N. (1994). *Time Series Prediction*, Forecasting the Future and Understanding the Past. Addison-Wesley, USA.

Wikipédia (2007). *A enciclopédia livre*. Retirado de [http://pt.wikipedia.org/wiki/Rede\\_neural](http://pt.wikipedia.org/wiki/Rede_neural) em 10.04.2007.

Wilks, D. S.(1995). *Statistical methods in the atmospheric Sciences*. Academic Press, NY, 468pp.



**ANEXOS**



ANEXO A: Dados Brutos recolhidos no HCM.

	DOAÇÕES DE SANGUE				INUTILIZADAS POR HIV
	Meses	Voluntários	Repositores	Total	
2000	Janeiro	573	475	1,048	100
	Fevereiro	485	353	838	75
	Março	770	556	1,326	115
	Abril	813	360	1,173	77
	Maiο	833	511	1,344	95
	Junho	761	422	1,272	113
	Julho	669	311	1,091	64
	Agosto	1025	251	1,336	73
	Setembro	553	194	804	42
	Outubro	627	226	821	52
	Novembro	807	306	1,033	52
	Dezembro	643	647	949	107
	TOTAL	8,559	4,612	13,171	965
2001	Janeiro	876	598	1,474	137
	Fevereiro	1,049	453	1,502	107
	Março	596	579	1,175	88
	Abril	746	475	1,221	69
	Maiο	694	575	1,269	78
	Junho	546	480	1,026	75
	Julho	847	325	1,172	59
	Agosto	817	214	1,031	50
	Setembro	1,134	210	1,344	61
	Outubro	874	329	1,203	28
	Novembro	582	456	1,038	82
	Dezembro	831	541	1,372	121
	TOTAL	9,592	5,235	14,827	955
2002	Janeiro	693	634	1,327	103
	Fevereiro	749	429	1,178	84
	Março	1,171	417	1,588	77
	Abril	795	375	1,170	90
	Maiο	1,915	348	2,263	107
	Junho	553	97	650	44
	Julho	759	242	1,001	26
	Agosto	1,509	269	1,778	95
	Setembro	790	130	920	57
	Outubro	845	315	1,160	76
	Novembro	767	295	1,062	60
	Dezembro	1,046	240	1,286	71
	TOTAL	11,592	3,791	15,383	890



2003	Janeiro	863	272	1,135	73
	Fevereiro	1,044	220	1,264	53
	Março	1,045	240	1,285	65
	Abril	946	234	1,180	42
	Maio	1,191	227	1,418	48
	Junho	1,037	187	1,224	15
	Julho	1,108	250	1,358	0
	Agosto	1,254	275	1,529	50
	Setembro	1,195	194	1,389	46
	Outubro	1,024	289	1,313	74
	Novembro	755	367	1,122	79
	Dezembro	1,106	369	1,475	114
	TOTAL	12,568	3,124	15,692	659
2004	Janeiro	667	313	980	57
	Fevereiro	704	322	1,026	69
	Março	1,183	379	1,562	79
	Abril	1,124	575	1,699	110
	Maio	1,017	493	1,510	95
	Junho	1,115	408	1,523	51
	Julho	1,068	314	1,382	53
	Agosto	1,271	217	1,488	59
	Setembro	1,057	214	1,271	36
	Outubro	727	221	948	40
	Novembro	783	254	1,037	51
	Dezembro	1,013	288	1,301	89
	TOTAL	11,729	3,998	15,727	789
2005	Janeiro	1,114	361	1,475	63
	Fevereiro	1,306	425	1,731	97
	Março	1,557	438	1,995	13
	Abril	1,457	452	1,909	77
	Maio	1,614	440	2,054	66
	Junho	1,469	407	1,876	61
	Julho	920	318	1,238	30
	Agosto	1,069	270	1,339	14
	Setembro	737	201	938	18
	Outubro	1,046	312	1,358	22
	Novembro	1,090	249	1,339	10
	Dezembro	998	218	1,216	33
	TOTAL	14,377	4,091	18,468	504

Fonte: HCM



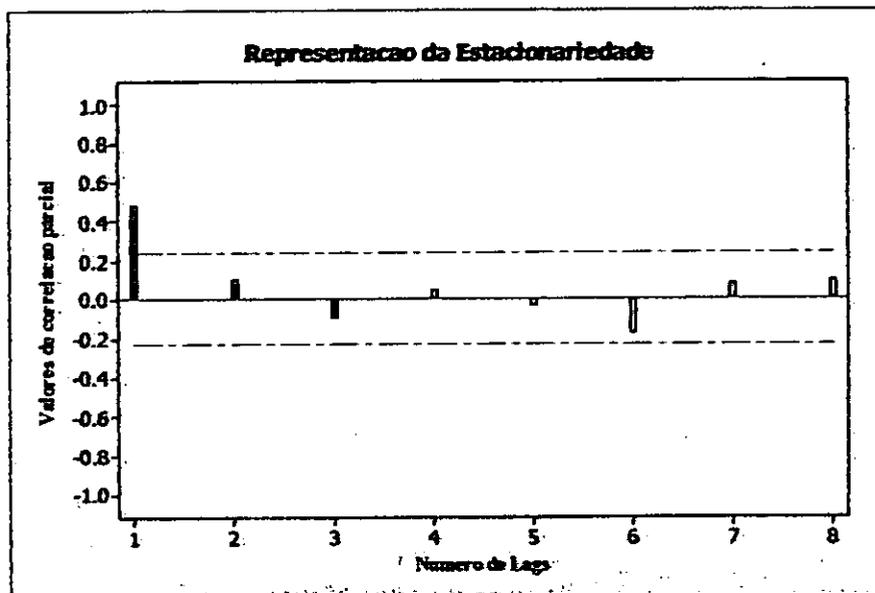
ANEXO B: Determinação da Primeira diferença dos dados Brutos

Mes/Ano	Dados Brutos de HIV	Primeira Diferença
Jan-00	100	-----
Fev-00	75	25
Mar-00	115	40
Abr-00	77	38
Mai-00	95	18
Jun-00	113	18
Jul-00	64	49
Ago-00	73	9
Set-00	42	31
Out-00	52	10
Nov-00	52	0
Dez-00	107	55
01-Jan	137	30
01-Fev	107	30
01-Mar	88	19
01-Abr	69	19
01-Mai	78	9
01-Jun	75	3
01-Jul	59	16
01-Ago	50	9
01-Set	61	11
01-Out	28	33
01-Nov	82	54
01-Dez	121	39
02-Jan	103	18
02-Fev	84	19
02-Mar	77	7
02-Abr	90	13
02-Mai	107	17
02-Jun	44	63
02-Jul	26	18
02-Ago	95	69
02-Set	57	38
02-Out	76	19
02-Nov	60	16
02-Dez	71	11
03-Jan	73	2
03-Fev	53	20
03-Mar	65	12
03-Abr	42	23
03-Mai	48	6
03-Jun	15	33
03-Jul	1	15
03-Ago	50	50
03-Set	46	4
03-Out	74	28
03-Nov	79	5



03-Dez	114	35
04-Jan	57	57
04-Fev	89	12
04-Mar	79	10
04-Abr	110	31
04-Mai	95	15
04-Jun	51	44
04-Jul	53	2
04-Ago	59	6
04-Set	38	23
04-Out	40	4
04-Nov	51	11
04-Dez	89	38
05-Jan	63	26
05-Fev	97	34
05-Mar	13	84
05-Abr	77	64
05-Mai	66	11
05-Jun	61	5
05-Jul	30	31
05-Ago	14	16
05-Set	18	4
05-Out	22	4
05-Nov	10	12
05-Dez	33	23

ANEXO C: Representação da não estacionariedade dos dados (Brutos) de HIV



Fonte: Adaptação do Autor



ANEXO D: Análise dos dados pelo software DTREG

Neural Network creation progress

Neural Network Build Progress

Analysis time:	00:03:33
Network evaluations:	4,452,963
Current error:	2.843032e-001
Hidden layer neurons:	10

Network building activity

- Creating cross-validation fold 2 of 4
- Using differential evolution to find global starting parameter values
- Creating initial population (12%)

Stop Analysis



### ANEXO E: Ambiente do software DTREG para a introdução dos dados

Model

Cascade Correlation | Discriminant Analysis | Logistic Regression | Class labels | Initial split  
Category weights | Misclassification | Missing data | Variable weights | DTL | Scoring | Translate | Misc.  
Design | Data | Variables | Single Tree | TreeBoost | Decision Tree Forest | SVM | MLFN Network | PNN/GRNN

----- Multi-Layer Feed-forward Neural Networks (MLFN) -----

Type of model to build  
MLFN neural network

Number of network layers  
 3 layers (1 hidden)  
 4 layers (2 hidden)

Automatic hidden layer neuron selection  
 Automatically optimize hidden layer 1  
Min: 2 Max: 30 Step: 2  
Max. steps without change: 4  
% rows to use for search: 100  
 Cross validate; folds: 4  
 Hold-out sample %: 20

Number of neurons for hidden layers  
Layer 1: 10 Layer 2: 4

Model testing and validation  
 No validation, use all data rows  
 Random percent: 20  
 V-fold cross-validation: 10  
 Leave-one-out validation

How to handle missing predictor variable values  
 Don't use rows with missing predictors  
 Replace missing predictors with medians

Options  
 Compute importance of variables

Hidden layer activation function  
Logistic

Output layer activation function  
Logistic

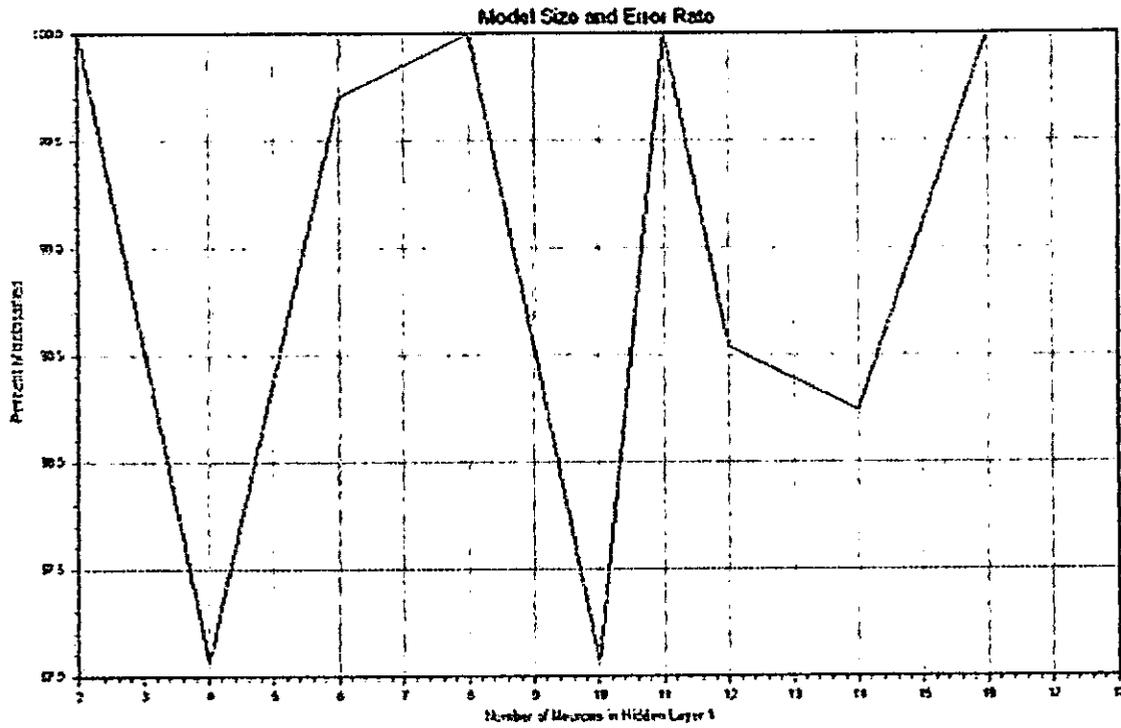
Advanced options

OK Cancel Apply



**ANEXO F: Gráfico da otimização do número de neurónios na camada escondida**

Error rate versus number of hidden neurons





Anexo G: Pesos usados para a rede

=====  
----- Misclassification Tables -----  
=====

--- Training Data ---

Category	-----Actual-----		-----Misclassified-----			
	Count	Weight	Count	Weight	Percent	Cost
1-Apr	0	0.0	0	0.0	0.000	0.000
1-Aug	1	0.6	1	0.6	100.000	1.000
1-Dec	1	0.5	1	0.5	100.000	1.000
1-Feb	0	0.0	0	0.0	0.000	0.000
1-Jan	1	0.6	1	0.6	100.000	1.000
1-Jul	1	0.5	1	0.5	100.000	1.000
1-Jun	0	0.0	0	0.0	0.000	0.000
1-Mar	0	0.0	0	0.0	0.000	0.000
1-May	1	0.8	1	0.8	100.000	1.000
1-Nov	0	0.0	0	0.0	0.000	0.000
1-Oct	0	0.0	0	0.0	0.000	0.000
1-Sep	0	0.0	0	0.0	0.000	0.000
2-Apr	0	0.0	0	0.0	0.000	0.000
2-Aug	0	0.0	0	0.0	0.000	0.000
2-Dec	0	0.0	0	0.0	0.000	0.000
2-Feb	1	0.6	1	0.6	100.000	1.000
2-Jan	0	0.0	0	0.0	0.000	0.000
2-Jul	1	0.8	1	0.8	100.000	1.000
2-Jun	1	0.5	1	0.5	100.000	1.000
2-Mar	0	0.0	0	0.0	0.000	0.000
2-May	1	0.9	1	0.9	100.000	1.000
2-Nov	0	0.0	0	0.0	0.000	0.000
2-Oct	1	0.8	1	0.8	100.000	1.000
2-Sep	0	0.0	0	0.0	0.000	0.000
3-Apr	1	0.7	1	0.7	100.000	1.000
3-Aug	0	0.0	0	0.0	0.000	0.000
3-Dec	1	0.6	1	0.6	100.000	1.000
3-Feb	1	0.8	1	0.8	100.000	1.000
3-Jan	1	0.5	1	0.5	100.000	1.000
3-Jul	1	0.5	1	0.5	100.000	1.000
3-Jun	0	0.0	0	0.0	0.000	0.000
3-Mar	1	0.5	1	0.5	100.000	1.000
3-May	1	0.9	1	0.9	100.000	1.000



3-Nov	1	0.8	1	0.8	100.000	1.000
3-Oct	1	0.9	0	0.0	0.000	0.000
3-Sep	0	0.0	0	0.0	0.000	0.000
4-Apr	1	0.6	1	0.6	100.000	1.000
4-Aug	0	0.0	0	0.0	0.000	0.000
4-Dec	0	0.0	0	0.0	0.000	0.000
4-Feb	1	0.8	1	0.8	100.000	1.000
4-Jan	0	0.0	0	0.0	0.000	0.000
4-Jul	1	1.0	1	1.0	100.000	1.000
4-Jun	1	0.6	1	0.6	100.000	1.000
4-Mar	0	0.0	0	0.0	0.000	0.000
4-May	1	0.5	1	0.5	100.000	1.000
4-Nov	0	0.0	0	0.0	0.000	0.000
4-Oct	0	0.0	0	0.0	0.000	0.000
4-Sep	1	0.7	1	0.7	100.000	1.000
5-Apr	0	0.0	0	0.0	0.000	0.000
5-Aug	1	1.0	1	1.0	100.000	1.000
5-Dec	1	0.8	1	0.8	100.000	1.000
5-Feb	1	0.8	1	0.8	100.000	1.000
5-Jan	0	0.0	0	0.0	0.000	0.000
5-Jul	1	0.8	1	0.8	100.000	1.000
5-Jun	0	0.0	0	0.0	0.000	0.000
5-Mar	1	0.8	1	0.8	100.000	1.000
5-May	1	0.5	1	0.5	100.000	1.000
5-Nov	0	0.0	0	0.0	0.000	0.000
5-Oct	1	0.6	1	0.6	100.000	1.000
5-Sep	0	0.0	0	0.0	0.000	0.000
Apr-00	0	0.0	0	0.0	0.000	0.000
Aug-00	1	0.7	1	0.7	100.000	1.000
Dec-00	0	0.0	0	0.0	0.000	0.000
Feb-00	1	0.6	1	0.6	100.000	1.000
Jan-00	1	0.5	1	0.5	100.000	1.000
Jul-00	0	0.0	0	0.0	0.000	0.000
Jun-00	1	0.7	1	0.7	100.000	1.000
Mar-00	1	0.7	1	0.7	100.000	1.000
May-00	0	0.0	0	0.0	0.000	0.000
Nov-00	0	0.0	0	0.0	0.000	0.000
Oct-00	1	0.6	1	0.6	100.000	1.000
Sep-00	1	0.6	1	0.6	100.000	1.000
Total	39	26.7	38	25.8	96.629	0.966



--- Validation Data ---

-----Actual-----			-----Misclassified-----			
Category	Count	Weight	Count	Weight	Percent	Cost
1-Apr	1	0.2	1	0.2	100.000	1.000
1-Aug	1	0.6	1	0.6	100.000	1.000
1-Dec	1	0.5	1	0.5	100.000	1.000
1-Feb	1	0.4	1	0.4	100.000	1.000
1-Jan	1	0.6	1	0.6	100.000	1.000
1-Jul	1	0.5	1	0.5	100.000	1.000
1-Jun	1	0.4	1	0.4	100.000	1.000
1-Mar	1	0.1	1	0.1	100.000	1.000
1-May	1	0.8	1	0.8	100.000	1.000
1-Nov	0	0.0	0	0.0	0.000	0.000
1-Oct	1	0.2	1	0.2	100.000	1.000
1-Sep	1	0.1	1	0.1	100.000	1.000
2-Apr	1	0.2	1	0.2	100.000	1.000
2-Aug	1	0.4	1	0.4	100.000	1.000
2-Dec	1	0.4	1	0.4	100.000	1.000
2-Feb	1	0.6	1	0.6	100.000	1.000
2-Jan	1	0.2	1	0.2	100.000	1.000
2-Jul	1	0.8	1	0.8	100.000	1.000
2-Jun	1	0.5	1	0.5	100.000	1.000
2-Mar	1	0.3	1	0.3	100.000	1.000
2-May	1	0.9	1	0.9	100.000	1.000
2-Nov	1	0.4	1	0.4	100.000	1.000
2-Oct	1	0.8	1	0.8	100.000	1.000
2-Sep	1	0.4	1	0.4	100.000	1.000
3-Apr	1	0.7	1	0.7	100.000	1.000
3-Aug	1	0.4	1	0.4	100.000	1.000
3-Dec	1	0.6	1	0.6	100.000	1.000
3-Feb	1	0.8	1	0.8	100.000	1.000
3-Jan	1	0.5	1	0.5	100.000	1.000
3-Jul	1	0.5	1	0.5	100.000	1.000
3-Jun	1	0.4	1	0.4	100.000	1.000
3-Mar	1	0.5	1	0.5	100.000	1.000
3-May	1	0.9	1	0.9	100.000	1.000
3-Nov	1	0.8	1	0.8	100.000	1.000
3-Oct	1	0.9	1	0.9	100.000	1.000



3-Sep	1	0.1	1	0.1	100.000	1.000
4-Apr	1	0.6	1	0.6	100.000	1.000
4-Aug	0	0.0	0	0.0	0.000	0.000
4-Dec	1	0.1	1	0.1	100.000	1.000
4-Feb	1	0.8	1	0.8	100.000	1.000
4-Jan	1	0.2	1	0.2	100.000	1.000
4-Jul	1	1.0	1	1.0	100.000	1.000
4-Jun	1	0.6	1	0.6	100.000	1.000
4-Mar	1	0.3	1	0.3	100.000	1.000
4-May	1	0.5	1	0.5	100.000	1.000
4-Nov	1	0.1	1	0.1	100.000	1.000
4-Oct	1	0.3	1	0.3	100.000	1.000
4-Sep	1	0.7	1	0.7	100.000	1.000
5-Apr	1	0.2	1	0.2	100.000	1.000
5-Aug	1	1.0	1	1.0	100.000	1.000
5-Dec	1	0.8	1	0.8	100.000	1.000
5-Feb	1	0.8	1	0.8	100.000	1.000
5-Jan	0	0.0	0	0.0	0.000	0.000
5-Jul	1	0.8	1	0.8	100.000	1.000
5-Jun	0	0.0	0	0.0	0.000	0.000
5-Mar	1	0.8	1	0.8	100.000	1.000
5-May	1	0.5	1	0.5	100.000	1.000
5-Nov	1	0.2	1	0.2	100.000	1.000
5-Oct	1	0.6	1	0.6	100.000	1.000
5-Sep	1	0.1	1	0.1	100.000	1.000
Apr-00	0	0.0	0	0.0	0.000	0.000
Aug-00	1	0.7	1	0.7	100.000	1.000
Dec-00	1	0.2	1	0.2	100.000	1.000
Feb-00	1	0.6	1	0.6	100.000	1.000
Jan-00	1	0.5	1	0.5	100.000	1.000
Jul-00	1	0.4	1	0.4	100.000	1.000
Jun-00	1	0.7	1	0.7	100.000	1.000
Mar-00	1	0.7	0	0.0	0.000	0.000
May-00	1	0.4	1	0.4	100.000	1.000
Nov-00	1	0.3	1	0.3	100.000	1.000
Oct-00	1	0.6	1	0.6	100.000	1.000
Sep-00	1	0.6	1	0.6	100.000	1.000
Total	67	34.1	66	33.4	97.947	0.979



## **ANEXO H: Estrutura da entrevista semi estruturada feita aos agentes do BS do HCM**

### **Questionário feito ao BS do HCM**

- O que é o Banco de Sangue?
- Quem deve se dirigir ao banco de Sangue?
- Existe alguma classificação quanto as pessoas que doam sangue?
- Como são classificados os dadores
- Quem é considerado dador voluntário?
- Quem é considerado dador repositório?
- Que medidas são usadas para garantir o sigilo dos dados, principalmente para os dadores positivos?
- Como se identifica o sangue infectado pelo HIV do não infectado?
- Qual é o tratamento do sangue infectado do não infectado?
- Como são armazenados os registos dos dadores de sangue?
- Em que altura do ano se regista maior número de dadores? Que motivos levam a esse aumento?
- São usadas algumas técnicas para prever o número de dadores?
- Que técnicas são usadas?
- As técnicas usadas são eficientes?



**ANEXO I: Dados fornecidos pelo software SITRAM**



## Doações por Tipo de Dador

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

Ano / Mes	N° de Doações				Total
	Voluntários		Repositores		
	N°	%	N°	%	
<b>2000</b>					
Janeiro	573	54	475	45	1059
Fevereiro	485	57	353	42	847
Marco	770	58	556	42	1336
Abril	813	68	360	30	1190
Maió	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
<b>Totais do Período</b>	<b>8559</b>	<b>64.11</b>	<b>4612</b>	<b>34.54</b>	<b>13351</b>

# 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	%
<b>2000</b>													
<b>Janeiro</b>													
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												
<b>Fevereiro</b>													
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												
<b>Marco</b>													
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												
<b>Abril</b>													
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	%
<b>Maio</b>													
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												
<b>Junho</b>													
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												
<b>Julho</b>													
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												
<b>Agosto</b>													
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												
<b>Setembro</b>													

Ano / Mes	Total	Faixa Etária											
		16 - 19		20 - 29		30 - 39		40 - 49		50 - 59		> 60	
		N°	%	N°	%	N°	%	N°	%	N°	%	N°	%
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	33.5E+03	27	3069	23 E+03	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	%	
<b>2001</b>					
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
<b>Totais do Período</b>	<b>9592</b>	<b>63.79</b>	<b>5235</b>	<b>34.81</b>	<b>15038</b>

# 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	%
<b>2001</b>													
<b>Janeiro</b>													
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												
<b>Fevereiro</b>													
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												
<b>Marco</b>													
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												
<b>Abril</b>													
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	41.4E+03	25	3056	20.3 E+03	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	%	
<b>2002</b>					
Janeiro	693	52	634	47	1339
Fevereiro	749	63	429	36	1187
Marco	1171	73	417	26	1600
Abril	795	66	375	31	1197
Maió	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
<b>Totais do Período</b>	<b>11592</b>	<b>74.55</b>	<b>3791</b>	<b>24.38</b>	<b>15549</b>

# 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	%
<b>2002</b>													
<b>Janeiro</b>													
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												
<b>Fevereiro</b>													
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												
<b>Marco</b>													
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												
<b>Abril</b>													
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	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
<b>Maio</b>													
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												
<b>Junho</b>													
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												
<b>Julho</b>													
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												
<b>Agosto</b>													
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												
<b>Setembro</b>													

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												
<b>Outubro</b>													
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												
<b>Novembro</b>													
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												
<b>Dezembro</b>													
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	51.2	21	2598	16.7	7.4	544	3.5		
Infectadas	1042	2	0.19	378	36.3	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	%	
<b>2003</b>					
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
<b>Totais do Período</b>	<b>12568</b>	<b>79.19</b>	<b>3124</b>	<b>19.68</b>	<b>15870</b>



# 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	%
<b>2003</b>													
<b>Janeiro</b>													
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												
<b>Fevereiro</b>													
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												
<b>Marco</b>													
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												
<b>Abril</b>													
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	%	
<b>2004</b>					
Janeiro	667	67	313	32	991
Fevereiro	704	68	322	31	1037
Marco	1183	75	379	24	1577
Abril	1124	66	575	34	1714
Maió	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
<b>Totais do Período</b>	<b>11729</b>	<b>73.94</b>	<b>3998</b>	<b>25.20</b>	<b>15862</b>

# 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	%
<b>2004</b>													
<b>Janeiro</b>													
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												
<b>Fevereiro</b>													
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												
<b>Marco</b>													
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												
<b>Abril</b>													
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	49.7	1703	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	%	
<b>2005</b>					
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
<b>Totais do Período</b>	<b>14377</b>	<b>77.42</b>	<b>4091</b>	<b>22.03</b>	<b>18570</b>



## 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	%
<b>2005</b>													
<b>Janeiro</b>													
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												
<b>Fevereiro</b>													
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												
<b>Marco</b>													
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												
<b>Abril</b>													
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	%
<b>Maio</b>													
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												
<b>Junho</b>													
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												
<b>Julho</b>													
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												
<b>Agosto</b>													
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												
<b>Setembro</b>													

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