



Ministério da  
Ciência e Tecnologia



INPE-15191-TDI/1300

## UTILIZAÇÃO DE REDE NEURAL LVQ PARA PREVISÃO DO NÍVEL DO RIO PARAGUAI

Yasuo Kono

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada,  
orientada pelos Drs. José Demisio Simões da Silva, e Rafael Duarte Coelho dos  
Santos, aprovada em 11 de dezembro de 2007

Registro do documento original:

<<http://urlib.net/sid.inpe.br/mtc-m17@80/2007/11.27.17.30>>

INPE  
São José dos Campos  
2008

## **PUBLICADO POR:**

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3945-6911/6923

Fax: (012) 3945-6919

E-mail: [pubtc@sid.inpe.br](mailto:pubtc@sid.inpe.br)

## **CONSELHO DE EDITORAÇÃO:**

### **Presidente:**

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

### **Membros:**

Dr<sup>a</sup> Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Haroldo Fraga de Campos Velho - Centro de Tecnologias Especiais (CTE)

Dr<sup>a</sup> Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Dr. Ralf Gielow - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr. Wilson Yamaguti - Coordenação Engenharia e Tecnologia Espacial (ETE)

## **BIBLIOTECA DIGITAL:**

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Jefferson Andrade Ancelmo - Serviço de Informação e Documentação (SID)

Simone A. Del-Ducca Barbedo - Serviço de Informação e Documentação (SID)

## **REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Marilúcia Santos Melo Cid - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

## **EDITORAÇÃO ELETRÔNICA:**

Viveca Sant´Ana Lemos - Serviço de Informação e Documentação (SID)



Ministério da  
Ciência e Tecnologia



INPE-15191-TDI/1300

## UTILIZAÇÃO DE REDE NEURAL LVQ PARA PREVISÃO DO NÍVEL DO RIO PARAGUAI

Yasuo Kono

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada,  
orientada pelos Drs. José Demisio Simões da Silva, e Rafael Duarte Coelho dos  
Santos, aprovada em 11 de dezembro de 2007

Registro do documento original:

<<http://urlib.net/sid.inpe.br/mtc-m17@80/2007/11.27.17.30>>

INPE  
São José dos Campos  
2008

Dados Internacionais de Catalogação na Publicação (CIP)

---

Kono, Yasuo.

K837u Utilização de rede neural LVQ para previsão do nível do Rio Paraguai / Yasuo Kono. – São José dos Campos : INPE, 2008.  
151 p. ; (INPE-15191-TDI/1300)

Dissertação (Mestrado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2007.

Orientadores : Drs. José Demisio Simões da Silva, e Rafael Duarte Coelho dos Santos.

1. LVQ. 2. Rede neural artificial. 3.Previsão. 4. Nível de Rio. 5. Rio Paraguai. I.Título.

CDU 681.3.019

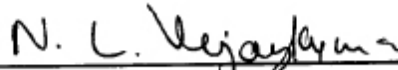
---

Copyright © 2008 do MCT/INPE. Nenhuma parte desta publicação pode ser reproduzida, armazenada em um sistema de recuperação, ou transmitida sob qualquer forma ou por qualquer meio, eletrônico, mecânico, fotográfico, reprográfico, de microfilmagem ou outros, sem a permissão escrita do INPE, com exceção de qualquer material fornecido especificamente com o propósito de ser entrado e executado num sistema computacional, para o uso exclusivo do leitor da obra.

Copyright © 2008 by MCT/INPE. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, microfilming, or otherwise, without written permission from INPE, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use of the reader of the work.

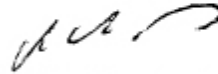
Aprovado (a) pela Banca Examinadora  
em cumprimento ao requisito exigido para  
obtenção do Título de Mestre em  
Computação Aplicada

Dr. Nandamudi Lankalapalli Vijaykumar



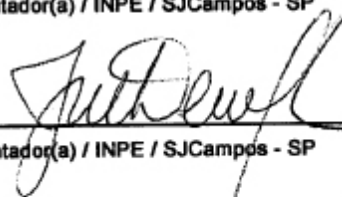
Presidente / INPE / SJC Campos - SP

Dr. Rafael Duarte Coelho dos Santos



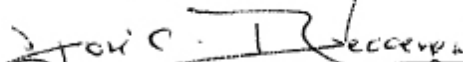
Orientador(a) / INPE / SJC Campos - SP

Dr. José Demisio Simões da Silva



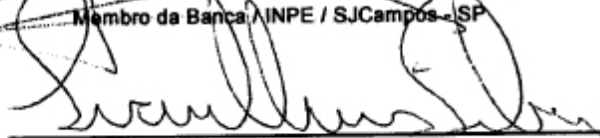
Orientador(a) / INPE / SJC Campos - SP

Dr. José Carlos Becceneri



Membro da Banca / INPE / SJC Campos - SP

Dr. Ivan Nunes da Silva



Convidado(a) / USP-São Carlos / São Carlos - SP

Aluno (a): Yasuo Kono

São José dos Campos, 11 de Dezembro de 2007

*“Seja como for, a grandiosa 'revolução humana' de uma única pessoa irá um dia impulsionar a mudança total do destino de um país e, além disso, será capaz de transformar o destino de toda a humanidade.”*

*Daisaku Ikeda*



*A meus pais Nobuo (in Memoriam) e Sayuri Kono, meus irmãos e a Fátima,  
companheira de todos os momentos.*





## **AGRADECIMENTOS**

Agradeço a todas as pessoas que me ajudaram a concluir este curso de Mestrado.

Ao Instituto Nacional de Pesquisas Espaciais - INPE pela permissão e oportunidade de continuar meus estudos e aperfeiçoamento tanto pessoal quanto profissional, assim como pela utilização de suas instalações.

Aos professores do INPE pelos valiosos ensinamentos e conhecimentos.

Ao Prof. Dr. José Carlos Becceneri pela orientação, apoio e principalmente pela paciência e coragem na realização deste trabalho.

Ao Prof. Dr. Rafael Duarte Coelho dos Santos pela orientação, apoio e dicas importantíssimas para a conclusão deste trabalho.

Ao orientador Prof. Dr. José Demísio Simões da Silva pela orientação, conhecimentos e entre muitas outras qualidades, pela perseverança em incentivar-me nos momentos cruciais.

Aos meus amigos de todas as horas, Amarildo, Antonio Egydio, Capra, Cassiano, Célia, Edenilse, Eduardo, Gatto, Isolette, Jadir, Joaquim, Luciana, Paulinho, Roberto, Rodolfo, e muitos outros, pelo incentivo e confiança.

À minha grande companheira e fonte de incentivos eternos, Fátima.

A meus pais pelos ensinamentos da importância da educação, dos estudos e do companheirismo, valores imutáveis nas relações humanas.

A meus irmãos, Miwa e Keiji, por vários tipos de apoio que recebi para todo este trabalho.



## RESUMO

Redes Neurais Artificiais têm sido utilizadas nos mais diversos campos do conhecimento para auxiliar na solução eficiente de problemas complexos e que, muitas vezes, processam grande volume de dados. Uma aplicação de sucesso é a previsão da vazão de um rio, utilizado para controle de reservatórios e barragens de água, bem como para alerta de enchente. A previsão de nível de rio também pode ser utilizada para esse alerta ou para conhecer o comportamento desse rio no futuro. Este trabalho de Dissertação de Mestrado propõe uma metodologia para realizar previsões de nível do rio Paraguai utilizando uma Rede Neural Artificial do tipo LVQ – *Learning Vector Quantization*. Os resultados dos estudos mostraram a capacidade da rede LVQ na previsão do nível do rio, gerando erros de 0% em alguns experimentos, utilizando um modelo gerado a partir dos dados observados e medidos. Portanto, o método proposto nesta dissertação apresenta-se como uma alternativa viável e de fácil calibração para a geração de alarmes em épocas de “cheia” ou de “vazante”.



# **USING LEARNING VECTOR QUANTIZATION NEURAL NETWORK TO PARAGUAY RIVER LEVEL PREDICTION**

## **ABSTRACT**

Artificial Neural Networks have been used in a variety of applications to assist in the solution of complex problems that need to process huge volumes of data. A successful application is the forecast of the outflow of a river, used to control reservoirs and water barrages, as well as for flooding alerts. The forecasting of a river level can also be used to know its future behavior. This dissertation proposes a methodology to carry out forecasts of the level of the Paraguay River, using a Learning Vector Quantization Artificial Neural Network (LVQ). The results of the studies showed the capacity of the LVQ to forecast the level of the river. Some experiments, using a model derived from observed and measured data, got 0% errors. Thus, the method proposed in this work is a viable alternative and easy to calibrate that can be used to generate alarms at flooding or low level phenomena.



## SUMÁRIO

	<u>Pág.</u>
<b>1 INTRODUÇÃO.....</b>	<b>23</b>
<b>2 FUNDAMENTOS TEÓRICOS.....</b>	<b>29</b>
2.1 Rede Neural Artificial .....	29
2.2 Formas de Aprendizado.....	32
2.3 Tipos de RNA.....	32
2.4 Análise de Séries Temporais .....	45
2.5 Utilização de RNA em Séries Temporais .....	50
2.6 Utilização de RNA em Hidrologia.....	52
<b>3 PREPARAÇÃO DOS DADOS.....</b>	<b>55</b>
3.1 Dados .....	55
3.2 Redução.....	62
3.3 Padrões de Nível do rio Paraguai.....	68
3.4 Geração dos dados de treinamento .....	69
<b>4 TREINAMENTO DA REDE LVQ, PREVISÕES E RESULTADOS.....</b>	<b>73</b>
4.1 Metodologia.....	73
4.2 Resultados de referência para comparação.....	77
4.3 Procedimentos Realizados para Treinamento da Rede LVQ .....	78
4.4 Previsões e Resultados obtidos .....	86
4.5 Análise dos Resultados .....	132
4.6 Conclusões dos resultados obtidos.....	140
<b>5 CONCLUSÕES E TRABALHOS FUTUROS.....</b>	<b>143</b>
5.1 Conclusões.....	143
5.2 Trabalhos Futuros .....	145
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>147</b>





## LISTA DE FIGURAS

2.1 - Ilustração de um neurônio biológico.....	30
2.2 - Esquema de um neurônio biológico .....	30
2.3 - Esquema simplificado de um neurônio artificial .....	30
2.4 - Esquema simplificado de uma rede neural artificial .....	30
2.5 - Rede neural artificial simples .....	31
2.6 - Esquema de uma rede neural artificial.....	31
2.7 - Esquema do Mapa de Kohonen.....	35
2.8 - Fluxograma genérico para treinamento de rede SOM .....	37
2.9 - Rede LVQ .....	38
2.10 - Exemplo da arquitetura de uma rede LVQ.....	40
2.11 - Fluxograma para algoritmo genérico de uma rede LVQ.....	43
2.12 - Níveis de altura do Rio Paraguai, em Ladário (MS), de 01/01/1900 até 31/03/2007 .....	46
3.1 - Ladário - MS.....	55
3.2 - Bacia do Alto Paraguai.....	56
3.3 - Classificação de ano de cheia ou de seca .....	57
3.4 - Régua de Medição do nível do Rio Paraguai em Ladário/MS.....	58
3.5 - Detalhe da Régua de Medição de Ladário/MS.....	58
3.6 - Perfil longitudinal do Rio Paraguai .....	59
3.7 - Perfil da Profundidade do Rio Paraguai - MCT .....	59
3.8 - Exemplo da variação do nível do Rio Paraguai.....	63
3.9 - Diferenças de representação das médias de medidas .....	64
4.1 - Metodologia para previsão utilizando rede LVQ.....	75
4.2 - Resultados da previsão da rede lvq_a3_d12_m3030_F20 .....	108
4.3 - Resultados da previsão da rede lvq_a3_d8_m3030_F20 .....	108
4.4 - Resultados da previsão da rede lvq_a3_d4_m3030_F20 .....	109
4.5 - Resultados da previsão da rede lvq_a3_d4_m3030_F100 .....	111
4.6 - Resultados da previsão da rede lvq_a3_d8_m3030_F100 .....	112
4.7 - Resultados da previsão da rede lvq_a3_d8_m100100_F100 .....	114
4.8 - Resultados da previsão da rede lvq_a3_d12_m100100_F100 .....	115
4.9 - Rede LVQ 100x100 neurônios, de 12 pesos por neurônio e 100 padrões discretizados e sua legenda de cores por soma dos valores peso de cada neurônio .....	118
4.10 - Rede LVQ 100x100 neurônios, de 12 pesos por neurônio e 100 padrões discretizados e sua legenda de cores por padrão discretizado.....	119
4.11 - Neurônio do resultado da previsão para dia 29, da rede LVQ 100x100 de 12 pesos por neurônio e 100 padrões discretizados.....	120
4.12 - Rede LVQ 100x100 neurônios, de 8 pesos por neurônio e 100 padrões discretizados e sua legenda de cores por soma dos valores peso de cada neurônio .....	123
4.13 - Rede LVQ 100x100 neurônios, de 8 pesos por neurônio e 100 padrões discretizados e sua legenda de cores por padrão discretizado.....	124

4.14 - Neurônio do resultado para previsão do dia 29, da rede LVQ 100x100 de 8 pesos por neurônio e 100 padrões discretizados.....	124
4.15 - Rede LVQ 30x30 neurônios, 8 pesos por neurônio e 100 padrões discretizados e sua legenda de cores por soma dos valores peso de cada neurônio .....	126
4.16 - Rede LVQ 30x30 neurônios, 8 pesos por neurônio e 100 padrões discretizados e sua legenda de cores por padrão discretizado.....	127
4.17 - Neurônio do resultado para previsão do dia 29, da rede LVQ 30x30 de 8 pesos por neurônio e 100 padrões discretizados.....	127
4.18 - Previsão para out. e nov.2007 – rede LVQ a3_d8_m3030_F100 .....	127
4.19 - Previsão para out. e nov.2007 – rede LVQ a3_d8_m100100_F100 .....	129

## LISTA DE TABELAS

3.1 - Exemplo das informações tabuladas dos dados de Ladário/MS.....	60
3.2 - Exemplo de medidas de nível do rio Paraguai obtidos em Ladário e os valores médios calculados em janelas temporais com número de medidas igual a: 3, 5, 10 e 15 medidas .....	66
3.3 - Exemplo de conteúdo dos arquivos gerados na amostragem 4 médias .	67
3.4 - Exemplo de conteúdo dos arquivos gerados na amostragem 8 médias .	67
3.5 - 10 padrões .....	68
3.6 - 20 padrões .....	68
3.7 - 30 padrões .....	68
3.8 - Exemplos de arquivo com 4 médias por entrada gerado para utilizar no treinamento da rede LVQ.....	70
3.9 - Exemplos de arquivo com 8 médias por entrada gerado para utilizar no treinamento da rede LVQ.....	70
3.10 - Parâmetros utilizados no treinamento da rede LVQ.....	71
4.1 - Resultados descritos por Li et al(1998).....	78
4.2 - Períodos utilizados para comparação dos resultados.....	79
4.3 - Exemplo do conteúdo de um arquivo de rede LVQ (média de 3 medidas, 8 médias por entrada, rede de 10x10 neurônios, 10 padrões utilizados) ..	81
4.4 - Exemplo de arquivo para treinar rede LVQ .....	82
4.5 - Exemplo de arquivo gerado no treinamento da rede LVQ .....	82
4.6 - Exemplo do cálculo do desvio padrão.....	83
4.7 - Exemplo dos resultados do cálculo da qualidade da rede LVQ. ....	84
4.8 - Resultado com 0% erro, em quantidade de redes LVQ, dos 7 períodos. 87	
4.9 - Resultado com 0% erro, por percentual de redes LVQ, dos 7 períodos .	88
4.10 - Resultados por número de medidas por média, redes LVQ dos 7 períodos .....	90
4.11 - Resultados por número de pesos / neurônio, redes LVQ dos 7 períodos .....	91
4.12 - Resultados por número de padrões utilizados nas redes LVQ, dos 7 períodos. ....	92
4.13 - Resultados por tamanho das redes LVQ (número de neurônios) , redes LVQ dos 7 períodos. ....	93
4.14 - Resumo dos resultados de previsão realizados pela rede LVQ 30x30 neurônios, treinada pelos parâmetros: 12 entradas, 30 medidas para média de cada entrada, 10 padrões discretizados.....	94
4.15 - Resultado Erros 0%, por quantidade de redes LVQ, previsão de 1 passo.....	96
4.16 - Resultado Erros 0%, por percentual de redes LVQ, previsão 1 passo. 97	
4.17 - Resultados por número de medidas por média, para previsão de 1 passo adiante.....	98
4.18 - Resultados por número de entradas, para previsão de 1 passo adiante. ....	99

4.19 - Resultados por número de padrões utilizados nas redes LVQ, para previsão de 1 passo adiante .....	99
4.20 - Resultados por tamanho das redes LVQ, para previsão de 1 passo adiante. ....	100
4.21 - Resultado com 0% erros, por quantidade de redes LVQ, para previsão de 21 passos adiante .....	101
4.22 - Resultado com 0% erros, por percentual de redes LVQ, para previsão de 21 passos adiante .....	102
4.23 - Resultados por número de medidas por média, para previsão de 21 passos adiante, consecutivos .....	103
4.24 - Resultados por número de entradas, para previsão de 21 passos adiante, consecutivos .....	103
4.25 - Resultados por número de padrões utilizados nas redes LVQ, para previsão de 21 passos adiante, consecutivos .....	104
4.26 - Resultados por tamanho das redes LVQ, para previsão de 21 passos adiante, consecutivos .....	104
4.27 - Exemplo de arquivo com referências para previsão, com 8 medidas. ....	105
4.28 - Resultados da previsão da rede lvq_a3_d12_m3030_F20 .....	108
4.29 - Resultados da previsão da rede lvq_a3_d8_m3030_F20 .....	109
4.30 - Resultados da previsão da rede lvq_a3_d4_m3030_F20 .....	110
4.31 - Resultados da previsão da rede lvq_a3_d4_m3030_F100 .....	112
4.32 - Resultados da previsão da rede lvq_a3_d8_m3030_F100 .....	113
4.33 - Resultados da previsão da rede lvq_a3_d8_m100100_F100 .....	115
4.34 - Resultados da previsão da rede lvq_a3_d12_m100100_F100 .....	116
4.35 - Resultados da previsão da rede lvq_a3_d12_m100100_F100, com os neurônios e suas localizações. ....	118
4.36 - Resultados da previsão da rede lvq_a3_d8_m100100_F100, com os neurônios e suas localizações. ....	122
4.37 - Resultados da previsão da rede lvq_a3_d8_m3030_F100, com os neurônios e suas localizações. ....	125
4.38 - Previsões de out. e nov./2007 – rede LVQ a3_d8_m3030_F100.....	129
4.39 - Previsões de out. e nov./2007 – rede LVQ a3_d8_m100100_F100....	131

## LISTA DE SIGLAS E ABREVIATURAS

ANA	Agencia Nacional de Águas
ANTAQ	Agência Nacional de Transportes Aquaviários
ART	Adaptive Resonance Theory – Teoria de Ressonância Adaptativa
ARTMAP	Adaptive Resonance Theory Map – Mapa de Rede ART
BAP	Bacia do Alto Paraguai
BPNN	Back Propagation Neural Network – Rede Neural com Retro Propagação
FAETEC	Faculdade de Tecnologia Thereza Porto Marques
IBGE	Instituto Brasileiro de Geografia e Estatística
INPE	Instituto Nacional de Pesquisas Espaciais
LVQ	Learning Vector Quantization
MCT	Ministério da Ciência e Tecnologia
MLP	Multi Layer Perceptron – Perceptron de Multi Camadas
MT	Ministério dos Transportes
PCD	Plataforma de Coleta de Dados
RNA	Rede Neural Artificial
SNNS	Stuttgart Neural Network Simulator – Simulador de Rede Neural Stuttgart
SOM	Self Organizing Map – Mapa Auto Organizado



## 1 INTRODUÇÃO

Uma das riquezas naturais que o Brasil possui é o imenso patrimônio hídrico, lembrado normalmente pelo Rio Amazonas e sua bacia de alcance quase continental. Não podemos esquecer, entretanto, as outras bacias fluviais formadas por malhas de rios, riachos e cursos de água que beneficiam a população brasileira como fonte para consumo humano, suprimento para agricultura, indústria e até para transportes e integração regional. Esse reconhecimento da importância da água percebe-se através da criação da Agência Nacional de Águas (ANA) (Lei nº. 9984 de 17/07/2000), que tem como missão regular o uso correto das águas dos rios e lagos, de forma sustentada, para evitar poluição e desperdício. O Ministério dos Transportes também dá muita importância às águas, como meio de transporte, através de uma entidade vinculada, a Agência Nacional de Transportes Aquaviários (ANTAQ), que trata dos transportes hidroviários marítimos e fluviais.

O Transporte Hidroviário permite escoar grande volume de carga com custos competitivos, isto é, a relação custo por tonelada transportada é a mais baixa entre outras formas de transporte (rodoviário, ferroviário e aéreo), além de ser baixo também o custo de implantação por km de transporte, principalmente o fluvial, que aproveita os caminhos naturais dos cursos de água como rios e lagos.

Portanto, conhecer o comportamento de um rio, seus períodos de cheia e de seca (ou vazante), é importante para o transporte de cargas e passageiros por hidrovias, e traz ainda benefícios em muitos aspectos, pois ajuda no planejamento e manuseio de recursos sociais, incrementa a economia e beneficia a população. Permite, por exemplo, manusear rebanhos, salvar equipamentos e deslocar famílias que estejam em áreas sob risco de enchentes ou preparar-se para períodos de estiagem mais significativos, muito



importante para o transporte fluvial por permitir prevenir-se dos baixos níveis de água do rio programando os trajetos e as cargas.

O controle e conhecimento total de um rio são impossíveis, porém a previsão do comportamento dele é possível, tornando-se uma informação muito importante para a definição de políticas de utilização e aproveitamento desse rio, podendo gerar impactos sócio-econômicos. Pelo termo “previsão do comportamento do rio”, entende-se a estimativa do nível ou da vazão que o rio poderá atingir em um determinado período de tempo. Atualmente existem vários métodos e processos que permitem realizar a previsão da vazão de um rio. Vazão significa o volume de água que passa em uma seção do rio em certo período de tempo. Se essa seção for constante é possível saber a área e, portanto, o nível que a água poderá atingir.

Segundo a Tabela de Modelos Utilizados no Gerenciamento dos Recursos Hídricos de Tucci (1998), o modelo precipitação-vazão permite a previsão da vazão de uma bacia a partir da precipitação; e o modelo vazão-vazão permite a previsão de cheias pelo cálculo da vazão de uma seção a partir de um ponto a montante. Ainda segundo a Tabela Relação entre as Áreas e os Modelos Tucci (1998), o modelo vazão-vazão relaciona-se com área de Calamidades e o modelo precipitação-vazão relaciona-se com a área de Transportes. Esta dissertação trata de um modelo de previsão de nível do rio Paraguai, utilizando dados de uma série histórica com mais de 100 anos de medidas diárias.

Na literatura existem vários títulos sobre o tema “comportamento de um rio e previsão da sua vazão”. Muller e Fill (2003), Kisi (2003), Atiya et al (1999), Li et al (1998) e Phien, Danh & Gupta (1999), por exemplo, utilizaram Redes Neurais Artificiais (RNA), enquanto que Nordemann (1998a, 1998b) e Olivo (2004) utilizaram modelos estatísticos.

Muller e Fill (2003) utilizaram RNA, de tipo MLP com *Backpropagation*, em problemas de vazão de um trecho do rio Iguaçu, entre Fluviópolis e União da

Vitória, em sete eventos ocorridos nos anos de 1982, 1983, 1987, 1990, 1992, 1993 e 1995, com o tempo de ocorrência estimado entre 5 e 100 anos. Essa rede foi composta por 2 neurônios na camada de entrada, 3 camadas intermediárias com 8 neurônios por camada e a camada de saída com um neurônio, além de experimentarem outras geometrias de rede de neurônios. As conclusões foram que utilizando os eventos de 1982, 1983, 1987 e 1990 para calibração e o restante para a verificação, os resultados foram satisfatórios e animadores, com erros médios obtidos na faixa de 3% a 8%. Por fim, concluíram também que, quanto maior e mais heterogêneo for o conjunto de eventos para calibração, melhor seriam os resultados obtidos.

Kisi (2003) discorreu em seu artigo sobre comparações numéricas e gráficas entre RNA e modelos de auto-regressão. Relatou sobre os valores de medidas de previsão de vazão de rios utilizando medidas contínuas de escoamento do rio obtidas em períodos de 7 e 4 anos dos rios Blackwater e Gila, nos E.U.A, e em um período de 2 anos de medidas de escoamento do Canal de Filyos da Turquia. Demonstrou pelos resultados obtidos que a utilização de RNA tende a gerar resultados melhores que algoritmo de auto-regressão.

Atiya et al (1999) utilizaram RNA para o problema da previsão do escoamento do rio Nilo, no Egito. Séries Temporais foram utilizadas para verificar o desempenho perante outros métodos de previsão através de RNA. Fizeram diversas comparações entre métodos diferentes para pré-processamento de entradas e resultados; entre tipos de técnicas utilizadas para resolver problemas de previsão de períodos mais distantes e entre métodos de previsão para um horizonte muito mais distante, com possibilidade de decompor o problema em partes menores de previsão de  $k$  períodos adiante.

Phien, Danh & Gupta (1999) utilizaram RNA com *Backpropagation* em seu artigo em que descreveram modelos utilizados para previsão de vazão em duas bacias (Da Nhim e La Nga) situadas nos planaltos do Vietnã em

comparação a modelos de tanque. Nas conclusões obtidas indicaram que modelos de RNA com *Backpropagation* são capazes de prever a vazão diária de um rio utilizando informações diárias de evaporação e de chuvas. E quanto maior for a área utilizada para obter as informações, mais preciso será a previsão. Também observaram que valores históricos de vazão são importantes, como nos modelos de regressão, e por fim, que a contribuição do item evaporação para ter uma previsão mais precisa é muito pequena.

Li et al (1998) utilizaram também RNA com *Backpropagation* implementado no aplicativo Stuttgart Neural Network Simulator (SNNS) desenvolvido pelo Instituto para Sistemas Paralelos e Distribuídos de Alta Performance da Universidade de Stuttgart. Conseguiram obter previsões de nível do rio Paraguai com até 4 meses de antecedência e até 17% de erros, utilizando 274 valores médios mensais de níveis do rio Paraguai.

Nordemann (1995a) e (1995b) utilizou métodos matemáticos como Análise de Fourier e método de regressão iterativa para obter séries temporais. Informou que o regime do Rio Paraguai era complexo e não estacionário, mesmo assim, obteve resultados proveitosos que mostraram diversas periodicidades que podiam, de certa forma, estarem relacionados com fenômenos geofísicos globais.

Olivo (2004) em sua tese utilizou técnicas como regressão múltipla por mínimos quadrados e modelos auto-regressivos para propor um sistema de alarme-resposta em tempo real. Esse alarme permite estimativas confiáveis para que autoridades competentes tomem decisões em relação às cheias de um rio em uma seção de interesse.

A maioria destes artigos referem-se a previsões de vazão de um rio, ou seja, o volume de um rio em função do tempo de modelos hidrológicos do tipo precipitação-vazão. Para isso necessitam informações como dados de

precipitação, taxas de evaporação, tipo de solo e porosidade do solo, lençol freático, declividade, cobertura vegetal da região, mata ciliar, velocidade da água, seção de um trecho do rio, entre muitas outras variáveis. Isso permite uma previsão mais completa e menos sujeita a variações, porém a um custo maior, pois há a necessidade de conhecer essas variáveis, muitas vezes complexas de se obter e de fusão que exige poder computacional elevado dependendo da quantidade disponível. Além disso, essas variáveis podem conter ruídos, falhas ou mesmo imprecisões e/ou inconsistências dada à diversidade de fontes de informações que os originam.

Portanto a possibilidade de se prever o nível de um rio, sem a necessidade de muitas informações além da observação da altura do rio, significa uma facilidade da metodologia proposta nesta dissertação, em relação a outros sistemas.

O objetivo deste trabalho é conceber um método simples e de fácil calibração para a previsão do nível do rio Paraguai, através do desenvolvimento de uma modelo baseado em uma RNA *Learning Vector Quantization* (LVQ) utilizando séries históricas de nível do rio Paraguai.

Um outro objetivo é tentar melhorar os resultados apresentados por Li et al (1998) utilizando outro método de determinação e previsão e que possa introduzir maior precisão e ou facilidade de implementação.

O LVQ é uma RNA de aprendizagem supervisionada e funciona como classificador de padrões que podem existir nos dados utilizados no treinamento da rede. Como a aprendizagem é supervisionada, a rede é treinada com auxílio externo até atingir resultados satisfatórios de classificação de padrões.

Os padrões, neste trabalho, serão os próprios níveis de altura que o rio Paraguai pode atingir, caracterizados através de intervalos. A previsão do nível será feita com base em medidas observadas em uma janela de tempo em dias. Essa é uma característica importante da metodologia que propõe a estimativa do nível do rio a partir do seu comportamento, oriundo da medição do nível em uma seqüência de dias anteriores, podendo ser em uma periodicidade de 2 dias, 3 dias, semanal, etc., dependendo da disponibilidade de dados.

As medidas podem ser realizadas visualmente ou por aparelhos automatizados, através de uma rede de observação ou de apenas uma estação de medição em uma localidade, exigindo poucos recursos humano e tecnológico especializados.

Este trabalho está estruturado da seguinte forma: a Introdução é apresentada no Capítulo 1. O Capítulo 2 apresenta a fundamentação teórica discorrendo sobre Redes Neurais Artificiais e sobre Análise de Séries Temporais. O Capítulo 3 é dedicado às informações sobre os dados utilizados e principalmente sobre a metodologia empregada na preparação dos dados. No Capítulo 4 os resultados das simulações e processamentos são apresentados. As conclusões obtidas e sugestões para trabalhos futuros são apresentadas no Capítulo 5. O documento termina com a apresentação das Referências Bibliográficas e os Anexos.

## **2 FUNDAMENTOS TEÓRICOS**

Neste Capítulo são descritos os fundamentos teóricos que norteiam esta dissertação de Mestrado. Inicialmente são apresentados conceitos sobre redes neurais artificiais, seguidas de breves discussões sobre alguns tipos de redes neurais, e enfatizando a rede LVQ. Em seguida é apresentada uma breve discussão sobre séries temporais. O Capítulo finaliza com alguns exemplos de utilização de redes neurais artificiais em hidrologia encontrados na literatura recente.

### **2.1 Rede Neural Artificial**

Segundo Fausett (1994), uma Rede Neural Artificial (RNA) é um sistema de processamento de informações que tem algumas características comuns com sistemas neurológicos biológicos. Os modelos de redes neurais foram desenvolvidos como generalizações de modelos matemáticos de cognição humana ou tomando como base um modelo computacional do neurônio biológico. As redes neurais artificiais, assim como existem, são sistemas computacionais que podem simular o funcionamento ou o modo como o cérebro humano processa as informações.

O cérebro é composto por bilhões de células nervosas (neurônios) que possuem extensões (axônios) com ramificações (dendritos) nas pontas e realizam conexões (conexões sinápticas) com outros neurônios. Estas conexões ligam um neurônio biológico a outros, formando uma rede com milhões de possibilidades de caminhos para as informações, em processamento, fluírem no nosso cérebro. A forma como são transmitidos os sinais elétricos entre cada conexão não será explicado neste trabalho, assumindo apenas que existe essa transmissão.

Fazendo uma analogia superficial, o equivalente simplificado de um neurônio biológico (Figura 2.1 e Figura 2.2) é o neurônio artificial (Figura 2.3), que é a unidade de processamento básica de uma Rede Neural Artificial.

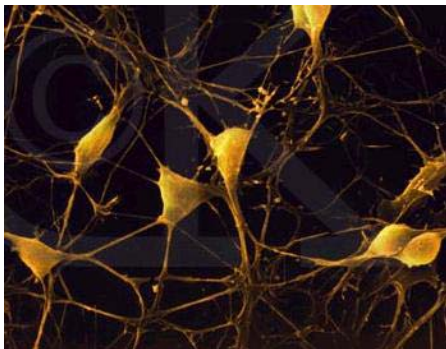


Figura 2.1 - Ilustração de um neurônio biológico

Fonte: LNCC, 2007

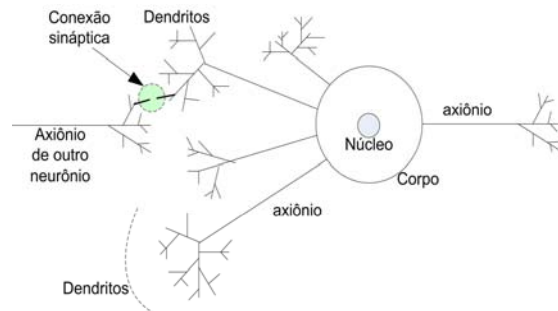


Figura 2.2 - Esquema de um neurônio biológico

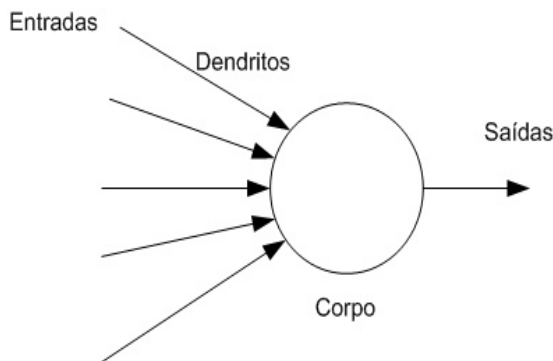


Figura 2.3 - Esquema simplificado de um neurônio artificial

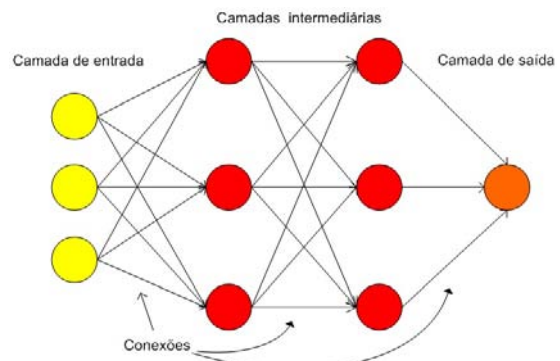


Figura 2.4 - Esquema simplificado de uma rede neural artificial

Em um processamento normal e simplificado de uma rede neural artificial, um conjunto de informações é fornecido como entrada em uma camada de neurônios artificiais. Cada entrada é repassada a outras camadas internas ou intermediárias, onde é realizado o processamento (Figura 2.3). Este processamento consiste em multiplicar o valor de cada entrada pelos valores atribuídos a cada conexão sináptica do neurônio artificial, valores esses chamados de pesos sinápticos (Figura 2.5). Os resultados das multiplicações são somados gerando um único valor, que é submetido a uma função que definirá o nível de ativação do neurônio. O resultado da ativação é passado a

outros neurônios artificiais (Figura 2.4). Uma RNA, então, consiste basicamente de conjuntos de nós ou neurônios artificiais organizados em camadas existindo: a camada de entrada, a camada de neurônios de saída (que fornece os resultados do processamento dos sinais pela rede) e um número de camadas intermediárias de neurônios entre a camada de entrada e a camada de saída.

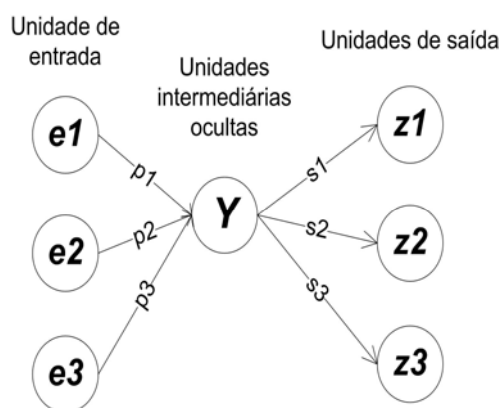


Figura 2.5 - Rede neural artificial simples

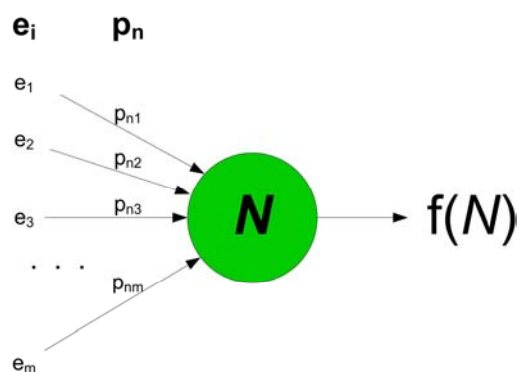


Figura 2.6 - Esquema de uma rede neural artificial

O funcionamento de uma RNA ocorre em 2 fases: a fase de treinamento dos neurônios e a fase de aplicação, conforme Fausett (1994). Na fase de treinamento a RNA é exposta a dados conhecidos na camada de entrada. Dependendo do tipo de RNA, também são apresentados dados de saída também conhecidos. No processamento dos dados pela RNA, seus pesos sinápticos são alterados para adequarem-se aos dados apresentados, de acordo com o objetivo de uso da RNA. Uma vez treinada, a RNA pode ser usada diretamente na aplicação, em que dados da mesma natureza daqueles usados no treinamento (mas não os mesmos), são apresentados a RNA que deve apresentar desempenho semelhante ao obtido durante a fase de treinamento. Este é o tipo de treinamento que é mais adequado para o método de previsão proposto nesta dissertação.



## **2.2 Formas de Aprendizado**

As formas (modos) mais usuais de treinamento (aprendizado) de uma rede neural são: supervisionada e não supervisionada. No modo supervisionado a rede neural artificial é alimentada, na entrada e na saída, por valores conhecidos. As discrepâncias entre os valores produzidos pela rede neural e aqueles fornecidos como referências, são utilizadas para ajustar os pesos da rede neural de forma que os resultados na saída sejam os mais próximos possíveis dos dados esperados e conhecidos previamente. Assim, ao encerrar esse processo, a rede neural estará treinada para ser utilizada com informações desconhecidas, mas da mesma natureza daqueles usados no treinamento.

No modo não supervisionado o processo de aprendizado utiliza apenas dados de entrada. O aprendizado ocorre pela adaptação dos pesos que buscam características estatísticas nos dados apresentados.

## **2.3 Tipos de RNA**

Existem diferentes formas de arranjos de neurônios que levam as diferentes modelos de redes neurais. A seguir são apresentados alguns tipos, com ênfase no modelo de rede neural utilizada nesta dissertação.

### **2.3.1 Perceptron de Múltiplas Camadas (MLP - Multi Layer Perceptron)**

A MLP é a RNA que possui os neurônios distribuídos em camadas. Assim, além da camada de entrada, a MLP possui a camada de neurônios intermediários ocultos (podendo se uma ou mais camadas) e camada de neurônios de saída. Os neurônios de entrada recebem os estímulos do meio externo, os neurônios de saída se comunicam com o exterior disponibilizando os resultados e os neurônios intermediários realizam transformações sobre os dados apresentados na entrada da rede.

Uma MLP tem alimentação “*feedforward*” em que o fluxo do processamento dos dados ocorre desde a camada de entrada, prosseguindo pelas camadas intermediárias, até a camada de saída. No ajuste dos pesos de uma MLP, o fluxo é retro-alimentado, com o cálculo dos ajustes dos pesos que depende do erro na saída da rede e do cálculo dos gradientes locais dos neurônios em cada camada. O algoritmo mais tradicional de treinamento de uma MLP é denominado retro-propagação do erro (*error backpropagation*).

### **2.3.2 ART - Adaptive Resonance Theory ou Teoria de Ressonância Adaptativa**

A rede neural ART foi desenvolvida por Stephen Grossberg em 1976 para modelar a arquitetura auto-organizada baseada em comportamentos biológicos para reconhecimento de padrões. A ART caracteriza-se pela capacidade de criar novas classes se nenhuma classe existente ressoar com o padrão de entrada apresentado à rede. Isto significa a habilidade de se manter o equilíbrio entre as propriedades de plasticidade e de estabilidade, preservando o aprendizado ou o conhecimento anterior. A plasticidade é o estado em que os parâmetros internos da rede podem ser modificados permitindo a capacidade de criar novas classes de reconhecimento. A estabilidade define a capacidade de adaptar ou agrupar os padrões de entrada similares na mesma classe de reconhecimento Beale e Jackson (1990).

As versões mais conhecidas do modelo ART são: ART-1, que permite aprender a classificar padrões de entrada binários em ordem arbitrária; ART-2, que aprende a classificar padrões de entrada analógicos ou binários; ART-3, que permite busca paralela, ou teste de hipóteses, em códigos com reconhecimento distribuído; ARTMAP, que estabelece associação entre padrões binários de entrada com binários de saída; Fuzzy ART, que é uma generalização da rede ART1 e permite reconhecer padrões analógicos e binários de forma mais simples que a ART2 e Fuzzy ARTMAP que é uma adaptação da rede ARTMAP

para utilizar padrões analógicos e binários tanto na entrada como na saída, conforme Balestrassi (2000).

A arquitetura de uma rede ART é composta por 2 camadas de neurônios: uma camada de entrada e comparação, e uma outra camada de saída e reconhecimento. As camadas são interconectadas, com uso de *feedback* da camada de saída para a de entrada, e também entre os neurônios laterais na camada de saída. Existem ainda, os pesos dos vetores *feedforward* da camada de entrada para a camada de saída. Há também para cada camada um controle de fluxo de dados entre as camadas para cada etapa do ciclo de operação, realizado por um controle lógico através de sinais. Entre as camadas de entrada e saída, existe também um mecanismo de *reset* e que tem a função principal de comparar as entradas por um mecanismo de verificação de similaridades e decidir a criação de um novo padrão para a entrada Silva, 2005.

### **2.3.3 SOM – Self Organizing Map**

SOM, também conhecido como Mapa de Kohonen, é uma RNA de classificação de dados e tem aprendizado não-supervisionado. O SOM foi desenvolvido por Teuvo Kohonen na década de 1980. Os estados internos da rede SOM se modificam para modelar as características encontradas nos dados de treinamento Beale e Jackson (1990). O objetivo de uma rede SOM é transformar um padrão de entrada de dimensão qualquer em um mapa discreto uni ou bidimensional, de forma adaptável e topologicamente ordenada. Silva, 2005. Geralmente, uma implementação de um SOM consiste em um arranjo de neurônios em duas dimensões (largura e altura) compondo um mapa. Os neurônios são arranjados ou organizados em um único plano. Todas as entradas são conectadas aos neurônios da rede. Cada neurônio da rede corresponde a uma saída. A realimentação das informações é limitada às conexões laterais dos neurônios vizinhos. A Figura 2.7 apresenta um exemplo de esquema do Mapa de Kohonen com as saídas em duas dimensões formadas por 4 neurônios de largura e 4 neurônios de altura, totalizando 16

neurônios. As entradas são representadas na figura pelos conjuntos verticais, em que cada entrada é composta por 6 elementos.

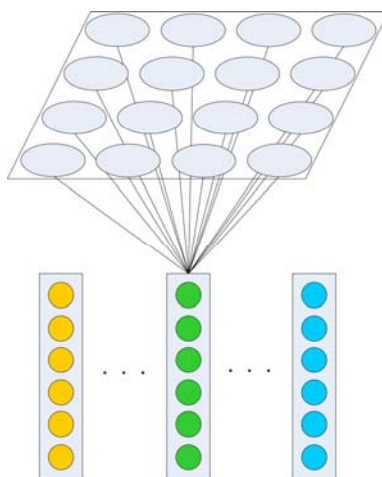


Figura 2.7 - Esquema do Mapa de Kohonen

O funcionamento do SOM busca encontrar os neurônios que se correlacionam ao máximo com as entradas de treinamento, melhorando a similaridade desses neurônios; e enquadrar as entradas na vizinhança mais próxima desses neurônios. Esse processo exige uma fase de competição em que procura-se o neurônio que mais se aproxima do padrão da entrada. Este neurônio é considerado como “vencedor” e os valores de seu vetor de pesos são alterados para melhorar a sensibilidade ao padrão de entrada, assim como os de seus vizinhos dentro de um raio. Como resultado, o algoritmo de aprendizado organiza os neurônios em vizinhanças locais que funcionam como classificadores dos dados de entrada e gera um mapa topográfico. Os conjuntos de neurônios organizados aleatoriamente na grade indicam a existência de áreas representativas no mapa auto-organizado.

Kohonen introduziu a idéia de vizinhança topológica, como sendo uma área em que seus limites ou fronteiras são alterados dinamicamente indicando a influência das modificações dos pesos dos nós vencedores. Inicialmente cada neurônio da SOM tem uma vizinhança que pode corresponder a toda rede. À

medida que ocorre o treinamento, o tamanho da vizinhança é reduzido por um valor predefinido até atingir um valor pré-estabelecido.

Portanto, o algoritmo de treinamento da SOM gera agrupamentos dos dados de treinamento, e tanto a seqüência quanto o número de agrupamentos dependem do modo como os dados de treinamento são apresentados à SOM. A atribuição de rótulos para identificar os agrupamentos obtidos no treinamento, pode classificar dados desconhecidos inseridos na SOM. A rotulação da classificação deve ser feita manualmente após o treinamento.

No treinamento, os ajustes dos pesos são realizados de acordo com uma taxa de adaptação ou de aprendizagem. Essa taxa de adaptação, assim como o raio de vizinhança são reduzidos a cada ciclo, seguindo alguma função de decremento.

O fim do treinamento ocorre quando é atingido algum parâmetro de finalização previamente estabelecido, que pode ser, por exemplo, o número de ciclos de processamento, o tempo de execução, o tamanho do raio de vizinhança ou o valor da taxa de aprendizagem.

A Figura 2.8 apresenta um exemplo simples de algoritmo para treinamento de uma SOM baseado no algoritmo de Fausett (1994). Destaca-se o cálculo da Distância Euclidiana para se obter o neurônio vencedor no processo competitivo e também o cálculo utilizado para atualizar os valores do vetor peso desse neurônio e dos que pertencem à área de influência da sua vizinhança.

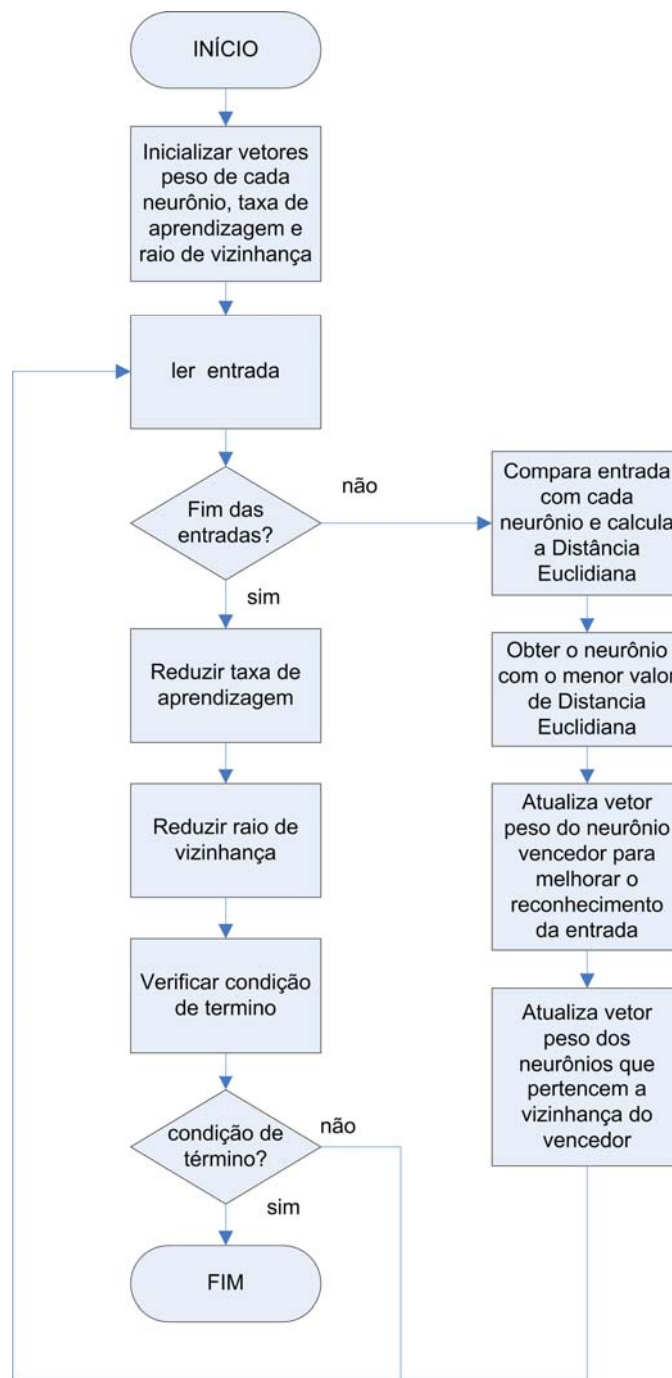


Figura 2.8 - Fluxograma genérico para treinamento de rede SOM

### 2.3.4 LVQ – Learning Vector Quantization

A Learning Vector Quantization (LVQ) Kohonen (1989a, 1990a) é um método de classificação de padrões, segundo Fausett (1994), que utiliza uma rede de camada única de neurônios artificiais posicionadas em duas ou mais dimensões para classificar as informações apresentadas em padrões ou classes distintas que existam. A LVQ utiliza aprendizagem do tipo supervisionado utilizando um conjunto de dados de entrada com padrões conhecidos Silva, 2005.

Na rede LVQ cada vetor de entrada é apresentado a todos os neurônios. Um processamento de comparação é realizado em cada neurônio gerando um resultado. Os resultados de todos os neurônios são apresentados em uma camada única destes neurônios, em que cada neurônio representa uma classe ou padrão das informações utilizadas no treinamento.

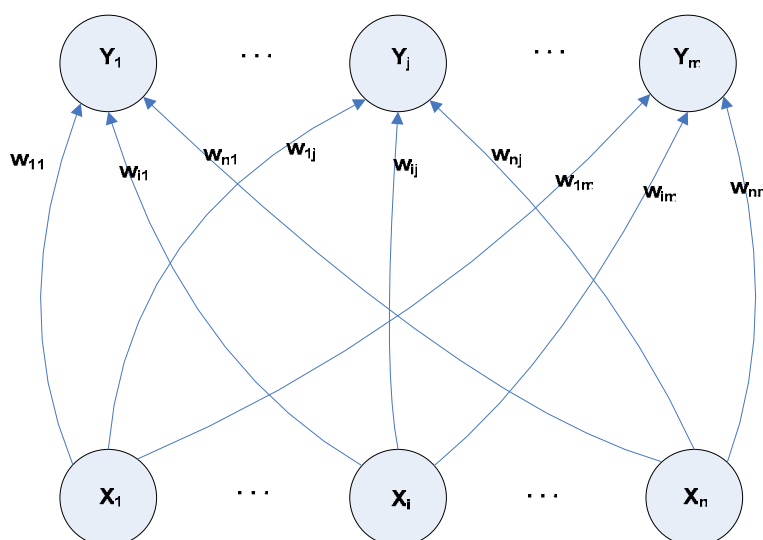


Figura 2.9 - Rede LVQ

Fonte: Fausett (1994 pág.188)

A Figura 2.9 apresenta a arquitetura de uma rede LVQ utilizada por Fausett (1994) para exemplificar seu funcionamento. O algoritmo procura os neurônios mais semelhantes a cada informação de entrada através de um processo de

competição. Na Figura 2.9, os elementos  $X_1, X_i, X_n$  representam as informações de entrada (vetores de treinamento),  $Y_1, Y_j, Y_m$  representam os neurônios artificiais (unidades de saída) e  $W_{11}, W_{i1}, W_{n1}, \dots, W_{nm}$  representam os elementos dos vetores-peso ( $W$ ) de cada um dos neurônios artificiais da rede LVQ.

Mais detalhadamente, cada neurônio artificial contém um conjunto de valores que forma um vetor ou lista de valores, chamado de vetor-peso. Isso possibilita que cada neurônio artificial possa reconhecer e representar os padrões encontrados nas informações processadas. Cada informação de entrada é composta por valores agrupados em vetores e o número de elementos de cada vetor de entrada define o número de elementos do vetor-peso de cada neurônio.

A Figura 2.10 mostra de outra forma a interpretação da Figura 2.9. Os elementos  $e_1, e_2, e_3$  representam as informações de entrada, cada um com seu vetor de valores ( $e_1 = \{e_{11}, e_{12}, \dots, e_{16}\}$ ,  $e_2 = \{e_{21}, e_{22}, \dots, e_{26}\}$  e  $e_3 = \{e_{31}, e_{32}, \dots, e_{36}\}$ ). Os elementos  $n_1, n_2, n_3, \dots, n_{16}$  representam os neurônios artificiais da rede LVQ com 4x4 neurônios. Os elementos  $n_{41}, n_{42}, \dots, n_{46}$  representam os valores-peso que compõem o vetor-peso do neurônio  $n_4$ . A Figura 2.10 informa também que a rede LVQ tem uma camada única de neurônios e duas dimensões, embora uma rede LVQ possa ter mais dimensões.



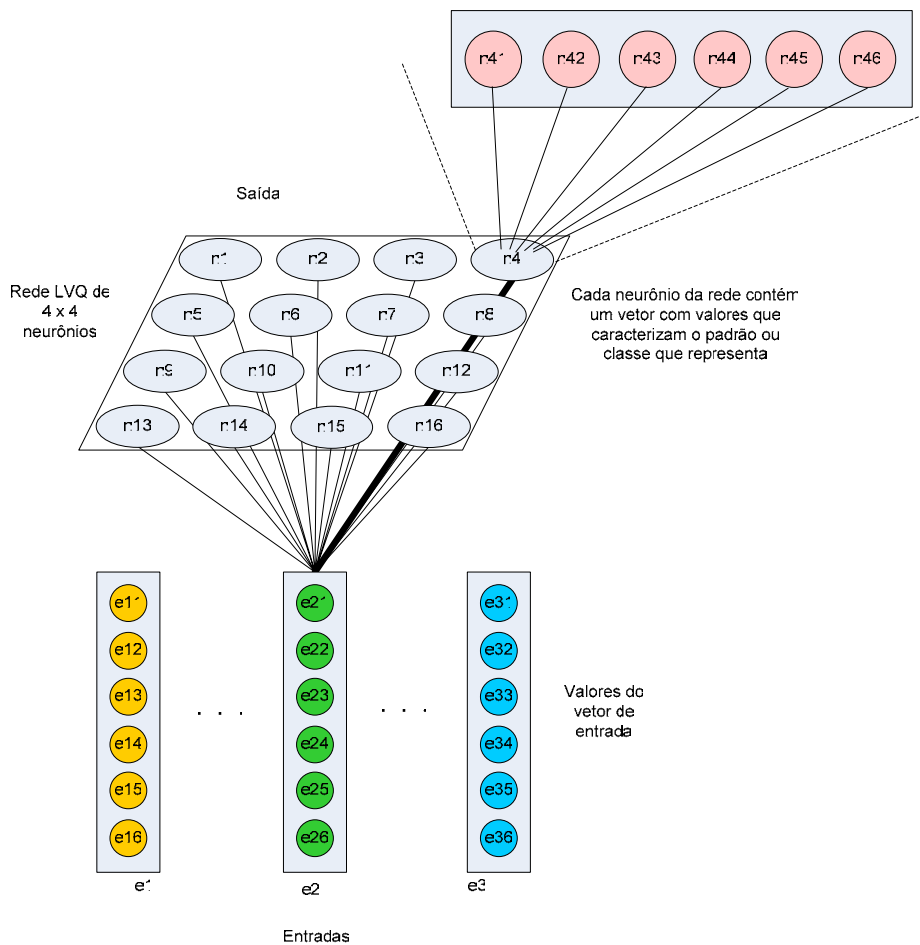


Figura 2.10 - Exemplo da arquitetura de uma rede LVQ

O reconhecimento do padrão de cada entrada realizado pelos neurônios é feito pela comparação de cada informação de entrada com cada neurônio da rede LVQ previamente ensinada na fase de treinamento.

No treinamento, os valores do vetor-peso são alterados à medida que as informações de treinamento são inseridas no processamento da rede LVQ. É um processo de comparação que verifica a semelhança entre o conjunto de valores de entrada e os valores do vetor-peso de cada neurônio. Em outras palavras, é um processo de competição em que se sobressai um neurônio vencedor, ou seja, o mais parecido com os valores de entrada. Lembramos que

cada conjunto de valores de entrada também forma um vetor (vetor-entrada), que tem o mesmo número de valores do vetor valores-peso.

Essa comparação utiliza o resultado do cálculo da distância euclidiana entre os valores do vetor de entrada e do vetor de pesos de cada neurônio. Assim, o neurônio que obtiver o menor valor calculado, será considerado o neurônio vencedor. Se esse neurônio vencedor é do mesmo padrão ou classe do vetor de entrada, os valores do seu vetor de pesos são alterados para efetuar uma comparação melhor com outros valores de entradas. Porém, se o padrão do vetor de entrada for diferente do padrão do neurônio vencedor, os valores do vetor de pesos do neurônio são alterados para diferenciá-lo mais dessa informação de entrada.

O processo anterior é realizado repetidamente por um número de ciclos e em cada ciclo a taxa de aprendizado é reduzida. Ao fim desse aprendizado ou treinamento a LVQ estará pronta para classificar outros dados.

Fausett (1994) também apresenta as evoluções da rede LVQ: LVQ2, LVQ2.1 Kohonen (1990a) e LVQ3 Kohonen (1990b). Na rede LVQ original somente o neurônio vencedor tinha seu vetor de pesos alterado. Na rede LVQ2, tanto o neurônio vencedor quanto o segundo neurônio mais parecido (vice) também são alterados, desde que satisfaçam critérios como: neurônio vencedor e vice sejam de padrões diferentes, padrão da informação de entrada seja o mesmo do vice; e a distancia euclidiana entre a entrada e o vencedor seja aproximadamente a mesma entre a entrada e o vice. Satisfeitos os critérios, o vetor de pesos do neurônio vencedor tem sua taxa de aprendizado diminuído enquanto o vetor de pesos do vice tem sua taxa de aprendizado aumentado.

A rede LVQ2.1, segundo Fausett (1994) e Kohonen (1990a), atualiza os dois neurônios com vetores de pesos mais semelhantes ao vetor de entrada. Para isso deve satisfazer condições como: um dos dois neurônios deve pertencer ao

mesmo padrão da informação de entrada e o outro neurônio não deve pertencer ao mesmo padrão da entrada; e a informação de entrada deve pertencer a uma janela ou intervalo de valores mínimo e máximo, calculados pela divisão da distância euclidiana entre a entrada e um dos neurônios semelhantes e a distância euclidiana entre a entrada e o outro neurônio semelhante. Satisfeitas as condições, o vetor de pesos do neurônio do mesmo padrão da entrada tem seus valores alterados proporcionalmente pela taxa de aprendizagem, enquanto que o outro neurônio tem os valores dos seus pesos reduzidos também proporcionalmente.

A LVQ3 possibilita que dois neurônios com vetores de pesos mais semelhantes ao vetor da informação de entrada realizem algum aprendizado desde que satisfaçam condições de pertencimento à janela de intervalo. Essa janela é calculada pela razão entre os valores de distâncias euclidianas do vetor de valores do vetor de entrada e do vetor de pesos de cada neurônio mais semelhante. Caso um neurônio pertença ao mesmo padrão da entrada e o outro neurônio não, a atualização dos vetores de pesos desses neurônios segue como realizado no LVQ2. Do contrário, o algoritmo da rede LVQ3 proporciona treinamento desde que a entrada e os neurônios mais semelhantes à entrada sejam do mesmo padrão, e a atualização dos valores do vetor de pesos é realizada em ambos os neurônios através de um novo índice de aprendizado calculado como um valor múltiplo (entre 0,1 e 0,5) do índice utilizado em situações de padrões diferentes dos neurônios semelhantes. A rede LVQ3 tem essa alteração no processo de treinamento para garantir que os valores do vetor de pesos continuem a classificação dos padrões e prevenir alterações indevidas.

A Figura 2.11 apresenta o fluxograma para um algoritmo genérico de uma rede LVQ, que segue a referência do algoritmo de rede LVQ de Fausett (1994), da página 188.

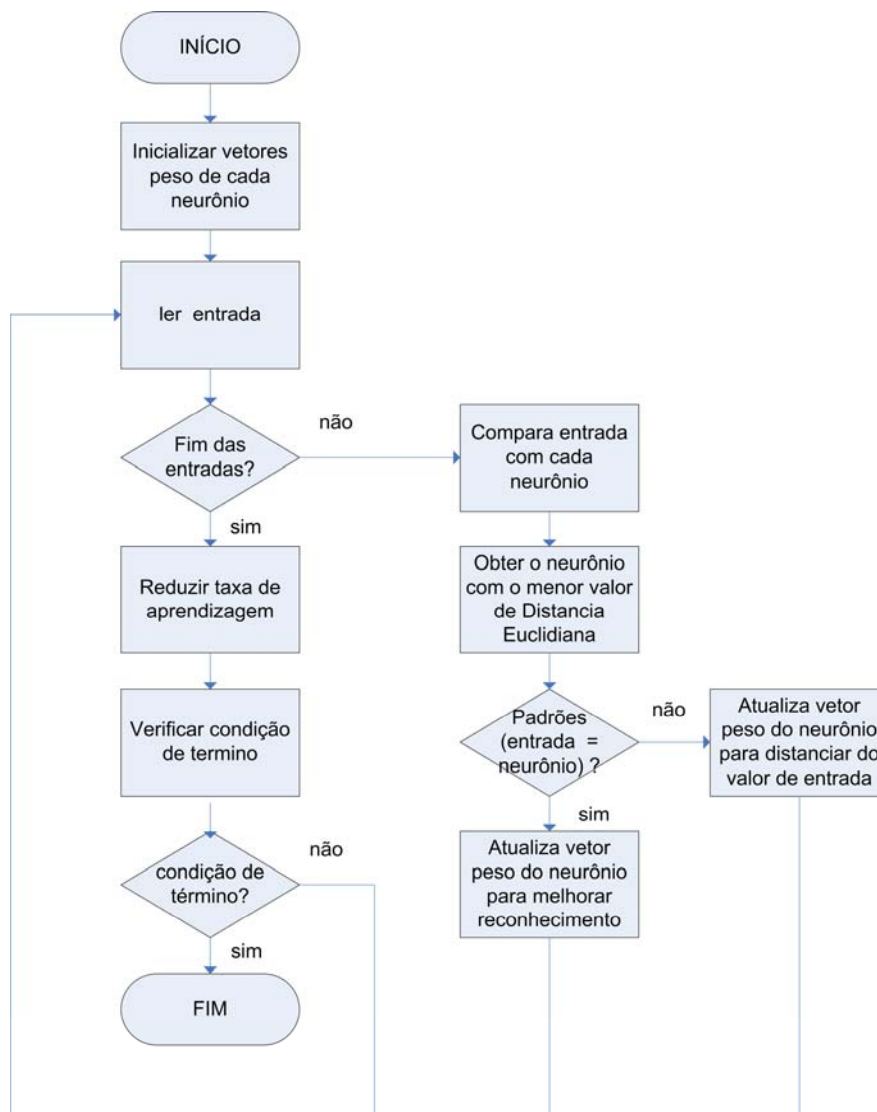


Figura 2.11 - Fluxograma para algoritmo genérico de uma rede LVQ

### 2.3.5 Comparação SOM X LVQ

As redes SOM e LVQ são parecidas quanto ao funcionamento, pois ambas utilizam processos competitivos para obter o neurônio vencedor e podem apresentar os resultados em um mapa bidimensional. Os neurônios da rede SOM a princípio não têm identificação sobre qual padrão representa, enquanto os neurônios da rede LVQ são rotulados e representam efetivamente algum

padrão encontrado nos dados de treinamento. Para o treinamento da rede SOM, o algoritmo utiliza treinamento não supervisionado enquanto que o algoritmo LVQ utiliza treinamento supervisionado. Essa diferença possibilita um treinamento mais eficaz à rede LVQ, através do “ajuste fino” nos valores dos vetores de pesos dos neurônios até alcançar os resultados desejados.

Uma outra diferença ocorre no processamento do neurônio vencedor. Na rede SOM, os valores do vetor do neurônio vencedor são alterados para melhorar a sensibilidade ao padrão das entradas, enquanto na rede LVQ, existe uma verificação do padrão do neurônio vencedor. Nessa verificação, se o padrão do neurônio for o mesmo da entrada, os valores do seu vetor são alterados para melhorar sua sensibilidade a esse padrão. Se forem de padrões diferentes, os valores são alterados para diferenciar melhor os padrões. E por fim, a rede SOM gera áreas de similaridade atualizando os vetores dos neurônios da vizinhança dos neurônios vencedores, enquanto na rede LVQ não existe esse processo de atualizar as vizinhanças. Podem ocorrer regiões próximas, mas será devido ao modo como as entradas são inseridas no treinamento da rede LVQ.

Portanto a escolha da rede LVQ neste trabalho de dissertação foi motivada pela maior possibilidade de acertos na classificação dos padrões das informações utilizadas. E foi utilizado o algoritmo da rede LVQ original, por ser de implementação simples e atende ao propósito de classificar os padrões de nível de rio, mesmo que suas variações apresentem-se como evoluções e melhorias.

## **2.4 Análise de Séries Temporais**

O conjunto de dados utilizado neste trabalho é composto por medidas de nível do rio Paraguai, anotadas diariamente desde 1900, constituindo uma série temporal extensa e confiável.

Uma série temporal é um conjunto de observações ordenadas no tempo Soares et al (1991). Uma análise de série temporal pode ser explicada como a observação de várias medidas obtidas em diferentes instantes de tempo e pode ser vista como uma realização de certo processo estocástico (Anteneodo (2004)) e, a partir dessas observações, tentar extrair informações que possam ser aproveitadas.

Segundo Migon (2004), uma série temporal é uma seqüência de observações ao longo do tempo, que podem ser espaçadas uniformemente e, se não o forem, podem ser utilizados artifícios como mudança na escala para contornar ausência de algumas observações, com atenção às escalas de medição. Em geral, os dados são obtidos ao longo do tempo de forma rotineira e contínua a intervalos supostamente constantes. Para análise da massa de dados, escolhe-se uma janela temporal, e dela tenta-se detectar padrões que possam se repetir no tempo, evidenciando um comportamento rotineiro que possa ser previsível.

Para Keogh e Pazzani (1999) as séries temporais correspondem a porções significativas dos dados armazenados em bancos de dados gerados por áreas de aplicação distintas como financeira, médica, engenharia, e até em banco de dados com informações sociológicas. Existem inúmeros testes estatísticos que podem ser aplicados sobre séries temporais, como as que determinam coeficientes de auto-correlação, tendências de medidas lineares, etc. A utilização dessas coleções de dados dependem das capacidades e das

habilidades humanas de visualizar os padrões embutidos e o formato desses dados graficamente.

Os métodos estatísticos de previsão de séries temporais, como a regressão múltipla por mínimos quadrados utilizada por Olivo (2004), baseiam-se na idéia de que as observações passadas da série contêm informações sobre o seu padrão de comportamento que podem ser usadas para fazer previsões futuras. A essência dos métodos de análise de séries temporais consiste em identificar o padrão da série utilizando modelos matemáticos e estatísticos, separando-o do ruído contido nas observações individuais, e utilizá-lo para previsões futuras da série Soares et al (1991). Ou seja, a análise das séries temporais tem como objetivo descrever e analisar o comportamento passado da série, visando à compreensão do comportamento da série e a consequente previsão de movimentos futuros. Fonseca et al (1988).

Um exemplo de série temporal é apresentado na Figura 2.12, que mostra os níveis diários do rio Paraguai, medidos entre o dia 1 de janeiro de 1900 até 31 de março de 2007.

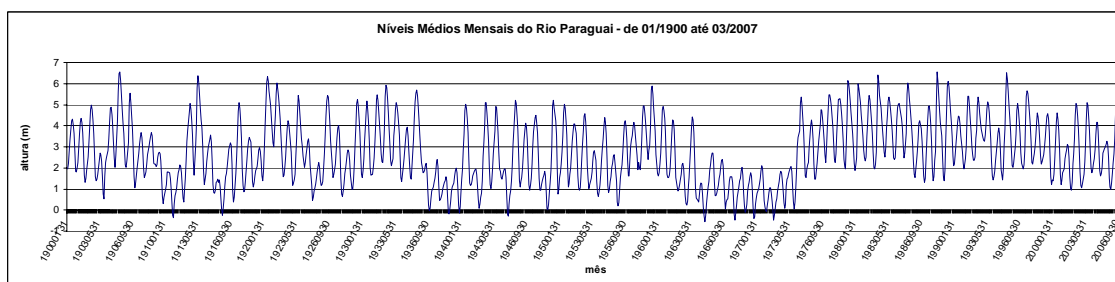


Figura 2.12 - Níveis de altura do Rio Paraguai, em Ladário (MS), de 01/01/1900 até 31/03/2007

Kazmier (1982) e Fonseca (1988), observando séries temporais dos dados de nível do rio Paraguai, identificam 4 influências ou componentes em uma análise de séries temporais:

- tendência secular (T): movimento regular nos valores da série temporal em um período de tempo muito longo;
- flutuações cíclicas (C): movimentos oscilatórios em torno da tendência de tempo muito longo;
- variações sazonais ou estacionais (S): movimentos para cima ou para baixo dos valores da série temporal, em relação à tendência secular, completados geralmente em períodos inferiores a um ano e repetição anual, sendo característicos movimentos semanais, quinzenais, mensais, bimestrais;
- variações irregulares ou aleatórias (I): são perturbações provocadas por fatores não regulares, erráticas e nunca atribuídas às influências cíclicas ou estacionais.

A análise de séries temporais pode ser realizada segundo um modelo univariado, em que as características de uma série são analisadas apenas a partir do comportamento da própria série, considerando-se fatores como tendência e sazonalidade, ou segundo um modelo multivariado, em que as características de uma série são entendidas em função do comportamento de outras séries, o que implica na consideração de relações de dependência e de causalidade.

As séries temporais podem ser diferenciadas entre séries contínuas e discretas. A série contínua tem medidas realizadas ao longo do tempo de forma contínua, tal como um termógrafo ou um sismógrafo. A série discreta tem as medidas realizadas em tempo específico igualmente espaçado, como por exemplo, a temperatura média mensal, ou os próprios dados acima ilustrados (que foram obtidos a cada hora).



Harvey (1981) e Morettin e Toloí (1981) apresentam basicamente dois enfoques para a análise de séries temporais, ambos objetivando construir modelos para as séries. No primeiro enfoque a análise é feita no domínio temporal, e os modelos propostos são modelos paramétricos (com número finito de parâmetros). O segundo enfoque realiza a análise no domínio das frequências e os modelos propostos são modelos não paramétricos. Segundo Morettin e Toloí (1981), com a disponibilidade de uma série temporal, é possível investigar o mecanismo gerador da série temporal, fazer previsões de valores futuros (por exemplo: de curto prazo para séries de vendas, ou longo prazo para séries demográficas ou produtividade), descrever o comportamento da série (tendências, ciclos, variações sazonais) ou procurar periodicidades importantes nos dados.

Harvey (1981) enfocam dois aspectos para o estudo de séries temporais – análise e modelagem. O propósito da análise é resumir as propriedades de uma série e destacar suas características importantes, e a razão principal para modelagem de uma série temporal é permitir a previsão de valores futuros.

Carvalho (2005) analisa séries temporais de fenômenos naturais considerando noções de processos estocásticos, já que não é simples conseguir de forma clara uma relação que possa ser aplicada a cada intervalo de tempo da série de dados do fenômeno em estudo, para obter uma previsão. Isso existe devido a infinidade de fatores que causam variações nas medidas. Por exemplo, uma série temporal de temperatura de um determinado sítio em dois anos distintos apresenta diferenças, notadas visualmente. Processos estocásticos são processos controlados por leis probabilísticas (Morettin e Toloí (1981)) caracterizados por um conjunto ordenado de variáveis aleatórias  $\{X(t)\}$ , que representa os possíveis valores da série temporal, em associação com a sua distribuição de probabilidades. Carvalho (2005).

Fonseca (1988) sugere a geração de gráficos para se ter uma visão comportamental das informações para análise de qualquer série temporal. Isso porque a representação gráfica pode permitir a identificação de movimentos de tendências sazonais, cíclicos ou aleatórios.

A análise ou avaliação da tendência em séries temporais pode ser realizada utilizando alguns métodos:

- Método das médias móveis – este método tenta suavizar as variações da série através de médias sucessivas. É de fácil aplicação e muito flexível, pois a tendência não é forçada a adaptar-se a qualquer função matemática. As médias móveis tem a propriedade de reduzir ou eliminar as flutuações das variações indesejáveis, sendo muito utilizado para “alisamento das séries temporais”. Spiegel (1993).

- Método dos mínimos quadrados – pode ser usado para avaliar a equação de uma reta ou curva de tendência apropriada. Fonseca(1988).

- Método das semimédias – separam-se os dados em duas partes iguais e determina-se a média de cada parte, obtendo-se assim dois pontos do gráfico das séries temporais. Desenha-se uma reta entre esses dois pontos e assim determinam-se os valores da tendência. É mais bem aplicado em tendências lineares ou semi-lineares, contudo pode ser aplicado a fragmentos de tendências lineares dos dados. Spiegel (1993).

A avaliação da sazonalidade ou estacionalidade determina se a série tem um fator sazonal, ou seja, características que se repetem a intervalos regulares de tempo. Spiegel (1993) enumerou alguns métodos para o cálculo do índice de estacionalidade:

- Método da percentagem média – dados de cada mês são expressos em percentual da média anual, e as percentagens dos meses correspondentes em anos diferentes são balanceadas utilizando uma média ou mediana.

- Método da tendência ou relação percentual – dados de cada mês são expressos em percentagem dos valores da tendência mensal.

- Método da média móvel percentual ou da relação entre as médias móveis – neste método calcula-se uma média móvel de 12 meses, também a média móvel dos 2 meses daquela calculada de 12 meses (chama-se a isso de média móvel centrada de 12 meses). Então os dados originais de cada mês são expressos como percentagens da média móvel centrada de 12 meses correspondente, e por fim calcula-se a média das percentagens dos meses correspondentes gerando o índice esperado.

- Método dos elos relativos – dados de cada mês são expressos em percentagens nos dados do mês anterior, em que as percentagens são chamadas de elos relativos. Dos 12 elos relativos obtêm-se as percentagens relativas de cada mês, após ajustes para 100% realizados pela tendência analisada, consegue-se o índice estacional desejado.

O tema Séries Temporais é abordado de forma simplificada nesta dissertação, já que o método neural proposto usa uma abordagem diferenciada dos métodos clássicos de análise de séries temporais.

## **2.5 Utilização de RNA em Séries Temporais**

A procura dos padrões em uma série temporal normalmente utiliza processos ou métodos estatísticos ou estocásticos, que têm tido ótimos resultados em séries comportadas e sem interferências provocadas por irregularidades na série. Com esses métodos conseguem-se boas previsões para períodos curtos

ou próximos. Porém é dependente, de certa forma, do conhecimento da natureza das informações, isto é, muitas vezes não se conseguem determinar todas as variáveis que influenciam o comportamento da série temporal. Também a previsão sofre impactos negativos quando existe a necessidade de acrescentar outras variáveis em um modelo já preparado por métodos estocásticos.

Uma outra abordagem em utilização, para a análise de séries temporais, é a aplicação de Redes Neurais Artificiais (RNA) para o reconhecimento de padrões e processamento de previsão.

Segundo Cortez (1997), as RNA conseguem obter padrões das informações da série temporal sendo imunes às interferências e têm facilidade em adaptar seu comportamento à medida que novos dados são introduzidos, sem necessidade de alterar sua estrutura.

As previsões de séries temporais são utilizadas para solucionar problemas encontrados em vários campos de aplicação, como medicina, aplicações financeiras, hidrologia, engenharia, entre outros. As características das RNA tornam-nas adequadas para realizar previsões, inclusive a longo prazo, como no artigo de Simon et al (2005). Através dos exemplos encontrados na literatura percebemos a utilização das RNA para atender as necessidades citadas:

- Simon et al (2005) descrevem uma metodologia desenvolvida especialmente para gerar previsões a tempo distante, utilizando uma aplicação dupla de SOM – Mapa Auto-Organizado.
- Ferrari et al (2006) relatam o estudo que utilizou 4 redes neurais (Wavelet, LMS, MLP e RBF) para auxiliar no exame e diagnóstico da Doença Pulmonar

Obstrutiva Crônica (DPOC), pois são muitos os parâmetros envolvidos na sua identificação.

- Oliveira et al (2006) estudaram a variabilidade do nível do mar e a maré meteorológica na Baía de Paranaguá – PR através de métodos estatísticos com utilização de RNA com *Backpropagation* para previsões desses fenômenos. Após resultados concluíram a possibilidade de experimentar outras arquiteturas que conseguissem suprir as faltas das estações meteorológicas próximas de estações maregráficas.

- Bem et al (2003) fizeram um estudo da utilização de RNA e seu potencial na análise de padrões de variabilidade em áreas do oceano dominadas por fortes fenômenos de mesoescala, apresentando ao final as vantagens e desvantagens dessa abordagem.

- Neto et al (2005) apresentam os resultados de previsão da vazão afluente mensal do rio São Francisco através de RNA de Elman. Essa rede processou os dados da série temporal das vazões afluentes medidas entre 1931 e 1991, após um pré-processamento. Os dados de 1992 até 1996 foram utilizados para testes de generalização da rede e na validação dos resultados. Os resultados indicam a possibilidade do uso das redes de Elman para previsão da vazão e que o erro médio obtido foi menor que 0,2%.

## **2.6 Utilização de RNA em Hidrologia**

Na literatura existem vários trabalhos sobre a utilização de Redes Neurais Artificiais em hidrologia. Em geral são aplicações para previsão de vazão de rio, previsão de nível de rio, gerenciamento de reservatórios, entre muitos outros. Nesta dissertação são apresentados alguns exemplos de tais aplicações enfatizando a idéia principal de um.

Li et al (1998) utilizaram uma rede de Perceptrons de Múltiplas Camadas (*Backpropagation Neural Network* - BPNN) na análise da série temporal de valores de nível do rio Paraguai. No trabalho são relatadas previsões de nível do rio Paraguai com até 4 meses de antecedência, e até 17% de erros, utilizando 274 valores médios mensais de níveis do rio Paraguai.

Phien, Danh & Gupta (1999) utilizaram um BPNN para previsão de vazão em duas bacias (Da Nhim e La Nga) situadas nos planaltos do Vietnã em comparação a modelos de tanque. Conclusões obtidas indicam que modelos de rede BPNN são capazes de prever a vazão diária de um rio utilizando informações diárias de evaporação e de chuvas, e que quanto maior for a área utilizada para obter as informações, mais preciso será a previsão. Também observaram que valores históricos de vazão são importantes, como nos modelos de regressão, e por fim, que a contribuição do item evaporação para uma previsão mais precisa é muito pouco.

Atiya et al (1999) utilizaram uma RNA para o problema da previsão do escoamento do rio Nilo, no Egito, utilizando séries temporais para verificar o desempenho perante outros métodos de previsão que utilizaram RNA. Eles fizeram diversas comparações entre métodos diferentes para pré-processar as entradas e seus resultados. Também realizaram comparações entre tipos de técnicas utilizadas para resolver problemas de previsão de períodos mais longos e ainda entre métodos de previsão para horizonte muito mais distante, com possibilidade de particionar o problema em partes menores de previsão de  $k$  períodos adiante.

Muller e Fill (2003) utilizaram uma rede Perceptron de Múltiplas Camada (*Multilayer Perceptron Network* - MLP) com o algoritmo de aprendizagem Backpropagation, nos problemas de vazão de um trecho do rio Iguaçu, entre Fluvópolis/PR e União da Vitória/PR, em sete eventos ocorridos nos anos de 1982, 1983, 1987, 1990, 1992, 1993 e 1995, com o tempo de ocorrência

estimado entre 5 e 100 anos. Essa rede foi composta por 2 neurônios na camada de entrada, 3 camadas intermediárias com 8 neurônios por camada e a camada de saída com um neurônio, após experimentarem outras geometrias de rede de neurônios. As conclusões foram que utilizando os eventos de 1982, 1983, 1987 e 1990 para calibração e os restantes para a verificação, os resultados foram satisfatórios e animadores com os erros médios obtidos na faixa de 3% a 8%. Principalmente se considerar que é um trabalho matemático sem considerar a teoria intrínseca do problema. Por fim, concluíram que quanto maior e mais heterogêneo for o conjunto de eventos para calibração, melhor seriam os resultados obtidos, considerando um possível aperfeiçoamento do método utilizado.

Kisi (2003) apresenta comparações numéricas e gráficas do uso de RNA e modelos de auto-regressão na obtenção de valores de medidas de previsão de vazão de rios, utilizando medidas contínuas de escoamento do rio obtidas em períodos de 7 e 4 anos dos rios Blackwater e Gila, nos EUA, e em um período de 2 anos de medidas de escoamento do Canal de Filyos da Turquia. Os resultados obtidos mostram que a utilização de RNA tende a gerar resultados melhores que algoritmos de auto-regressão.

Baratti et al (2003) descrevem a utilização de uma rede MLP para previsão de vazão da bacia do Rio Tirso, na Sardenha (Itália). As conclusões foram que modelos baseados em RNA podem gerar ferramentas úteis de previsão de vazão e de estruturas hidrológicas para problemas em sistemas de planejamento e gerenciamento de recursos hídricos, utilizando dados mensais. Para dados diários, perceberam a necessidade de métodos e de redes neurais mais sofisticadas.

### 3 PREPARAÇÃO DOS DADOS

Este Capítulo apresenta os dados utilizados neste trabalho de dissertação. São descritos também os procedimentos realizados na preparação das informações para serem utilizados no treinamento da rede LVQ. Inicialmente são apresentadas as informações sobre os dados, prosseguindo pela explicação sobre a redução dos dados utilizados, os padrões discretizados de nível de rio e a geração dos dados de treinamento da rede LVQ.

#### 3.1 Dados

As informações que compõem o conjunto de dados utilizados neste trabalho de dissertação de Mestrado são os valores medidos e observados de uma régua de medição de nível do rio Paraguai, obtidos diariamente desde 1900, compondo uma base de dados com mais de 107 anos de medidas.

Essa régua de medição está localizada no município de Ladário (Figura 3.1) que tem 343 km<sup>2</sup> de área territorial, uma população em 2007 de 17906 habitantes e está situado às margens do rio Paraguai (IBGE, 2007).

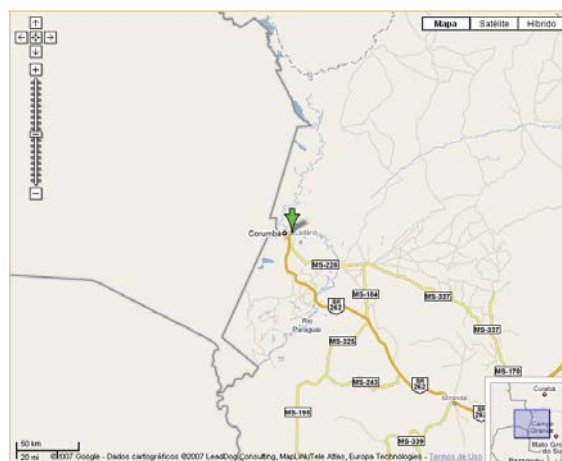


Figura 3.1 - Ladário - MS

Fonte: Google Maps (2007)

O rio Paraguai tem uma extensão de 2.621 km até sua foz no rio Paraná, sendo 1620 km no Brasil. É o principal dreno coletor das águas da Bacia do



Alto Paraguai (BAP) (Figura 3.2), que compreende o Pantanal e os planaltos adjacentes. Assim, comporta-se tal qual um escoador de águas do Pantanal.



Figura 3.2 - Bacia do Alto Paraguai

Fonte: Ministério dos Transportes (2005)

Entre os vários postos de medição de níveis de água do Rio Paraguai, destaca-se o do Ladário, vizinho de Corumbá, por ser o que mais dispõe de dados de toda uma rede instalada no Pantanal. O observatório de medição de Ladário acumula informações de registros diários desde o ano 1900 a 2007 sem falhas e se destaca também porque por ali passa 81% de vazão média do volume de água da Bacia do Alto Paraguai (BAP) que sai do território brasileiro para o território paraguaio. Atualmente pode-se acompanhar o nível do rio Paraguai diariamente, via internet, através do link <http://www.corumbaonline.com.br/nautica.asp>.

Dessa forma, essa régua de medição do nível do rio Paraguai (Figura 3.4 e Figura 3.5), instalada no 6º Distrito Naval da Marinha Brasileira (em Ladário, MS) constitui-se no principal referencial do regime hidrológico da BAP, possibilitando até caracterizar um dado período de seca ou de cheia. Considera-se ano de cheia quando o nível máximo do rio Paraguai em Ladário supera o nível de alerta de enchente, historicamente em 4,0 metros, caso contrário, ano de seca se ficar abaixo desses 4,0 metros. Seguindo esse princípio, pode-se dizer, em ano de super seca, se o nível do rio Paraguai ficará abaixo de 2,0 m. Na Figura 3.3 detalha-se melhor essa classificação de ano de cheia, em cheia pequena quando o nível do rio Paraguai atinge de 4,0 m a 5,0, de cheia normal entre 5,0 m a 6,0 m e ano de super cheia quando atinge nível máximo acima de 6,0.

Pico das cheias do Rio Paraguai, em Ladário, de 1973 a 2006

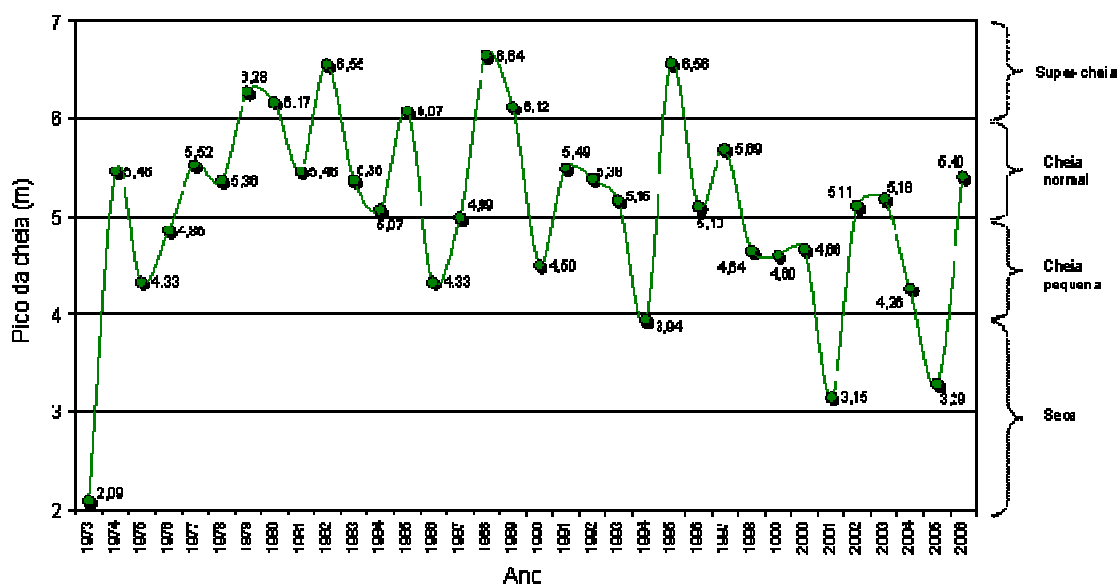


Figura 3.3 - Classificação de ano de cheia ou de seca

Fonte: Embrapa (2007)

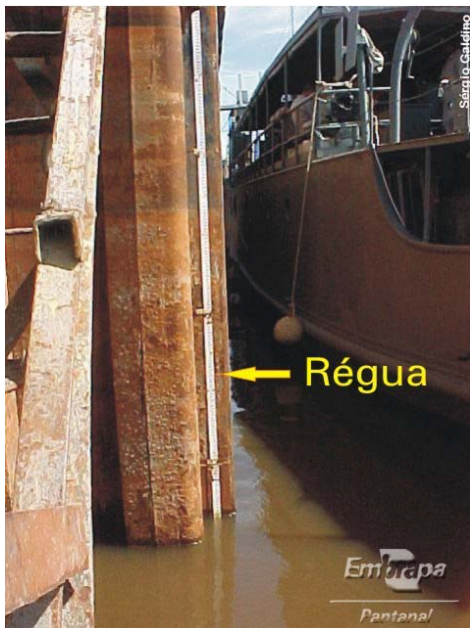


Figura 3.4 - Régua de Medição do nível do Rio Paraguai em Ladário/MS  
Fonte: Sérgio Galdino (2005)



Figura 3.5 - Detalhe da Régua de Medição de Ladário/MS  
Fonte: Sérgio Galdino (2005)

Vale observar que as cheias do Pantanal devem-se às características do relevo e à concentração das chuvas durante o verão na região da Bacia do Alto Paraguai. O relevo, na parte elevada da bacia possui altitudes entre 200 e 700 metros. O Pantanal, por sua vez, apresenta uma declividade de 0,7 a 5 cm/km no sentido norte-sul e entre 7 a 50 cm/km no sentido leste-oeste, é praticamente plano como podemos ver na Figura 3.6. Na figura se observa a declividade do rio Paraguai em relação à sua extensão desde Cáceres (MT) até sua foz no Rio Paraná. Soma-se a isso, o volume de água que vem do planalto (parte elevada da bacia) e o volume das chuvas locais. Ocorrem então as enchentes, pois o escoamento das águas é muito lento, favorecendo inundações em áreas extensas (Embrapa(2003)). Nessas ocasiões o ecossistema do Pantanal pode chegar a 138 mil km<sup>2</sup> dos 366 mil km<sup>2</sup> que a bacia do rio Paraguai ocupa de acordo com o Plano de Conservação da Bacia do Alto Paraguai Cristina – Radiobrás (2003).



Figura 3.6 - Perfil longitudinal do Rio Paraguai

Fonte: Ministério dos Transportes (2006).

Na Figura 3.6 pode-se observar a declividade do Rio Paraguai, que ao longo de 2300 km, desce de uma altitude de 110 m em Cáceres (MT) para 44 m na sua foz no rio Paraná em território argentino. Resulta então em uma declividade média de 3,2 cm/km. Percebemos ainda na Figura 3.5 os símbolos representativos de portos fluviais, cotas altimétricas, declividade média nos trechos, entre outras informações.

A Figura 3.7 exhibe a profundidade do rio Paraguai, importante para a navegação fluvial.

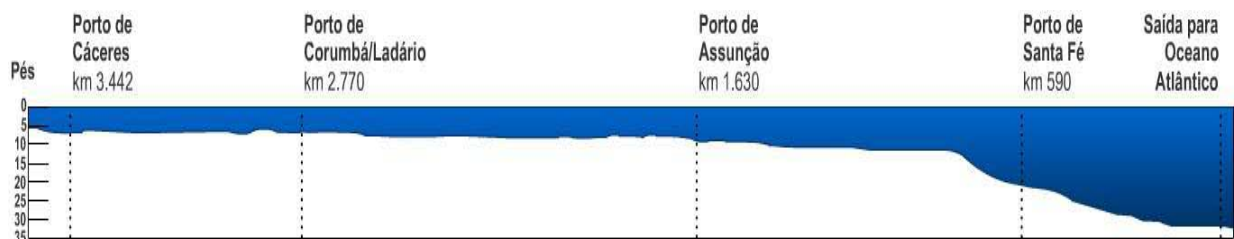


Figura 3.7 - Perfil da Profundidade do Rio Paraguai - MCT

As medidas obtidas em Ladário são exemplificadas na Tabela 3.1, com as colunas de informações:

- Dia, mês e ano: informam a data em que foi realizada a observação da régua de medição do nível do rio Paraguai;
- Altura da coluna de água: informa a medida observada do nível do rio Paraguai.

Tabela 3.1 - Exemplo das informações tabuladas dos dados de Ladário/MS

Níveis do Rio Paraguai na Estação Limnimétrica de Ladário (MS) 1900 - 2007			
DIA	MES	ANO	Altura da Coluna de Água (m)
1	1	1900	1,85
2	1	1900	1,85
3	1	1900	1,85
4	1	1900	1,85
5	1	1900	1,85
6	1	1900	1,85
7	1	1900	1,85
8	1	1900	1,85
9	1	1900	1,85
10	1	1900	1,85
11	1	1900	1,85
12	1	1900	1,85
13	1	1900	1,86
14	1	1900	1,87
15	1	1900	1,87
16	1	1900	1,87
17	1	1900	1,87
18	1	1900	1,89
19	1	1900	1,93
20	1	1900	1,98
21	1	1900	1,99
...			...
11	6	2007	4,95
12	6	2007	4,94

Esses valores fizeram parte do trabalho de Li et al (1998) sobre a utilização de redes neurais para previsão do nível do rio Paraguai para períodos com até 4 meses de antecedência. O trabalho explica como foram realizados os processos para a previsão dos níveis do rio Paraguai. Esse processo utilizou-se da média mensal dos níveis do rio Paraguai e técnicas de Redes Neurais para gerar os valores da previsão, obtendo os resultados aceitáveis para determinados períodos, dado certo período de treinamento e obtidas taxas de erros de aproximadamente 17%.

Este trabalho baseia-se na hipótese de melhoria dos resultados apresentados por Li et al (1998) utilizando outros métodos de determinação e previsão, que

possa introduzir maior precisão e ou facilidade de implementação. O trabalho justifica-se pela necessidade da indicação de ocorrência de “cheia” ou “vazante”, ajudando no planejamento e manuseio de equipamentos sociais, trazendo benefícios econômicos e sociais.

Uma das técnicas pensadas é o uso de uma SOM em que cada posição (neurônios) do mapa representa uma informação e geralmente evidenciam-se regiões com similaridades.

Outra técnica considerada é o LVQ – *Learning Vector Quantization*, uma rede semelhante ao SOM, também simples de implementar e criar visualizações dos resultados, e que tem como diferencial, a separação maior dos padrões não-correspondentes nas informações de entrada.

Para utilizar esses métodos, é necessário primeiramente se obter padrões dos dados do estudo. Esses padrões podem ser obtidos através da própria aplicação de rede neural, com escolha de métodos mais adequados ao trabalho, que pode ser o próprio SOM, mas, neste trabalho, optou-se pela obtenção por meios empíricos e observacionais.

Os padrões obtidos podem ser utilizados no algoritmo do LVQ, que necessita conhecer os padrões que existem no conjunto de dados, ou seja, é necessário estabelecer rótulos nos dados de entrada.

A utilização do LVQ tem como objetivo classificar os resultados semelhantes do SOM de uma forma melhorada ou detalhada, com suas classes de padrões bem separadas em um mapa. Isto é mapear os padrões conhecidos, conforme Brownlee (2005). Obtido esse mapa, é possível utilizar algum método de visualização para demonstrar as características obtidas. De maneira geral, o SOM fornece um mapa ou matriz em que se podem identificar os vários padrões que possam existir nos dados de entrada. A matriz (mapa) pode ter

regiões mais destacadas e outras regiões mais discretas. Enfim, se consegue distinguir claramente uma classificação dos dados. Com o uso do LVQ, essas regiões podem ser melhoradas e detalhadas.

Os resultados dos experimentos são confrontados com os dados reais, com a correspondente verificação da percentagem de acerto e de erros.

### **3.2 Redução**

O procedimento utilizado para reduzir a quantidade de dados foi realizado através de amostragens do conjunto de medidas. Os estudos realizados através de gráficos gerados por medidas consecutivas e com os valores reduzidos demonstraram que a diferença entre o processamento dos dados originais e dos dados das amostragens é pequena, para reduções até 15 medidas consecutivas, possibilitando utilizar esse procedimento sem perdas significativas de informações.

#### **3.2.1 Redução dos dados de medidas de nível do rio Paraguai**

Foi observado que a variação das medidas de nível do Rio Paraguai de um dia para outro é em média 0,02m, e em muitas ocasiões a variação é nula ou abaixo de 0,01m.

O conjunto de gráficos que compõe a Figura 3.8 é um exemplo da variação do nível do rio Paraguai. No primeiro gráfico do conjunto da linha superior, no gráfico da esquerda, são representados os níveis medidos do dia 1 ao dia 5 do mês de janeiro de 1900. O segundo gráfico apresenta as medidas do dia 1 ao dia 10. O gráfico da direita, ainda da linha superior, estão as medidas do dia 1 ao dia 15. Na segunda linha, o primeiro gráfico pela esquerda, apresenta as medidas do dia 1 a 30 de janeiro. O próximo gráfico apresenta as medidas de 1 de janeiro até 15 de fevereiro de 1900 e o seguinte, as medidas de 1 de janeiro

até 28 de janeiro de 1900. O gráfico maior apresenta as medidas de 1 de janeiro de 1900 até 31 de março de 1900.

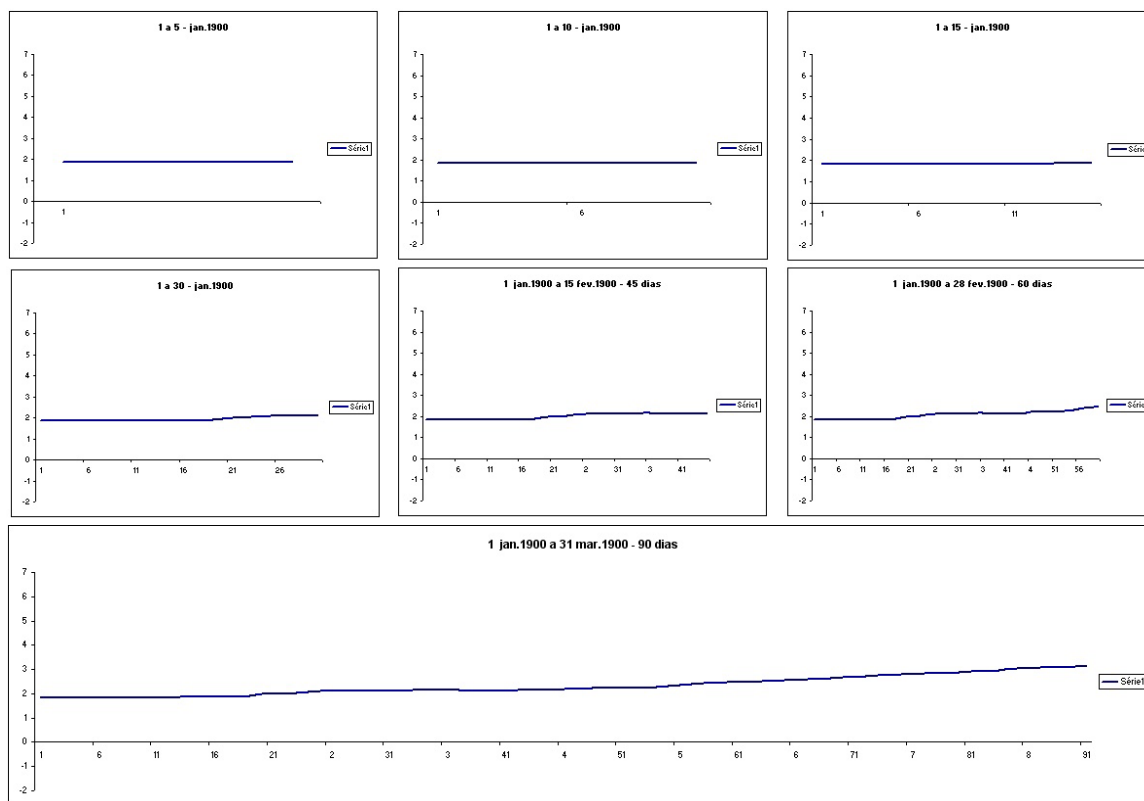


Figura 3.8 - Exemplo da variação do nível do Rio Paraguai.

Pela observação dos gráficos, notou-se uma variação significativa somente a partir de 15 medidas seguidas. Essa variação tornou-se mais evidente ao considerar um período maior, principalmente no último gráfico, o maior, que apresenta medidas diárias de 3 meses, de janeiro a 31 de março, totalizando 90 dias.

A redução foi utilizada na fase de pré-processamento para tratar estas informações de medida de nível do rio de forma adequada, sem perder os detalhes nem utilizar dados repetidos que não influenciem o resultado final do trabalho de forma significativa. Calcular a média simples dos valores de algum período, qualquer que seja, causa a perda de detalhes. Para evitar isso, recorreu-se a um processo de cálculo de médias de um período com deslocamento de janela desse período, sobrepondo em alguns valores. Com



esse método ainda temos perdas, porém é menor que o cálculo da média simples.

Por estas observações, utilizou-se a redução através de médias com deslocamento de uma medida para os experimentos deste trabalho. Os gráficos da Figura 3.9 abaixo apresentam as medidas de nível do rio com redução em média semanal, quinzenal, mensal, 45 dias e 60 dias, para o ano de 1900. Na coluna esquerda são apresentados os gráficos com marcação das medidas e na coluna direita os gráficos sem as marcações.

Os gráficos da primeira linha apresentam as medidas diárias do ano 1900, sem redução.

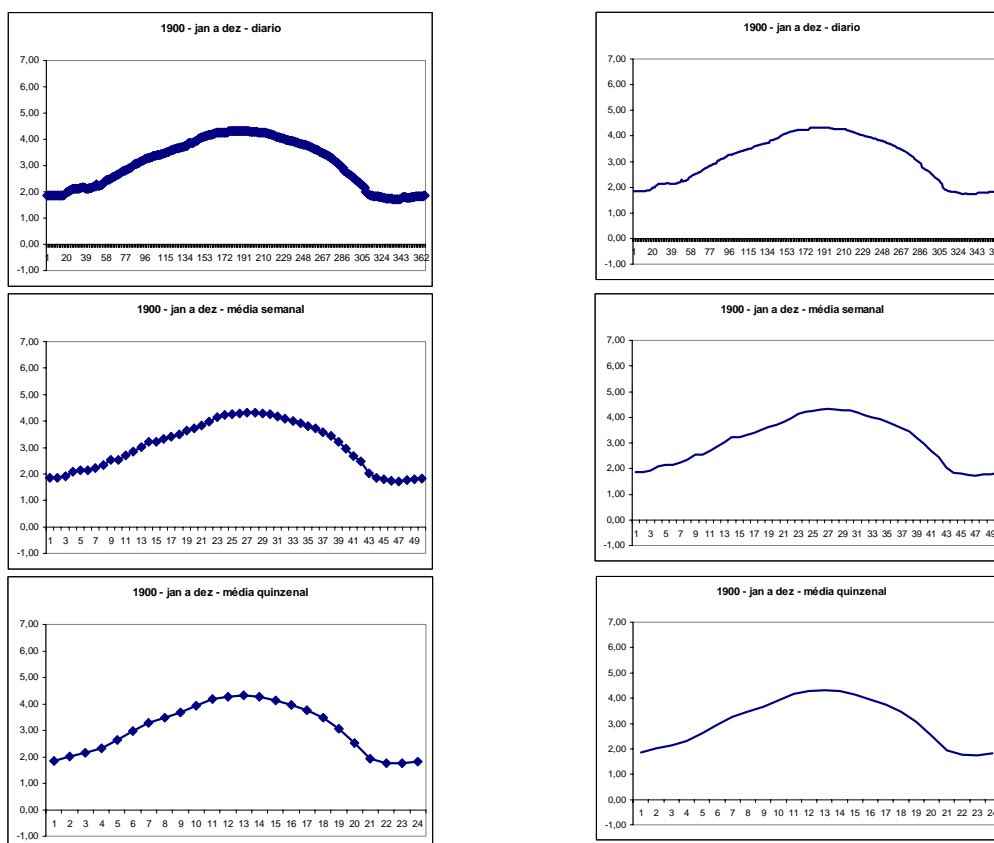


Figura 3.9 - Diferenças de representação das médias de medidas.

(continua)

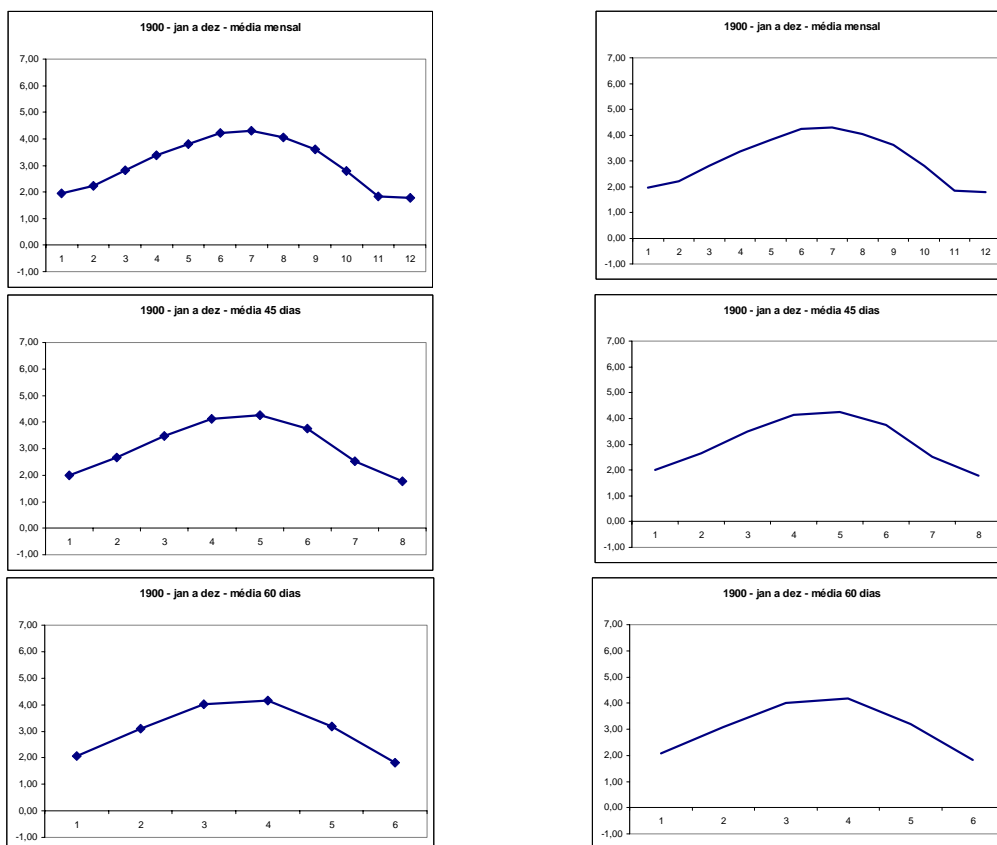


Figura 3.9 - Diferenças na representação das médias de medidas.

(conclusão)

Percebeu-se que, quanto maior o número de medidas usadas para o cálculo da média, mais detalhes da evolução do nível do rio são ignorados, como se pode observar nos gráficos da Figura 3.9.

Tabela 3.2 - Exemplo de medidas de nível do Rio Paraguai obtidos em Ladário e os valores médios calculados, em janelas temporais com número de medidas igual a: 3, 5, 10 e 15 medidas

DIA	MES	ANO	Altura da Coluna de água	Data da Amostra	Médias	Data da Amostra	Médias	Data da Amostra	Médias	Data da Amostra	Médias
1	1	1900	1,85	<b>A = 3 medidas</b>		<b>A = 5 medidas</b>		<b>A = 10 medidas</b>		<b>A = 15 medidas</b>	
2	1	1900	1,85								
3	1	1900	1,85	3/1/1900	1,85						
4	1	1900	1,85								
5	1	1900	1,85			5/1/1900	1,85				
6	1	1900	1,85	6/1/1900	1,85						
7	1	1900	1,85								
8	1	1900	1,85								
9	1	1900	1,85	9/1/1900	1,85						
10	1	1900	1,85			10/1/1900	1,85	10/1/1900	1,85		
11	1	1900	1,85								
12	1	1900	1,85	12/1/1900	1,85						
13	1	1900	1,86								
14	1	1900	1,87								
15	1	1900	1,87	15/1/1900	1,87	15/1/1900	1,86			15/1/1900	1,85
16	1	1900	1,87								
17	1	1900	1,87								
18	1	1900	1,89	18/1/1900	1,88						
19	1	1900	1,93								
20	1	1900	1,98			20/1/1900	1,91	20/1/1900	1,88		
21	1	1900	1,98	21/1/1900	1,97						
22	1	1900	2,00								
23	1	1900	2,05								
24	1	1900	2,07	24/1/1900	2,04						
25	1	1900	2,1			25/1/1900	2,04				
26	1	1900	2,12								
27	1	1900	2,12	27/1/1900	2,11						
28	1	1900	2,12								
29	1	1900	2,12								
30	1	1900	2,12	30/1/1900	2,12	30/1/1900	2,12	30/1/1900	2,08	30/1/1900	2,02
...			...	...	...	...	...	...	...	...	...

A Tabela 3.2 apresenta um pequeno exemplo dos dados originários de Ladário. A primeira coluna (DIA MÊS ANO) contém a data em que foi realizada a medição do nível do rio Paraguai. A segunda coluna (“Altura da Coluna de Água”) contém o valor que foi anotado pela observação da régua de medidas existente no observatório. As colunas restantes apresentam os valores de média das janelas temporais definidas pelo parâmetro A (número de medidas para média). Essas colunas apresentam as informações:

- Data da Amostra: indica a data da última medida utilizada para o cálculo da média das medidas;
- Médias: indica a média calculada das medidas em quantidades definidas pelo parâmetro A.

Após a obtenção das médias, foi gerado um novo conjunto de dados composto pelo agrupamento de um número (parâmetro D) desses valores consecutivos,

formando um vetor de médias. O valor do parâmetro D corresponde, portanto, ao valor da dimensão desse vetor, e é informado na execução desse procedimento. A última média do vetor de cada registro fornece a data da última medida dessa janela temporal para datar esse registro. Assim, entende-se que os outros valores correspondem aos períodos antecedentes a essa data do registro.

Tabela 3.3 - Exemplo de conteúdo dos arquivos gerados na amostragem 4 médias

D = 4 médias				data da última medida da janela temporal da última média do registro
média 1	média 2	média 3	média 4	
1,85	1,85	1,85	1,85	19000112
1,85	1,85	1,85	1,87	19000115
1,85	1,85	1,87	1,88	19000118
1,85	1,87	1,88	1,97	19000121
1,87	1,88	1,97	2,04	19000124
1,88	1,97	2,04	2,11	19000127
1,97	2,04	2,11	2,12	19000130
2,04	2,11	2,12	2,13	19000202
2,11	2,12	2,13	2,16	19000205
2,12	2,13	2,16	2,13	19000208
2,13	2,16	2,13	2,13	19000211
2,16	2,13	2,13	2,15	19000214
2,13	2,13	2,15	2,2	19000217
...	...	...	...	...

entrada\_lvq\_a3\_d4.txt

Tabela 3.4 - Exemplo de conteúdo dos arquivos gerados na amostragem 8 médias

D = 8 médias								data da última medida da janela temporal da última média do registro
média 1	média 2	média 3	média 4	média 5	média 5	média 7	média 8	
1,85	1,85	1,85	1,85	1,87	1,88	1,97	2,04	19000124
1,85	1,85	1,85	1,87	1,88	1,97	2,04	2,11	19000127
1,85	1,85	1,87	1,88	1,97	2,04	2,11	2,12	19000130
1,85	1,87	1,88	1,97	2,04	2,11	2,12	2,13	19000202
1,87	1,88	1,97	2,04	2,11	2,12	2,13	2,16	19000205
1,88	1,97	2,04	2,11	2,12	2,13	2,16	2,13	19000208
1,97	2,04	2,11	2,12	2,13	2,16	2,13	2,13	19000211
2,04	2,11	2,12	2,13	2,16	2,13	2,13	2,15	19000214
2,11	2,12	2,13	2,16	2,13	2,13	2,15	2,2	19000217
2,12	2,13	2,16	2,13	2,13	2,15	2,2	2,25	19000220
2,13	2,16	2,13	2,13	2,15	2,2	2,25	2,25	19000223
2,16	2,13	2,13	2,15	2,2	2,25	2,25	2,35	19000226
2,13	2,13	2,15	2,2	2,25	2,25	2,35	2,45	19000301
...	...	...	...	...	...	...	...	...

entrada\_lvq\_a3\_d8.txt

A Tabela 3.3 e a Tabela 3.4 apresentam dois exemplos de conjuntos de dados após a redução. As colunas tituladas “média 1”, “média 2”, e assim por diante, contêm as médias já calculadas. A última coluna apresenta a data da última medida do período utilizado para calcular o valor do último elemento do vetor de cada registro, isto é, a data da última medida da série temporal do registro.

### 3.3 Padrões de Nível do rio Paraguai

Para representar os vários níveis que o rio pode ter, foram discretizados em um número de padrões informado por parâmetro (F) os valores do nível mínimo até o nível máximo, de todo o conjunto de medidas processadas.

Um algoritmo de discretização foi desenvolvido para a criação desses padrões. O algoritmo cria um outro arquivo contendo as informações exemplificadas nas Tabelas 3.5, 3.6 e 3.7.

Tabela 3.5 - 10 padrões

Padrões		
De: -0.61 ate: 6.64 em 10 faixas		
Padrão	Nível inicio	Nível fim
F-0.61	-0,61	0,19
F0.20	0,20	0,99
F1.00	1,00	1,80
F1.81	1,81	2,60
F2.61	2,61	3,41
F3.42	3,42	4,21
F4.22	4,22	5,02
F5.03	5,03	5,82
F5.83	5,83	6,63
F6.64	6,64	7,44

faixas\_a3\_F10.txt

Tabela 3.6 - 20 padrões

Padrões		
De: -0.6 ate: 6.61 em 20 faixas		
Padrão	Nível inicio	Nível fim
F-0.60	-0,6	-0,23
F-0.22	-0,22	0,15
F0.16	0,16	0,53
F0.54	0,54	0,91
F0.92	0,92	1,29
F1.30	1,3	1,67
F1.68	1,68	2,05
F2.06	2,06	2,43
F2.44	2,44	2,81
F2.82	2,82	3,18
F3.19	3,19	3,56
F3.57	3,57	3,94
F3.95	3,95	4,32
F4.33	4,33	4,7
F4.71	4,71	5,08
F5.09	5,09	5,46
F5.47	5,47	5,84
F5.85	5,85	6,22
F6.23	6,23	6,6
F6.61	6,61	6,98

faixas\_a3\_F20.txt

Tabela 3.7 - 30 padrões

Padrões		
De: -0.6 ate: 6.61 em 30 faixas		
Padrão	Nível inicio	Nível fim
F-0.60	-0,6	-0,36
F-0.35	-0,35	-0,11
F-0.10	-0,1	0,14
F0.15	0,15	0,38
F0.39	0,39	0,63
F0.64	0,64	0,88
F0.89	0,89	1,13
F1.14	1,14	1,38
F1.39	1,39	1,63
F1.64	1,64	1,88
F1.89	1,89	2,12
F2.13	2,13	2,37
F2.38	2,38	2,62
F2.63	2,63	2,87
F2.88	2,88	3,12
F3.13	3,13	3,37
F3.38	3,38	3,62
F3.63	3,63	3,87
F3.88	3,88	4,11
F4.12	4,12	4,36
F4.37	4,37	4,61
F4.62	4,62	4,86
F4.87	4,87	5,11
F5.12	5,12	5,36
F5.37	5,37	5,61
F5.62	5,62	5,85
F5.86	5,86	6,1
F6.11	6,11	6,35
F6.36	6,36	6,6
F6.61	6,61	6,85

faixas\_a3\_F30.txt

Nas Tabelas 3.5, 3.6 e 3.7 são apresentados alguns exemplos de padrões utilizados para gerar os dados de treinamento. Na primeira tabela, foram criados 10 padrões, na segunda 20 e na terceira, 30 padrões. As colunas “Padrão” apresentam os identificadores da faixa de nível, isto é, o padrão ou classe de nível. As colunas “Nível início” apresentam os valores iniciais dessa faixa de nível e as colunas “Nível fim” apresentam os valores finais da faixa. Os limites de cada faixa, as identificações dos padrões e os valores de nível máximo e mínimo de cada tabela são obtidos pelo próprio algoritmo que gera os padrões. Apenas o número de faixas ou padrões é fornecido como parâmetro (F).

### **3.4 Geração dos dados de treinamento**

O arquivo de vetores de entrada para o treinamento da rede LVQ utiliza os dados de médias associados com os padrões estabelecidos pelo processo de discretização. O processamento utiliza o último valor do vetor de médias de cada registro para obter o padrão que representa esse vetor, e chamamos esse valor de conseqüente. O padrão encontrado é aquele que, na comparação do conseqüente com os níveis inicial e final de cada faixa, situar-se nesse intervalo.

Assim cada registro do arquivo de medidas gera um registro no novo arquivo, isto é, contém o vetor de médias, a data da última medida da série temporal do registro e o padrão encontrado pelo procedimento descrito. Exemplos de conjuntos de dados gerados podem ser observados na Tabela 3.8 e na Tabela 3.9.

Tabela 3.8 - Exemplo de arquivo com 4 médias por entrada, gerado para utilizar no treinamento da rede LVQ

D = 4 médias

média 1	média 2	média 3	média 4	Padrão	data da última medida da janela temporal da última média do registro
1,85	1,85	1,85	1,85	F1.81	19000112
1,85	1,85	1,85	1,87	F1.81	19000115
1,85	1,85	1,87	1,88	F1.81	19000118
1,85	1,87	1,88	1,97	F1.81	19000121
1,87	1,88	1,97	2,04	F1.81	19000124
1,88	1,97	2,04	2,11	F1.81	19000127
1,97	2,04	2,11	2,12	F1.81	19000130
2,04	2,11	2,12	2,13	F1.81	19000202
2,11	2,12	2,13	2,16	F1.81	19000205
2,12	2,13	2,16	2,13	F1.81	19000208
2,13	2,16	2,13	2,13	F1.81	19000211
2,16	2,13	2,13	2,15	F1.81	19000214
2,13	2,13	2,15	2,2	F1.81	19000217
...	...	...	...	...	...

ent\_lvq\_a3\_d4\_F10.txt

Tabela 3.9 - Exemplo de arquivo com 8 médias por entrada, gerado para utilizar no treinamento da rede LVQ

D = 8 médias

média 1	média 2	média 3	média 4	média 5	média 6	média 7	média 8	Padrão	data da última medida da janela temporal da última média do registro
1,85	1,85	1,85	1,85	1,87	1,88	1,97	2,04	F1.81	19000124
1,85	1,85	1,85	1,87	1,88	1,97	2,04	2,11	F1.81	19000127
1,85	1,85	1,87	1,88	1,97	2,04	2,11	2,12	F1.81	19000130
1,85	1,87	1,88	1,97	2,04	2,11	2,12	2,13	F1.81	19000202
1,87	1,88	1,97	2,04	2,11	2,12	2,13	2,16	F1.81	19000205
1,88	1,97	2,04	2,11	2,12	2,13	2,16	2,13	F1.81	19000208
1,97	2,04	2,11	2,12	2,13	2,16	2,13	2,13	F1.81	19000211
2,04	2,11	2,12	2,13	2,16	2,13	2,13	2,15	F1.81	19000214
2,11	2,12	2,13	2,16	2,13	2,13	2,15	2,2	F1.81	19000217
2,12	2,13	2,16	2,13	2,13	2,15	2,2	2,25	F1.81	19000220
2,13	2,16	2,13	2,13	2,15	2,2	2,25	2,25	F1.81	19000223
2,16	2,13	2,13	2,15	2,2	2,25	2,25	2,35	F1.81	19000226
2,13	2,13	2,15	2,2	2,25	2,25	2,35	2,45	F1.81	19000301
...	...	...	...	...	...	...	...	...	...

ent\_lvq\_a3\_d8\_F10.txt

Os procedimentos descritos utilizaram os parâmetros para gerar os conjuntos intermediários e o conjunto que será utilizado no treinamento da rede LVQ.

Esses parâmetros são:

Tabela 3.10 - Parâmetros utilizados no treinamento da rede LVQ

Amostragem	A=3; A=10; A=15	é número de medidas para o cálculo das médias
Dimensão	D=4. D=8	é o número de médias por registro
Padrão	F=10, F=20, F=30	é o número de padrões gerados e utilizados no treinamento da rede LVQ

Os conjuntos gerados foram nomeados utilizando esses parâmetros para facilitar o reconhecimento do conteúdo. Então como exemplo, nos arquivos da Tabela 3.3 e Tabela 3.4, o código “a3” informa que os arquivos contêm médias obtidas de 3 medidas. O código “d4” informa que o arquivo contém 4 médias por registro e o código “d8” que o arquivo contém 8 médias por registro.

Nos exemplos da Tabela 3.8 e Tabela 3.9, o nome do arquivo informa que o mesmo contém médias de 3 medidas, 4 médias por registro (Tabela 3.8) ou 8 médias por registro (Tabela 3.9) e utilizou 10 faixas de nível para obter o padrão de cada registro e rotulá-lo.

Ao fim desses procedimentos, se obtém os dados a serem utilizados para treinar a rede LVQ.

É importante destacar que para cada arquivo gerado com parâmetros de amostragem diferente, deve-se criar um arquivo de padrões discretizados utilizando as mesmas informações que geraram aquele arquivo. Isso porque pode ocorrer de se alterar o valor do nível mínimo e/ou máximo de rio nas informações utilizadas. Portanto os padrões não conseguem representar todos os níveis de rio do arquivo gerado, produzindo ao final do processo, uma rede LVQ não confiável.





## **4 TREINAMENTO DA REDE LVQ, PREVISÕES E RESULTADOS**

Neste Capítulo são apresentados os experimentos realizados para o treinamento da rede LVQ, os experimentos para a previsão do nível do rio Paraguai e as análises dos resultados obtidos. Inicialmente é apresentada a metodologia proposta e utilizada neste trabalho. A seguir são expostos os dados do artigo de Li et al (1998) para explicar o treinamento das várias redes LVQ de períodos selecionados, prosseguindo pela apresentação dos procedimentos para o treinamento das redes LVQ adotados pela metodologia proposta, os resultados dos processamentos e suas análises. Por fim, são apresentadas as conclusões obtidas neste Capítulo.

Observamos que os experimentos foram processados de forma que permitissem comparações com os resultados obtidos e descritos no artigo de Li et al (1998), por terem utilizadas as mesmas informações dos nossos experimentos, isto é, as medidas de nível do rio Paraguai.

Essa comparação simples permite avaliar e validar a metodologia utilizada para realizar previsões utilizando redes LVQ.

### **4.1 Metodologia**

A metodologia utilizada neste trabalho é composta por fases que seguem uma ordem de procedimentos separados em: pré-processamento, treinamento da rede LVQ, previsão e geração do mapa da rede LVQ.

A fase de pré-processamento é responsável por gerar os dados para o treinamento utilizando as medidas históricas de nível do rio Paraguai, como explicado no item 3.2 Redução. Os resultados desta fase são os arquivos de entrada para o treinamento da rede LVQ.

A fase de treinamento é a que cria a rede LVQ, utilizando parâmetros de configuração como exposto na subseção 4.3.1 Treinamento da Rede LVQ, pág. 79.

Na próxima fase as previsões são processadas, utilizando a rede LVQ gerada e as medidas de referência. A previsão utiliza referências que são as medidas anteriores ao período ou data que se quer prever.

A última fase é a da confecção do mapa da rede LVQ utilizada na previsão. Esta fase pode ser executada antes da previsão, pois somente necessita da existência da rede LVQ para seu processamento.

A Figura 4.1 apresenta a metodologia proposta, através de um diagrama de fluxo de dados.

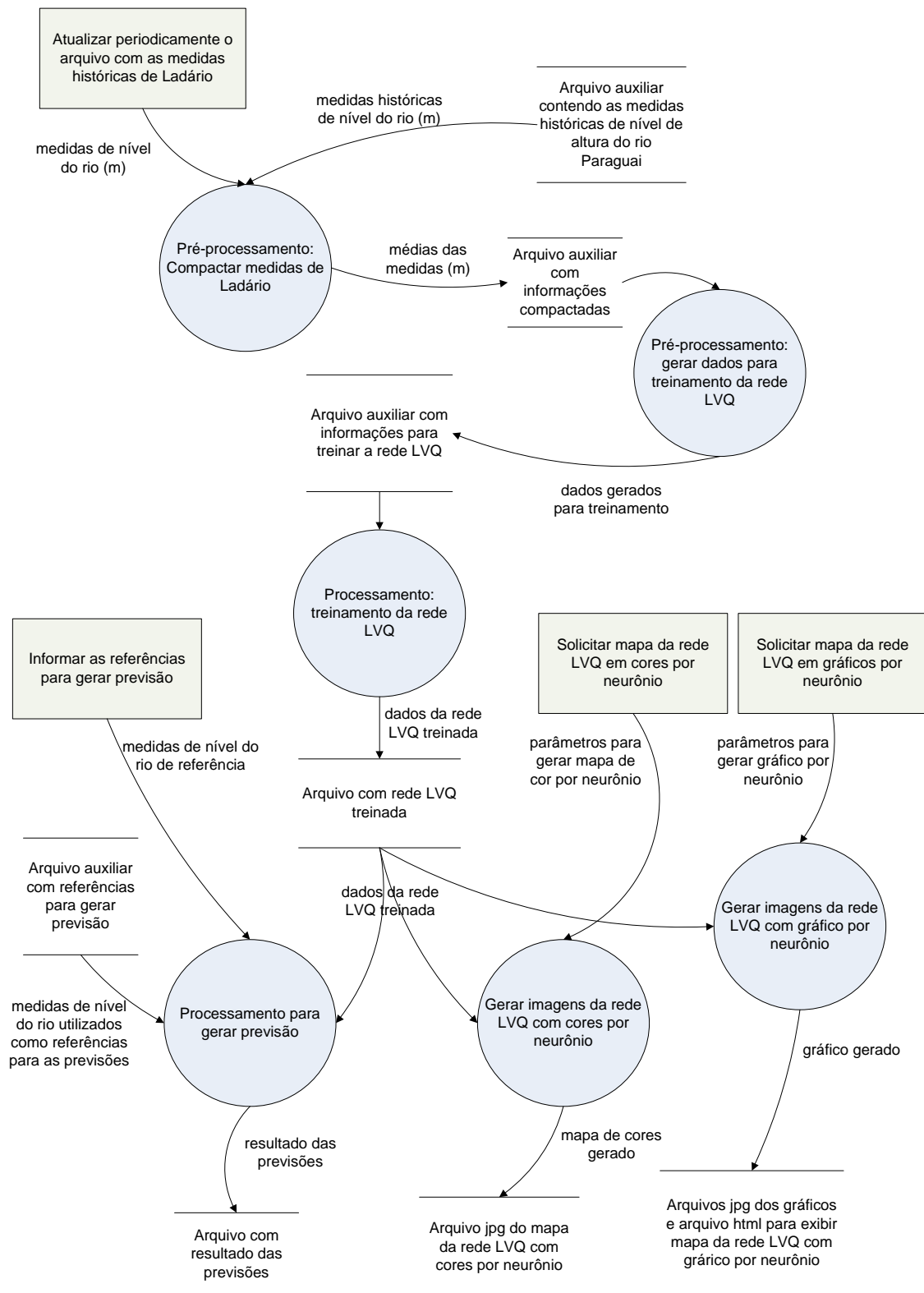


Figura 4.1 – Metodologia para previsão utilizando rede LVQ

Na Figura 4.1 são apresentados dois procedimentos de pré-processamento: um para a redução das medidas de Ladário e o procedimento para gerar o arquivo para treinar a rede LVQ. Observam-se as intervenções para atualizar as medidas de nível do rio Paraguai originadas de Ladário, informar as medidas de referência para processar previsão e para processar os mapas da rede LVQ informando os parâmetros necessários.

Nesta metodologia, o resultado da previsão processada é gerado em um arquivo texto que contém informações como: padrão do nível previsto, posição do neurônio que resultou a previsão, em linha e coluna. Pode-se também gerar 3 tipos de mapas da rede LVQ: um mapa com gráficos de linha, um mapa com neurônios colorizados por padrão e um mapa com neurônios colorizados pelos próprios pesos, sendo estes dois últimos tipos apresentados adiante.

O funcionamento do diagrama apresentado na Figura 4.1 é explanado nas linhas seguintes. As figuras retangulares representam as interações externas através de comandos ou inclusão de dados. As figuras em linhas paralelas representam os depósitos de dados. As figuras redondas representam os processamentos e as setas representam os fluxos de dados.

Inicialmente é atualizado o arquivo de medidas históricas de nível do rio Paraguai, com as medidas observadas mais recentes, realizado por algum operador da metodologia. Após é executada a primeira fase do pré-processamento, que gera um arquivo com as medidas reduzidas para a fase seguinte. Nesta fase são utilizados parâmetros que indicam o número de medidas utilizadas no cálculo das médias para a redução dos dados de entrada e os nomes do arquivo de entrada e de saída.

Na segunda fase do pré-processamento é gerado o arquivo para treinar a rede LVQ. São utilizados parâmetros que informam o número de valores peso de

cada neurônio da rede LVQ a ser treinada, o número de padrões discretizados a serem utilizados no treinamento e os nomes do arquivo de entrada e de saída desta fase.

O próximo passo ou processamento é o treinamento da rede LVQ, utilizando os dados gerados pelo pré-processamento e as informações dos parâmetros que indicam as características da rede a ser treinada. Ao final deste passo um arquivo é gerado com a rede LVQ treinada.

Esse arquivo é então utilizado para gerar os mapas da rede LVQ em gráfico de cada neurônio, em mapa de cores de cada neurônio, e gerar as previsões de nível do rio Paraguai.

Para gerar os mapas da rede LVQ, utilizam-se arquivos de configuração com os parâmetros das características solicitadas, conforme indicam os processos de geração das imagens da rede LVQ e seus fluxos de dados. Esses processos geram os mapas armazenados em arquivos.

Para gerar as previsões, utilizam-se medidas anteriores às previsões dos períodos solicitados, como referências nos processamentos, juntamente com os parâmetros que caracterizam essas previsões. Também esse processo gera um arquivo que contém os resultados das previsões solicitadas.

## **4.2 Resultados de referência para comparação**

Li et al (1998) descreve a utilização de Rede Neural Artificial com Retro Propagação para prever o nível médio do Rio Paraguai com 4 meses de antecedência. No artigo, a RNA utilizada foi alimentada com 274 valores médios mensais de nível do rio, calculados a partir de um conjunto de medidas obtidas diariamente em Ladário (MS), do período de maio de 1972 até fevereiro

de 1995. Através dela, os autores realizaram previsões de níveis médios do rio para os meses de março a junho de 1995.

A Tabela 4.1 apresenta os resultados obtidos nos processamentos descritos por Li et al (1998), para realizar as previsões de nível do Rio Paraguai.

Tabela 4.1 - Resultados descritos por Li et al (1998)

Resumo dos resultados apresentados no artigo

qtde de valores médios mensais	dados obtidos		para prever		Resultados obtidos				
	de	até	de	até					
274	mai/72	few/95	mar/95	jun/95	erro relativo	mar/95 7%	abr/95 17%	mai/95 13%	jun/95 13%
564	jan/00	dez/47	jan/48	abr/48	erro relativo	jan/47 24%	few/47 3%	mar/47 28%	abr/47 60%
344	jan/00	set/30	out/30	jan/31	erro relativo	out/30 5%	now/30 7%	dez/30 8%	jan/31 4%
700	jan/00	abr/59	mai/59	ago/59	erro relativo	mai/59 8%	jun/59 10%	jul/59 11%	ago/59 8%
861	jan/00	set/72	out/72	jan/73	erro relativo	out/72 33%	now/72 sem inf.	dez/72 47%	jan/73 70%
135	abr/83	few/95	mar/95	jun/95	erro relativo	mar/95 16%	abr/95 22%	mai/95 16%	jun/95 14%
1142	jan/00	few/95	mar/95	jun/95	erro relativo	mar/95 16%	abr/95 22%	mai/95 13%	jun/95 11%

A primeira coluna apresenta a quantidade de valores médios mensais utilizados no artigo. A coluna “qtde de valores médios mensais” apresenta a quantidade de valores médios mensais utilizados nos experimentos. As colunas “dados obtidos” e “para prever” mostram os períodos ou janelas temporais utilizados para se obter as médias mensais e os meses para serem previstos. Por fim, a coluna “Resultados obtidos” apresenta os valores de erros relativos calculados para cada mês previsto.

### 4.3 Procedimentos Realizados para Treinamento da Rede LVQ

Para criar condições semelhantes às dos experimentos do artigo, foram criados 7 tipos de arquivos contendo os mesmos períodos utilizados. Esses arquivos foram gerados na fase de pré-processamento, em que para cada arquivo foi selecionado um conjunto de medidas de nível do rio Paraguai, como é apresentado na Tabela 4.2.

Tabela 4.2 – Períodos utilizados para comparação dos resultados

Período	Medidas do intervalo
1900_jan_a_1930_set	de janeiro de 1900 até setembro de 1930
1900_jan_a_1947_dez	de janeiro de 1900 até dezembro de 1947
1900_jan_a_1959_abr	de janeiro de 1900 até abril de 1959
1900_jan_a_1972_set	de janeiro de 1900 até setembro de 1972
1900_jan_a_1995_fev	de janeiro de 1900 até fevereiro de 1995
1972_mai_a_1995_fev	de maio de 1972 até fevereiro de 1995
1983_abr_a_1995_fev	de abril de 1983 até fevereiro de 1995
1900_jan_a_2007_jul	de janeiro de 1900 até julho de 2007

A coluna Período apresenta o período utilizado para o treinamento das redes LVQ e a coluna Medidas do intervalo informa o início e o final do período selecionado. A última linha da Tabela 4.2 é destacada por não ser parte da comparação, e refere-se a todo o conjunto de medidas que foi utilizado para o processamento de previsões para o mês de agosto.

Uma seleção manual, de execução simples e rápida, foi realizada para gerar os arquivos referentes a cada período descrito.

#### 4.3.1 Treinamento da Rede LVQ

O treinamento consistiu em apresentar um arquivo de entrada e obter um arquivo de saída com a rede LVQ treinada. O término de cada treinamento gerou arquivo em formato texto contendo os valores-peso de cada neurônio da rede LVQ.

Como parâmetros para o treinamento das redes LVQ foram utilizadas:

- taxa de aprendizagem com o valor 0,9;
- taxa de decréscimo da aprendizagem com o valor 0,01 e
- taxa de aprendizagem mínima com o valor 0,01.

Outras configurações também foram utilizadas:

- número de medidas utilizadas para os cálculos das médias;
- número de médias por registro de entrada;



- tamanho das redes LVQ em quantidade de neurônios por linha e por coluna;
- número de padrões utilizados no treinamento das redes LVQ.

Foram treinadas 2100 redes, nas configurações com médias obtidas de 3, 10 e 30 medidas, 4, 6, 8 e 12 médias por registro de entrada, redes com 10x10, 20x20, 30x30, 50x50 e 100x100 neurônios e utilização de 10, 20, 30, 50 e 100 padrões de nível de rio, isto é, 3 x 4 x 5 x 5, totalizam 300 tipos de redes LVQ, que são processadas para cada um dos 7 períodos selecionados.

Aos 7 arquivos dos períodos selecionados foi acrescentado um arquivo que contém medidas de nível do rio Paraguai de janeiro de 1900 até 31 de julho 2007. Portanto, outras 300 redes LVQ foram treinadas com as mesmas configurações anteriores. Treinou-se ao todo, 2400 redes LVQ diferentes.

A Tabela 4.3 apresenta um exemplo de parte do conteúdo de um arquivo com rede LVQ já treinada. As colunas Linha e Coluna posicionam o neurônio na rede LVQ, isto é, como coordenadas X e Y de uma matriz bi-dimensional. As colunas valor peso 1, valor peso 2, ..., valor peso 8, correspondem aos valores peso de cada neurônio, que é representada na Tabela 4.3 como uma linha. A coluna qtde ocorrências informa o número de reconhecimentos realizados pelo neurônio na fase de treinamento e a coluna padrão informa a classe de nível de nível a que o neurônio pertence.

Tabela 4.3 - Exemplo do conteúdo de um arquivo de rede LVQ (média de 3 medidas, 8 médias por entrada, rede de 10x10 neurônios, 10 padrões utilizados)

a3\_d8\_m1010\_F10.txt

linha	coluna	valor peso 1	valor peso 2	valor peso 3	valor peso 4	valor peso 5	valor peso 6	valor peso 7	valor peso 8	qtde ocorrências	padrão
0	0	2,4835	2,3685	2,252	2,1357	2,0257	1,9216	1,8259	1,7404	137	F1.00
0	1	1,6382	1,7052	1,7948	1,8953	1,9971	2,0951	2,1872	2,2675	76	F1.81
0	2	0,8314	0,7634	0,6951	0,6274	0,5635	0,5016	0,4434	0,391	109	F0.20
0	3	0,5408	0,5661	0,5972	0,6328	0,67	0,7079	0,7451	0,7825	158	F0.20
0	4	2,2333	2,3099	2,3999	2,5045	2,6197	2,7395	2,8612	2,9749	82	F2.61
0	5	3,0728	3,0911	3,1091	3,1269	3,1435	3,159	3,1728	3,1864	101	F2.61
0	6	2,1906	2,2493	2,3155	2,3889	2,4629	2,5313	2,5956	2,657	139	F2.61
0	7	3,2554	3,3521	3,4577	3,5657	3,6845	3,8218	3,9825	4,1604	39	F3.42
0	8	4,1602	4,2669	4,3665	4,4599	4,546	4,6245	4,6963	4,7612	56	F4.22
0	9	3,6982	3,6339	3,5677	3,5022	3,4331	3,3561	3,2731	3,184	153	F2.61
1	0	0,9086	0,9143	0,9239	0,9373	0,9524	0,9707	0,992	1,0144	212	F1.00
1	1	0,1818	0,1561	0,1312	0,1048	0,0784	0,0531	0,0322	0,0181	121	F-0.61
1	2	-0,3893	-0,411	-0,4234	-0,4284	-0,426	-0,4167	-0,4012	-0,3806	95	F-0.61
1	3	2,4319	2,385	2,3362	2,2898	2,2469	2,2069	2,1728	2,1427	143	F1.81
1	4	3,6557	3,7175	3,7802	3,8476	3,921	3,9978	4,0757	4,1557	112	F3.42
1	5	2,915	3,0121	3,1014	3,1856	3,2692	3,355	3,4412	3,5291	114	F3.42
1	6	2,108	2,0678	2,029	1,993	1,9571	1,9229	1,8893	1,8608	148	F1.81
1	7	2,7669	2,8502	2,9422	3,0313	3,1144	3,1933	3,2719	3,3537	110	F2.61
1	8	2,1382	2,1739	2,214	2,2555	2,2962	2,3414	2,3912	2,4418	189	F1.81
1	9	4,3202	4,4989	4,6649	4,8145	4,9459	5,0633	5,1696	5,2609	34	F5.03
2	0	2,7059	2,6521	2,5911	2,5299	2,472	2,4202	2,3744	2,3314	130	F1.81
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
9	8	4,6411	4,8103	4,9601	5,0967	5,2188	5,3265	5,4222	5,5028	35	F5.03
9	9	3,5	3,62	3,7805	4,0224	4,3731	4,7533	5,0799	5,3308	18	F5.03

#### 4.3.2 Medida de Qualidade dos Neurônios da Rede LVQ

Em uma rede LVQ é interessante ter neurônios com uma medida de qualidade associada a ele. Isto é, se o conjunto dos neurônios dessa rede classifica bem os dados que são apresentados. Uma forma de medir essa qualidade pode ser a utilização do valor do desvio padrão de um neurônio em a relação a todas as séries temporais associadas a ele. O desvio padrão informa o grau de dispersão de dados em relação a uma média. Portanto, podem-se obter a distância que existe entre um ponto e os outros pontos de um conjunto, sendo esse ponto um neurônio e os outros pontos, os dados de treinamento.

Os resultados baixos de valores de desvio padrão indicam que os neurônios classificam bem as séries temporais que reconhece, portanto são bem coerentes e coesos. Se a rede LVQ tem muitos neurônios com valor de desvio padrão alto, essa rede não classifica bem os dados informados a ele.

Tabela 4.4 – Exemplo de arquivo para treinar rede LVQ

entrada_lvq_a3_d4_m1010_F10					
v1	v2	v3	v4	padrão	data
1,85	1,85	1,85	1,85	F1.81	19000112
1,85	1,85	1,85	1,87	F1.81	19000115
1,85	1,85	1,87	1,88	F1.81	19000118
1,85	1,87	1,88	1,97	F1.81	19000121
1,87	1,88	1,97	2,04	F1.81	19000124
1,88	1,97	2,04	2,11	F1.81	19000127
1,97	2,04	2,11	2,12	F1.81	19000130
2,04	2,11	2,12	2,13	F1.81	19000202
2,11	2,12	2,13	2,16	F1.81	19000205
2,12	2,13	2,16	2,13	F1.81	19000208
2,13	2,16	2,13	2,13	F1.81	19000211
2,16	2,13	2,13	2,15	F1.81	19000214
2,13	2,13	2,15	2,2	F1.81	19000217
2,35	2,45	2,52	2,58	F1.81	19000307
2,45	2,52	2,58	2,65	F2.61	19000310
2,52	2,58	2,65	2,72	F2.61	19000313
2,58	2,65	2,72	2,79	F2.61	19000316
2,65	2,72	2,79	2,84	F2.61	19000319
2,72	2,79	2,84	2,91	F2.61	19000322
...	...	...	...	...	...

Tabela 4.5 – Exemplo de arquivo gerado no treinamento da rede LVQ

lvq_a3_d4_m1010_F10							
L	C	p1	p2	p3	p4	qt	padrão
0	0	4,0269	4,1074	4,1842	4,2597	94	F4.22
0	1	4,3351	4,449	4,555	4,6528	69	F4.22
0	2	4,4946	4,615	4,7244	4,8211	63	F4.22
0	3	4,7036	4,7905	4,8672	4,9325	87	F4.22
0	4	3,7464	3,92	4,1543	4,4374	20	F4.22
0	5	3,8669	4,2118	4,6619	5,0491	11	F5.03
0	6	4,4568	4,8447	5,1739	5,4341	14	F5.03
0	7	5,9168	5,9202	5,9157	5,9029	112	F5.83
0	8	6,1385	6,147	6,146	6,1376	92	F5.83
0	9	6,4478	6,4574	6,4552	6,4435	78	F5.83
1	0	2,9351	2,9592	2,983	3,0081	120	F2.61
1	1	3,075	3,129	3,1802	3,2329	139	F2.61
1	2	4,3083	4,268	4,2282	4,1872	175	F3.42
1	3	4,1038	4,2429	4,3849	4,5227	60	F4.22
1	4	4,8936	4,9622	5,0212	5,0707	95	F5.03
1	5	5,0786	5,1495	5,2108	5,2633	66	F5.03
1	6	5,1126	5,3094	5,4668	5,6048	21	F5.03
1	7	5,3473	5,3407	5,3295	5,3127	154	F5.03
1	8	5,7082	5,6843	5,6568	5,6274	98	F5.03
...	...	...	...	...	...	...	...
8	3	1,8332	1,8502	1,8618	1,8746	182	F1.81
...	...	...	...	...	...	...	...
9	8	-0,1363	-0,1431	-0,1433	-0,1353	146	F-0.61
9	9	-0,4372	-0,4444	-0,4435	-0,4339	84	F-0.61

A Tabela 4.4 mostra uma parte do arquivo de dados de entrada do treinamento de uma rede LVQ. Cada linha é formada por um vetor de 4 elementos, cada um contém a média obtida de 3 medidas de nível do rio Paraguai. Na Tabela 4.5 é apresentado o exemplo de parte do arquivo da rede LVQ de 10x10 neurônios com 4 valores-peso em cada neurônio, gerado pelo treinamento que utilizou o arquivo da Tabela 4.4.

A qualidade dos neurônios foi obtida através dos processamentos que calcularam os valores de desvio padrão através a aplicação da fórmula:

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} \quad (4.1)$$

$X_i$  - representa os valores de cada vetor do arquivo utilizado para treinamento  
(v1, v2, v3, v4 da Tabela 4.4)

$\bar{X}$  - representa os valores de cada vetor do arquivo gerado no treinamento (p1, p2, p3, p4 da Tabela 4.5)

$n$  - número de elementos por vetor

Cada informação de entrada foi processada para se obter o valor do desvio padrão entre seus valores e os valores de cada um dos neurônios da rede LVQ já treinada. Esses valores de desvios padrão foram armazenados e acumulados de forma a representar a soma dos desvios padrão de cada neurônio da rede LVQ.

Ao final do processamento, um arquivo foi gerado com os valores acumulados por cada neurônio, o maior valor de desvio padrão obtido de cada neurônio e o valor total acumulado por toda a rede LVQ.

De cada linha do arquivo utilizado para o treinamento (Tabela 4.4), os valores v1, v2, v3 e v4 foram apresentados ao arquivo gerado no treinamento da rede LVQ, isto é, para cada neurônio que tem os valores-peso p1, p2, p3 e p4. O cálculo do desvio padrão é exemplificado na Tabela 4.6.

Tabela 4.6 – Exemplo do cálculo do desvio padrão.

n	$X_i$	$\bar{X}$	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$
4	$v_n - p_n$	$X_i / n$		
1	0,0168	-0,00495	0,02175	0,00047306
2	-0,0002	-0,00495	0,00475	2,2562E-05
3	-0,0118	-0,00495	-0,00685	4,6922E-05
4	-0,0246	-0,00495	-0,01965	0,00038612
soma	-0,0198		soma	0,00092867
			div / n	0,00023217
			raiz	0,01523704

Na Tabela 4.6 vemos o exemplo do desvio padrão da primeira linha do arquivo de entrada. O neurônio que corresponde à entrada é o da linha 8 e coluna 3

(8,3), apresentada na Tabela 4.7. A coluna n mostra o número do elemento de cada vetor, de entrada e do vetor peso, ambos com mesma quantidade de elementos.

A coluna Xi apresenta os valores da diferença entre os valores do vetor (v) do arquivo de entrada e os valores do vetor peso (p) do arquivo de rede LVQ. A coluna X apresenta o valor da média dos valores de Xi, isto é, a somatória de todos os valores cálculos por v menos p e depois dividido por 4 (número de valores v e p). A coluna (Xi – X) apresenta o valor da Xi subtraído de X. A coluna (Xi – X) \*\* 2 apresenta o quadrado do valor da coluna (Xi-X). Por fim, dos valores da ultima coluna é obtido a média da qual é calculada a raiz quadrada, resultando no valor do desvio padrão dessa primeira linha do arquivo de entrada.

Dessa forma é obtido o valor do desvio padrão para todas as linhas do arquivo de entrada, lembrando que os neurônios “mortos”, ou seja, que não há ocorrências de semelhança com os dados de entrada, são ignorados para o cálculo do desvio padrão.

Tabela 4.7 – Exemplo dos resultados do cálculo da qualidade da rede LVQ.

qualid_res_prev_lvq_teste_a3_d4_m1010_F10.txt				
Y= linha	X= coluna	Padrão	Qt Ocorrências na Rede	Desvio Padrão
0	0	F4.22	94	5,6241
0	1	F4.22	69	4,3289
0	2	F4.22	63	3,9258
0	3	F4.22	87	5,2938
0	4	F4.22	20	1,8946
0	5	F5.03	11	1,7214
0	6	F5.03	14	2,2659
0	7	F5.83	112	9,625
0	8	F5.83	92	8,7616
0	9	F5.83	78	7,7216
1	0	F2.61	120	4,614
1	1	F2.61	139	6,4447
1	2	F3.42	175	6,4213
1	3	F4.22	60	4,1812
1	4	F5.03	95	5,2867
1	5	F5.03	66	4,0037
1	6	F5.03	21	2,086
1	7	F5.03	154	7,8923
1	8	F5.03	98	6,6777
1	9	F5.03	144	7,3779
2	0	F2.61	82	4,4945
2	1	F2.61	121	5,4875
2	2	F2.61	158	7,5901
2	3	F4.22	86	4,0359
2	4	F2.61	96	3,8319
2	5	F3.42	135	4,5266
2	6	F3.42	163	5,8815
...	...	...	...	...

A Tabela 4.7 apresenta um exemplo dos resultados do processamento da qualidade dos neurônios da rede LVQ de 10 por 10 neurônios. As colunas Y e X contêm a posição do neurônio na rede LVQ, em linhas e colunas. A coluna “Padrão” contém o padrão que o neurônio representa; a coluna “Qt. Ocorrências na Rede” apresenta o número de correspondências que o neurônio reconheceu no treinamento. A coluna “Desvio Padrão” apresenta a soma dos desvios padrão de cada neurônio.

Os valores de desvio padrão calculados mostraram sua utilidade representando bem a semelhança entre os valores de entrada e valores-peso dos neurônios vencedores.

Os valores de desvio padrão da rede 10x10 são maiores que os das redes 20x20 e 30x30. Esse resultado é porque em redes menores os neurônios reconhecem um conjunto maior de dados de entrada, classificando-os de forma mais generalizada. Portanto, o resultado do cálculo do desvio padrão tenderá a ser maior que nas redes maiores. Nas redes maiores, os valores de desvio padrão são menores mostrando que uma rede maior tem maiores possibilidades de existir neurônios melhor treinados e classificar de forma mais precisa.

Em algumas ocasiões o valor de desvio padrão foi zero, isto é, uma identificação perfeita entre o neurônio e o valor de entrada. Embora o resultado seja ótimo, não se pode afirmar que o neurônio consegue classificar perfeitamente os dados de entrada, pois há poucas ocorrências de desvio padrão zero. Para indicar uma representatividade melhor desses neurônios, o número de ocorrências que indica os reconhecimentos dos dados de entrada deveria ser maior, bem como o número de neurônios com esse desvio padrão também deveria ser maior. Esses valores sugerem que a rede seja coerente e que consegue realizar uma boa classificação dos dados de entrada.

#### **4.4 Previsões e Resultados obtidos**

A Tabela 4.8 e Tabela 4.9 mostram os resultados das previsões com erros 0%, nas diversas configurações de redes LVQ de 7 períodos utilizadas para previsão de um passo adiante por vez, em 4 passos consecutivos. A Tabela 4.8 apresenta as quantidades de redes LVQ com resultado erro 0% e a Tabela 4.9 mostra os valores percentuais de cada configuração de rede LVQ, em relação ao total de 301 redes em que os resultados foram de erro 0%. As ausências de valores indicam que não houve redes com este resultado. Na Tabela 4.8, a coluna Medidas para média informa o número de medidas utilizadas para o cálculo do valor de cada elemento do vetor de entrada. A coluna Entradas informa a quantidade de médias ou elementos do vetor de entrada. A coluna Total apresenta a quantidade de redes LVQ que resultaram em erro 0%, para cada linha da coluna Padrão discretizado. Esta coluna mostra os números de padrões utilizados pelas redes LVQ. Os títulos 10x10, 20x20, 30x30, 50x50 e 100x100 representam os tamanhos das redes LVQ utilizadas e suas colunas mostram as quantidades de redes LVQ com erro 0% para cada configuração de tamanho de rede.

Tabela 4.8 - Resultado com 0% erros, em quantidade de redes LVQ, dos 7 períodos.

Quantidades de Redes LVQ por resultados com 0% de erros - 7 períodos selecionados									
Número de	de			Tamanho da rede LVQ ( neurônios x neurônios)					
	Entradas	Total	Padrões discretizados	10x10	20x20	30x30	50x50	100x100	
Medidas para média	média de 3 medidas	4	22	10	3	4	5	5	5
			9	20	1	2	2	2	2
			10	30	2	2	2	2	2
			14	50	2	3	3	3	3
			12	100	3	3	3	3	3
			67		8	14	15	15	15
	8	14	10	2	2	3	3	4	
		12	20	2	2	2	3	3	
		10	30	2	2	2	2	2	
		15	50	2	4	3	3	3	
		12	100	2	2	2	3	3	
		63		10	12	12	14	15	
	12	14	10	3	3	2	3	3	
		11	20	2	1	2	3	3	
		13	30	3	1	3	3	3	
		17	50	3	3	3	4	4	
		11	100	2	1	2	3	3	
		66		13	9	12	16	16	
			196		31	35	39	45	46
média de 10 medidas	4	11	10	2	2	2	2	3	
		5	20	1	1	1	1	1	
		5	30	1	1	1	1	1	
		5	50	1	1	1	1	1	
		3	100	1	1	1			
	29		6	6	6	5	6		
	8		10						
			20						
			30						
			50						
			100						
	12		10						
		20							
		30							
		50							
		100							
		29		6	6	6	5	6	
média de 30 medidas	4	6	10	2	1	1	1	1	
		5	20		1	1	1	2	
		4	30		1	1	1	1	
		5	50		1	1	1	2	
		3	100		1	1	1	1	
		23		2	5	4	5	7	
	8	14	10	1	3	4	3	3	
		8	20		1	2	3	2	
		6	30		2		2	2	
		3	50			1	1	1	
		2	100			1		1	
		33		1	6	8	9	9	
	12	10	10	2	3	2	2	1	
		6	20	1	1	1	2	1	
		1	30			1			
		2	50			1	1		
		1	100		1				
		20		3	5	5	5	2	
		76		6	16	17	19	18	
Total de redes LVQ:		301		43	57	62	69	70	



Tabela 4.9 - Resultado com 0% erro, por percentual de redes LVQ, dos 7 períodos.

Percentual de Redes LVQ por resultados com 0% de erros - 7 períodos selecionados								
Número de		Tamanho da rede LVQ ( neurônios x neurônios)						
Medidas para média	Entradas	Total	Padrões discretizados	10x10	20x20	30x30	50x50	100x100
média de 3 medidas	4	7,31%	10	1,00%	1,33%	1,66%	1,66%	1,66%
		2,99%	20	0,33%	0,66%	0,66%	0,66%	0,66%
		3,32%	30	0,66%	0,66%	0,66%	0,66%	0,66%
		4,65%	50	0,66%	1,00%	1,00%	1,00%	1,00%
		3,99%	100	1,00%	1,00%	1,00%	1,00%	1,00%
	22,26%		2,66%		4,65%	4,98%	4,98%	4,98%
	8	4,65%	10	0,66%	0,66%	1,00%	1,00%	1,33%
		3,99%	20	0,66%	0,66%	0,66%	1,00%	1,00%
		3,32%	30	0,66%	0,66%	0,66%	0,66%	0,66%
		4,98%	50	0,66%	1,33%	1,00%	1,00%	1,00%
		3,99%	100	0,66%	0,66%	0,66%	1,00%	1,00%
	20,93%		3,32%		3,99%	3,99%	4,65%	4,98%
12	4,65%	10	1,00%	1,00%	0,66%	1,00%	1,00%	
	3,65%	20	0,66%	0,33%	0,66%	1,00%	1,00%	
	4,32%	30	1,00%	0,33%	1,00%	1,00%	1,00%	
	5,65%	50	1,00%	1,00%	1,00%	1,33%	1,33%	
	3,65%	100	0,66%	0,33%	0,66%	1,00%	1,00%	
21,93%		4,32%		2,99%	3,99%	5,32%	5,32%	
65,12%		10,30%		11,63%	12,96%	14,95%	15,28%	
média de 10 medidas	4	3,65%	10	0,66%	0,66%	0,66%	0,66%	1,00%
		1,66%	20	0,33%	0,33%	0,33%	0,33%	0,33%
		1,66%	30	0,33%	0,33%	0,33%	0,33%	0,33%
		1,66%	50	0,33%	0,33%	0,33%	0,33%	0,33%
		1,00%	100	0,33%	0,33%	0,33%	0,33%	0,33%
	9,63%		1,99%		1,99%	1,99%	1,66%	1,99%
	8		10					
			20					
			30					
			50					
			100					
	12		10					
		20						
		30						
		50						
		100						
9,63%		1,99%		1,99%	1,99%	1,66%	1,99%	
média de 30 medidas	4	1,99%	10	0,66%	0,33%	0,33%	0,33%	0,33%
		1,66%	20		0,33%	0,33%	0,33%	0,66%
		1,33%	30		0,33%	0,33%	0,33%	0,33%
		1,66%	50		0,33%	0,33%	0,33%	0,66%
		1,00%	100		0,33%	0,33%	0,33%	0,33%
	7,64%		0,66%		1,66%	1,33%	1,66%	2,33%
	8	4,65%	10	0,33%	1,00%	1,33%	1,00%	1,00%
		2,66%	20		0,33%	0,66%	1,00%	0,66%
		1,99%	30		0,66%		0,66%	0,66%
		1,00%	50			0,33%	0,33%	0,33%
		0,66%	100			0,33%	0,33%	0,33%
	10,96%		0,33%		1,99%	2,66%	2,99%	2,99%
12	3,32%	10	0,66%	1,00%	0,66%	0,66%	0,33%	
	1,99%	20	0,33%	0,33%	0,33%	0,66%	0,33%	
	0,33%	30			0,33%			
	0,66%	50			0,33%	0,33%		
	0,33%	100		0,33%				
6,64%		1,00%		1,66%	1,66%	0,66%	0,66%	
25,25%		1,99%		5,32%	5,65%	6,31%	5,98%	
Total de redes LVQ:		100,00%		14,29%	18,94%	20,60%	22,92%	23,26%

A leitura da Tabela 4.8, para a primeira linha pode ser feita da forma: “das redes LVQ geradas por 3 medidas para cálculo de média, 4 médias por entrada e utilização de 10 padrões discretizados, foram obtidas 22 redes com erro 0% para previsão de um passo adiante, por 4 passos consecutivos. Os resultados estão distribuídos em 3 redes com tamanho 10x10, 4 para 20x20, 5 com tamanho 30x30, 5 para rede 50x50 e 5 redes com tamanho 100x100 neurônios”. E a leitura dos valores totais da Tabela 4.8 pode ser feita como: “das 301 redes LVQ com erro 0% para previsão de um passo adiante, por 4 passos consecutivos, 43 delas ocorreram com rede de 10x10 neurônios, 57 em rede de 20x20, 62 em rede de 30x30, 69 com tamanho 50x50 neurônios e 70 em redes de 100x100 neurônios”.

Para a Tabela 4.9 aplica-se a mesma leitura, porém em percentual de cada configuração em relação ao total de 301 redes LVQ com erro 0%.

As Tabelas 4.10, 4.11, 4.12 e 4.13, apresentam os resultados por tipo de configuração das redes LVQ das previsões de um passo adiante, por 4 passos consecutivos. Nas tabelas menores são mostrados os valores em quantidades e nas maiores, os valores detalhados por período selecionado, com apresentação percentual em relação ao total de 1575 redes LVQ processadas para previsão.

Tabela 4.10 - Resultados por número de medidas por média, redes LVQ dos 7 períodos.

Número de Medidas para Média							
% erros	a03		a10		a30		Total
0%	196	12,44%	29	1,84%	76	4,83%	301 19,11%
1900 jan a 1930 set	0		0		4		4
1900 jan a 1947 dez	73		23		27		123
1900 jan a 1959 abr	54		1		25		80
1900 jan a 1972 set	24		0		20		44
1900 jan a 1995 fev	31		0		0		31
1972 mai a 1995 fev	10		5		0		15
1983 abr a 1995 fev	4		0		0		4
> 0% e <= 17%	298	18,92%	145	9,21%	68	4,32%	511 32,44%
1900 jan a 1930 set	65		6		1		72
1900 jan a 1947 dez	2		12		15		29
1900 jan a 1959 abr	21		43		50		114
1900 jan a 1972 set	30		0		0		30
1900 jan a 1995 fev	44		25		2		71
1972 mai a 1995 fev	65		30		0		95
1983 abr a 1995 fev	71		29		0		100
> 17%	31	1,97%	351	22,29%	381	24,19%	763 48,44%
1900 jan a 1930 set	10		69		70		149
1900 jan a 1947 dez	0		40		33		73
1900 jan a 1959 abr	0		31		0		31
1900 jan a 1972 set	21		75		55		151
1900 jan a 1995 fev	0		50		73		123
1972 mai a 1995 fev	0		40		75		115
1983 abr a 1995 fev	0		46		75		121
Total	525		525		525		1575

Número de Medidas para Média				
% erros	a03	a10	a30	Total
0%	196	29	76	301
<= 17%	298	145	68	511
> 17%	31	351	381	763
Total	525	525	525	1575

A Tabela 4.10 mostra as quantidades de redes LVQ por número de medidas utilizadas na média (colunas a3, a10 e a30), pelos 7 períodos e por tipo de erro (coluna % erros).

A Tabela 4.11 mostra as quantidades de redes por número de entradas (colunas d04, d08 e d12), pelos 7 períodos e por tipo de erro (coluna % erros).

Tabela 4.11 - Resultados por número de pesos / neurônio, redes LVQ dos 7 períodos.

Número de Entradas							
% erros	d04		d08		d12		Total
0%	119	7,56%	96	6,10%	86	5,46%	301 19,11%
1900 jan a 1930 set	0		4		0		4
1900 jan a 1947 dez	67		28		28		123
1900 jan a 1959 abr	20		34		26		80
1900 jan a 1972 set	13		19		12		44
1900 jan a 1995 fev	5		6		20		31
1972 mai a 1995 fev	10		5		0		15
1983 abr a 1995 fev	4		0		0		4
> 0% e <= 17%	229	14,54%	165	10,48%	117	7,43%	511 32,44%
1900 jan a 1930 set	30		27		15		72
1900 jan a 1947 dez	3		19		7		29
1900 jan a 1959 abr	55		35		24		114
1900 jan a 1972 set	10		10		10		30
1900 jan a 1995 fev	45		20		6		71
1972 mai a 1995 fev	40		25		30		95
1983 abr a 1995 fev	46		29		25		100
> 17%	177	11,24%	264	16,76%	322	20,44%	763 48,44%
1900 jan a 1930 set	45		44		60		149
1900 jan a 1947 dez	5		28		40		73
1900 jan a 1959 abr	0		6		25		31
1900 jan a 1972 set	52		46		53		151
1900 jan a 1995 fev	25		49		49		123
1972 mai a 1995 fev	25		45		45		115
1983 abr a 1995 fev	25		46		50		121
Total	525		525		525		1575

Número de Médias por Entrada / Registro				
% erros	d04	d08	d12	Total
0%	119	96	86	301
<= 17%	229	165	117	511
> 17%	177	264	322	763
Total	525	525	525	1575

A Tabela 4.12 mostra as quantidades de redes por padrão discretizados (colunas F010, F020, F030, F050 e F100), pelos 7 períodos e tipo de erros (coluna %erros).

Tabela 4.12 - Resultados por número de padrões utilizados nas redes LVQ, dos 7 períodos.

Padrões discretizados de Nível do Rio Paraguai							
% erros	F010	F020	F030	F050	F100	Total	
0%	91 5,78%	56 3,56%	49 3,11%	61 3,87%	44 2,79%	301	19,11%
1900 jan a 1930 set	4	0	0	0	0	4	
1900 jan a 1947 dez	29	26	25	23	20	123	
1900 jan a 1959 abr	23	10	17	19	11	80	
1900 jan a 1972 set	11	7	3	14	9	44	
1900 jan a 1995 fev	14	4	4	5	4	31	
1972 mai a 1995 fev	6	9	0	0	0	15	
1983 abr a 1995 fev	4	0	0	0	0	4	
> 0% e <= 17%	84 5,33%	121 7,68%	119 7,56%	88 5,59%	99 6,29%	511	32,44%
1900 jan a 1930 set	21	16	15	10	10	72	
1900 jan a 1947 dez	0	12	7	4	6	29	
1900 jan a 1959 abr	16	30	22	19	27	114	
1900 jan a 1972 set	0	15	15	0	0	30	
1900 jan a 1995 fev	7	17	16	15	16	71	
1972 mai a 1995 fev	24	11	20	20	20	95	
1983 abr a 1995 fev	16	20	24	20	20	100	
> 17%	140 8,89%	138 8,76%	147 9,33%	166 10,54%	172 10,92%	763	48,44%
1900 jan a 1930 set	20	29	30	35	35	149	
1900 jan a 1947 dez	16	7	13	18	19	73	
1900 jan a 1959 abr	6	5	6	7	7	31	
1900 jan a 1972 set	34	23	27	31	36	151	
1900 jan a 1995 fev	24	24	25	25	25	123	
1972 mai a 1995 fev	15	25	25	25	25	115	
1983 abr a 1995 fev	25	25	21	25	25	121	
Total	315	315	315	315	315	1575	

Padrões de Nível do Rio Paraguai						
% erros	F010	F020	F030	F050	F100	Total
0%	91	56	49	61	44	301
<= 17%	84	121	119	88	99	511
> 17%	140	138	147	166	172	763
Total	315	315	315	315	315	1575

A Tabela 4.13 mostra as quantidades de redes por tamanho (colunas 10x10, 20x20, 30x30, 50x50 e 100x100 neurônios), pelos 7 períodos e por tipo de erro (coluna %erros).

Tabela 4.13 - Resultados por tamanho das redes LVQ (número de neurônios), redes LVQ dos 7 períodos.

Tamanho de rede LVQ ( neurônios X neurônios )						
% erros	10x10	20x20	30x30	50x50	100x100	Total
0%	43 2,73%	57 3,62%	62 3,94%	69 4,38%	70 4,44%	301 19,11%
1900 jan a 1930 set	0	1	1	1	1	4
1900 jan a 1947 dez	22	27	25	25	24	123
1900 jan a 1959 abr	12	11	13	21	23	80
1900 jan a 1972 set	1	10	11	11	11	44
1900 jan a 1995 fev	5	4	8	7	7	31
1972 mai a 1995 fev	3	3	3	3	3	15
1983 abr a 1995 fev	0	1	1	1	1	4
> 0% e <= 17%	103 6,54%	110 6,98%	105 6,67%	97 6,16%	96 6,10%	511 32,44%
1900 jan a 1930 set	14	16	14	14	14	72
1900 jan a 1947 dez	4	6	7	6	6	29
1900 jan a 1959 abr	24	27	27	19	17	114
1900 jan a 1972 set	6	6	6	6	6	30
1900 jan a 1995 fev	16	16	12	13	14	71
1972 mai a 1995 fev	19	19	19	19	19	95
1983 abr a 1995 fev	20	20	20	20	20	100
> 17%	169 10,73%	148 9,40%	148 9,40%	149 9,46%	149 9,46%	763 48,44%
1900 jan a 1930 set	31	28	30	30	30	149
1900 jan a 1947 dez	19	12	13	14	15	73
1900 jan a 1959 abr	9	7	5	5	5	31
1900 jan a 1972 set	38	29	28	28	28	151
1900 jan a 1995 fev	24	25	25	25	24	123
1972 mai a 1995 fev	23	23	23	23	23	115
1983 abr a 1995 fev	25	24	24	24	24	121
Total	315	315	315	315	315	1575

Tamanho de rede LVQ						
% erros	m1010	m2020	m3030	m5050	m100100	Total
0%	43	57	62	69	70	301
< = 17%	103	110	105	97	96	511
> 17%	169	148	148	149	149	763
Total	315	315	315	315	315	1575

Após os 1575 processamentos de previsão, foi escolhida a rede LVQ de 30x30 neurônios, gerada e treinada utilizando 30 medidas para média, 12 entradas, 20 padrões discretizados com seus resultados para comparação aos do artigo mencionado (Tabela 4.1).

A Tabela 4.14 tem a mesma apresentação da Tabela 4.1 para facilitar a comparação. Em ambas as tabelas, são apresentados os resultados percentuais, por previsão para um passo, dois passos, três e quatro passos adiante. Observa-se que cada passo nesta previsão corresponde a um mês, e os valores previstos pela rede LVQ foram para a primeira medida de nível do mês a ser previsto.

Tabela 4.14 - Resumo dos resultados de previsão realizados pela rede LVQ 30x30 neurônios, treinada pelos parâmetros: 12 entradas, 30 medidas para média de cada entrada, 10 padrões discretizados.

Resultados obtidos pela Rede LVQ a30\_d12\_m3030\_F20

qtde de valores médios mensais	dados obtidos		para prever		Resultados obtidos				
	de	até	de	até					
274	mai/72	fev/95	mar/95	jun/95	erro relativo	mar/95 0,00%	abr/95 0,00%	mai/95 0,00%	jun/95 0,00%
576	jan/00	dez/47	jan/48	abr/48	erro relativo	jan/47 0,00%	fev/47 4,68%	mar/47 14,73%	abr/47 24,44%
369	jan/00	set/30	out/30	jan/31	erro relativo	out/30 0,00%	nov/30 0,00%	dez/30 0,00%	jan/31 0,00%
712	jan/00	abr/59	mai/59	ago/59	erro relativo	mai/59 68,18%	jun/59 0,00%	jul/59 14,00%	ago/59 45,00%
873	jan/00	set/72	out/72	jan/73	erro relativo	out/72 29,90%	nov/72 91,74%	dez/72 48,80%	jan/73 35,17%
143	abr/83	fev/95	mar/95	jun/95	erro relativo	mar/95 0,00%	abr/95 0,00%	mai/95 0,00%	jun/95 0,00%
1291	jan/00	fev/95	mar/95	jun/95	erro relativo	mar/95 23,97%	abr/95 48,40%	mai/95 42,01%	jun/95 29,00%

Estão destacados os valores obtidos pela rede LVQ em comparação aos do artigo em:

- resultados com erros 0%;
- resultados com percentual de erro abaixo do apresentado pelo artigo, porém acima de 0% e;
- os resultados com percentual de erro acima do apresentado pelo artigo.

#### **4.4.1 Outros Processamentos de Previsão**

Foram processados previsões para alguns dias do mês de Agosto de 2007 através das 300 redes LVQ treinadas com as medidas de 1900 até 2007 e diversas configurações.

Inicialmente, previsões para 1 dia, 2 dias, 3 dias, até para 21 dias, ou seja, 1 passo, 2 passos, 3 passos, até 21 passos, sempre com a data 31 de julho de 2007 como referencia e daí para adiante.

Nesses processamentos, o resultado de um passo é utilizado para o processamento do passo seguinte. Por exemplo, para a previsão de 4 passos consecutivos, o processamento realizou a previsão para 1º dia, o resultado foi acrescentado para a previsão do dia seguinte (2º dia) com descarte da medida mais “antiga”; o resultado desse processamento (2º dia) foi acrescentado para a previsão do 3º dia (com descarte da medida mais “antiga”); o resultado da previsão do 3º foi acrescentado para a previsão do 4º dia e por fim, o resultado da previsão do 4º dia que é o objetivo da previsão. E assim foi feito para todos os processamentos.

Os resumos dos resultados podem ser observados nas Tabelas 4.15, 4.16, 4.17, 4.18, 4.19 e 4.20, em que as quantidades de redes LVQ são apresentadas com suas configurações. Também são apresentados os resultados percentuais de cada configuração, das situações com erros 0%.



A Tabela 4.15 mostra os resultados da previsão para 1 passo adiante com as quantidades de redes LVQ por configuração.

Tabela 4.15 - Resultado Erros 0%, por quantidade de redes LVQ, previsão de 1 passo.

<b>Número de redes LVQ com resultado 0% de erros para previsão de 1 passo adiante</b>									
<b>Conjunto de medidas de janeiro de 1900 até julho de 2007</b>									
Medidas para média	Número de Entradas	Padrões discretizados	Tamanho da rede LVQ (neurônios x neurônios)					Total	
			10x10	20x20	30x30	50x50	100x100		
3 média de 3 medidas	4	10	1	1	1	1	1	5	
		20	1	1	1	1	1	5	
		30							
		50							
		100							
				<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>10</b>
	8	10						1	1
		20	1	1	1	1	1	1	5
		30							
		50							
100									
			<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>2</b>	<b>6</b>	
12	10								
	20	1	1	1	1	1		5	
	30								
	50								
	100								
			<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>5</b>	
			<b>4</b>	<b>4</b>	<b>4</b>	<b>4</b>	<b>5</b>	<b>21</b>	
10 média de 10 medidas	4	10							
		20							
		30							
		50							
		100							
	8	10							
		20							
		30							
		50							
		100							
	12	10							
		20							
30									
50									
100									
30 média de 30 medidas	4	10							
		20	1		1		1	3	
		30	1	1	1	1	1	5	
		50				1	1	2	
		100							
				<b>2</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>3</b>	<b>10</b>
	8	10						1	2
		20	1	1	1	1	1	1	4
		30		1	1	1			3
		50							
		100							
				<b>1</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>9</b>
12	10								
	20	1		1	1	1		3	
	30			1	1	1		3	
	50								
	100								
			<b>1</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>6</b>	
			<b>4</b>	<b>3</b>	<b>6</b>	<b>6</b>	<b>6</b>	<b>25</b>	
Total de redes LVQ:			<b>8</b>	<b>7</b>	<b>10</b>	<b>10</b>	<b>11</b>	<b>46</b>	

A Tabela 4.16 apresenta as quantidades percentuais em relação ao total de resultados com erros 0%.

Tabela 4.16 - Resultado Erros 0%, por percentual de redes LVQ, previsão 1 passo.

<b>Número de redes LVQ com resultado 0% de erros para previsão de 1 passo adiante</b>								
<b>Conjunto de medidas de janeiro de 1900 até julho de 2007</b>								
Número de		Tamanho da rede LVQ (neurônios x neurônios)						
Medidas para média	Entradas	Padrões discretizados	10x10	20x20	30x30	50x50	100x100	Total
3 média de 3 medidas	4	10	2,17%	2,17%	2,17%	2,17%	2,17%	10,87%
		20	2,17%	2,17%	2,17%	2,17%	2,17%	10,87%
		30						
		50						
		100						
			<b>4,35%</b>	<b>4,35%</b>	<b>4,35%</b>	<b>4,35%</b>	<b>4,35%</b>	<b>21,74%</b>
	8	10					2,17%	2,17%
		20	2,17%	2,17%	2,17%	2,17%	2,17%	10,87%
		30						
		50						
		100						
			<b>2,17%</b>	<b>2,17%</b>	<b>2,17%</b>	<b>2,17%</b>	<b>4,35%</b>	<b>13,04%</b>
12	10							
	20	2,17%	2,17%	2,17%	2,17%	2,17%	10,87%	
	30							
	50							
	100							
		<b>2,17%</b>	<b>2,17%</b>	<b>2,17%</b>	<b>2,17%</b>	<b>2,17%</b>	<b>10,87%</b>	
	<b>8,70%</b>	<b>8,70%</b>	<b>8,70%</b>	<b>8,70%</b>	<b>10,87%</b>	<b>45,65%</b>		
10 média de 10 medidas	4	10						
		20						
		30						
		50						
		100						
	8	10						
		20						
		30						
		50						
		100						
12	10							
	20							
	30							
	50							
	100							
30 média de 30 medidas	4	10						
		20	2,17%		2,17%		2,17%	6,52%
		30	2,17%	2,17%	2,17%	2,17%	2,17%	10,87%
		50				2,17%	2,17%	4,35%
		100						
			<b>4,35%</b>	<b>2,17%</b>	<b>4,35%</b>	<b>4,35%</b>	<b>6,52%</b>	<b>21,74%</b>
	8	10	2,17%				2,17%	4,35%
		20		2,17%	2,17%	2,17%	2,17%	8,70%
		30		2,17%	2,17%	2,17%		6,52%
		50						
		100						
			<b>2,17%</b>	<b>4,35%</b>	<b>4,35%</b>	<b>4,35%</b>	<b>4,35%</b>	<b>19,57%</b>
12	10							
	20	2,17%		2,17%	2,17%	2,17%	6,52%	
	30			2,17%	2,17%	2,17%	6,52%	
	50							
	100							
		<b>2,17%</b>		<b>4,35%</b>	<b>4,35%</b>	<b>2,17%</b>	<b>13,04%</b>	
	<b>8,70%</b>	<b>6,52%</b>	<b>13,04%</b>	<b>13,04%</b>	<b>13,04%</b>	<b>54,35%</b>		

As Tabelas 4.17, 4.18, 4.19 e 4.20 apresentam os resultados por tipo de configuração das redes LVQ, em valores de quantidade de redes LVQ e sua representação percentual em relação ao total de redes LVQ processadas na previsão de um passo adiante. Nas tabelas, em cada célula são apresentadas:

- na primeira linha, o valor da quantidade de redes LVQ que atendem ao tipo de erros e sua representação em percentual em relação ao total dos casos da linha da tabela;
- na segunda linha de cada célula, as representações em percentual dos valores das quantidades de redes LVQ em relação ao total da coluna da tabela;
- na última linha de cada célula, as representações em percentual dos valores da quantidade de redes LVQ em relação ao total geral da respectiva tabela (última célula das linhas e das colunas).

A Tabela 4.17 mostra as quantidades de redes LVQ por número de medidas utilizadas no cálculo da média (colunas a3, a10 e a30) e por tipo de erro (coluna % erros).

Tabela 4.17 - Resultados por número de medidas por média, para previsão de 1 passo adiante.

Número de Medidas para Média - de 1900 até 2007							
% erros	a03		a10		a30		Total
0%	21	28,00%	0	0,00%	25	33,33%	46
		45,65%		0,00%		54,35%	
		9,33%		0,00%		11,11%	
> 0% e <= 17%	54	72,00%	75	100,00%	50	66,67%	179
		30,17%		41,90%		27,93%	
		24,00%		33,33%		22,22%	
> 17%	0	0,00%	0	0,00%	0	0,00%	0
		0,00%		0,00%		0,00%	
		0,00%		0,00%		0,00%	
Total	75		75		75		225

A Tabela 4.18 mostra as quantidades de redes por número de entradas (colunas d04, d08 e d12) e por tipo de erro (coluna % erros).

Tabela 4.18 - Resultados por número de entradas, para previsão de 1 passo adiante.

Número de Entradas - de 1900 até 2007							
% erros	d04		d08		d12		Total
0%	20	26,67%	15	20,00%	11	14,67%	46
		43,48%		32,61%		23,91%	
		8,89%		6,67%		4,89%	
> 0% e <= 17%	55	73,33%	60	80,00%	64	85,33%	179
		30,73%		33,52%		35,75%	
		24,44%		26,67%		28,44%	
> 17%	0	0,00%	0	0,00%	0	0,00%	0
		0,00%		0,00%		0,00%	
		0,00%		0,00%		0,00%	
Total	75		75		75		225

A Tabela 4.19 mostra as quantidades de redes em número de padrões discretizados (colunas F010, F020, F030, F050 e F100) e por tipo de erros (coluna % erros).

Tabela 4.19 - Resultados por número de padrões utilizados nas redes LVQ, para previsão de 1 passo adiante.

Padrões discretizados de Nível do Rio Paraguai - de 1900 até 2007											
% erros	F010		F020		F030		F050		F100		Total
0%	8	17,78%	25	55,56%	11	24,44%	2	4,44%	0	0,00%	46
		17,39%		54,35%		23,91%		4,35%		0,00%	
		3,56%		11,11%		4,89%		0,89%		0,00%	
> 0% e <= 17%	37	82,22%	20	44,44%	34	75,56%	43	95,56%	45	100,00%	179
		20,67%		11,17%		18,99%		24,02%		25,14%	
		16,44%		8,89%		15,11%		19,11%		20,00%	
> 17%	0	0,00%	0	0,00%	0	0,00%	0	0,00%	0	0,00%	0
		0,00%		0,00%		0,00%		0,00%		0,00%	
		0,00%		0,00%		0,00%		0,00%		0,00%	
Total	45		45		45		45		45		225

A Tabela 4.20 mostra as quantidades de redes por tamanho (colunas 10x10, 20x20, 30x30, 50x50 e 100x100 neurônios) e tipo de erro (coluna % erros).

Tabela 4.20 - Resultados por tamanho das redes LVQ, para previsão de 1 passo adiante.

Tamanho de rede LVQ - de 1900 até 2007											
% erros	10x10 neurônios		20x20 neurônios		30x30 neurônios		50x50 neurônios		100x100 neurônios		Total
0%	8	17,78%	7	15,56%	10	22,22%	10	22,22%	11	24,44%	46
		17,39%		15,22%		21,74%		21,74%		23,91%	
		3,56%		3,11%		4,44%		4,44%		4,89%	
> 0% e <= 17%	37	82,22%	38	84,44%	35	77,78%	35	77,78%	34	75,56%	179
		20,67%		21,23%		19,55%		19,55%		18,99%	
		16,44%		16,89%		15,56%		15,56%		15,11%	
> 17%	0	0,00%	0	0,00%	0	0,00%	0	0,00%	0	0,00%	0
		0,00%		0,00%		0,00%		0,00%		0,00%	
		0,00%		0,00%		0,00%		0,00%		0,00%	
Total	45		45		45		45		45		225

Nas linhas de cada tabela são informados os percentuais em relação: ao total do tipo de erro, ao total da coluna e ao total de redes LVQ utilizados na previsão (225 redes).

São apresentados nas próximas tabelas, os resultados da previsão, em que a partir de uma data é previsto o nível de 21 passos adiante, com utilização apenas dos valores calculados e previstos pelo processamento. Foram processadas 4725 previsões, que é o valor acumulado para o processamento por passo (225 processamentos) em 21 passos de previsão (21 x 225 = 4725). A Tabela 4.21 apresenta os resultados desse processamento com 0% erros em 490 redes LVQ, classificadas pelas configurações que geraram essas redes.

Tabela 4.21 - Resultado com 0% erros, por quantidade de redes LVQ, para previsão de 21 passos adiante.

<b>Número de redes LVQ com resultado 0% de erros para previsão de 21 passos adiante</b>								
<b>Conjunto de medidas de janeiro de 1900 até julho de 2007</b>								
Número de		Tamanho da rede LVQ (neurônios x neurônios)						
Medidas para média	Entradas	Padrões discretizados	10x10	20x20	30x30	50x50	100x100	Total
3 média de 3 medidas	4	10	21	21	21	21	21	105
		20	7	7	7	7	7	35
		30						
		50						
		100						
			<b>28</b>	<b>28</b>	<b>28</b>	<b>28</b>	<b>28</b>	<b>140</b>
	8	10					21	21
		20	7	7	7	7	7	35
		30						
		50						
100								
		<b>7</b>	<b>7</b>	<b>7</b>	<b>7</b>	<b>28</b>	<b>56</b>	
12	10							
	20	7	7	7	7	7	35	
	30							
	50							
	100							
		<b>7</b>	<b>7</b>	<b>7</b>	<b>7</b>	<b>7</b>	<b>35</b>	
		<b>42</b>	<b>42</b>	<b>42</b>	<b>42</b>	<b>63</b>	<b>231</b>	
10 média de 10 medidas	4	10						
		20						
		30						
		50						
		100						
	8	10						
		20						
		30						
		50						
		100						
12	10							
	20							
	30							
	50							
	100							
30 média de 30 medidas	4	10		19				19
		20	1	6	2		2	11
		30	6	2	10	2	10	30
		50	7		2	4	4	17
		100	1		2			4
			<b>15</b>	<b>28</b>	<b>16</b>	<b>6</b>	<b>16</b>	<b>81</b>
	8	10	21	8	1	1	21	52
		20	13	4	5	3	3	28
		30	3	1	2	2	1	9
		50		1	1			2
100					1		1	
		<b>37</b>	<b>14</b>	<b>9</b>	<b>7</b>	<b>25</b>	<b>92</b>	
12	10	2	2	9	14	9	36	
	20	3	4	10	6	4	27	
	30			1	4	3	8	
	50	3		3	4		10	
	100		1	1	2	1	5	
		<b>8</b>	<b>7</b>	<b>24</b>	<b>30</b>	<b>17</b>	<b>86</b>	
		<b>60</b>	<b>49</b>	<b>49</b>	<b>43</b>	<b>58</b>	<b>259</b>	
Total de redes LVQ:			<b>102</b>	<b>91</b>	<b>91</b>	<b>85</b>	<b>121</b>	<b>490</b>

Na Tabela 4.22 os resultados são apresentados em quantidade de redes em valores percentuais de cada resultado do total de redes LVQ com 0% erro.

Tabela 4.22 - Resultado com 0% erros, por percentual de redes LVQ, para previsão de 21 passos adiante.

<b>Número de redes LVQ com resultado 0% de erros para previsão de 21 passos adiante</b>									
<b>Conjunto de medidas de janeiro de 1900 até julho de 2007</b>									
Medidas para média	Número de Entradas	Padrões discretizados	Tamanho da rede LVQ (neurônios x neurônios)					Total	
			10x10	20x20	30x30	50x50	100x100		
3 média de 3 medidas	4	10	4,29%	4,29%	4,29%	4,29%	4,29%	21,43%	
		20	1,43%	1,43%	1,43%	1,43%	1,43%	7,14%	
		30							
		50							
		100							
				<b>5,71%</b>	<b>5,71%</b>	<b>5,71%</b>	<b>5,71%</b>	<b>28,57%</b>	
	8	10					4,29%	4,29%	
		20	1,43%	1,43%	1,43%	1,43%	1,43%	7,14%	
		30							
		50							
		100							
				<b>1,43%</b>	<b>1,43%</b>	<b>1,43%</b>	<b>1,43%</b>	<b>5,71%</b>	
12	10								
	20	1,43%	1,43%	1,43%	1,43%	1,43%	7,14%		
	30								
	50								
	100								
			<b>1,43%</b>	<b>1,43%</b>	<b>1,43%</b>	<b>1,43%</b>	<b>7,14%</b>		
			<b>8,57%</b>	<b>8,57%</b>	<b>8,57%</b>	<b>8,57%</b>	<b>47,14%</b>		
10 média de 10 medidas	4	10							
		20							
		30							
		50							
		100							
	8	10							
		20							
		30							
		50							
		100							
	12	10							
		20							
30									
50									
100									
30 média de 30 medidas	4	10		3,88%				3,88%	
		20	0,20%	1,22%	0,41%		0,41%	2,24%	
		30	1,22%	0,41%	2,04%	0,41%	2,04%	6,12%	
		50	1,43%		0,41%	0,82%	0,82%	3,47%	
		100	0,20%	0,20%	0,41%			0,82%	
				<b>3,06%</b>	<b>5,71%</b>	<b>3,27%</b>	<b>1,22%</b>	<b>3,27%</b>	<b>16,53%</b>
	8	10	4,29%	1,63%	0,20%	0,20%	4,29%	10,61%	
		20	2,65%	0,82%	1,02%	0,61%	0,61%	5,71%	
		30	0,61%	0,20%	0,41%	0,41%	0,20%	1,84%	
		50		0,20%	0,20%			0,41%	
		100				0,20%		0,20%	
				<b>7,55%</b>	<b>2,86%</b>	<b>1,84%</b>	<b>1,43%</b>	<b>5,10%</b>	<b>18,78%</b>
12	10	0,41%	0,41%	1,84%	2,86%	1,84%	7,35%		
	20	0,61%	0,82%	2,04%	1,22%	0,82%	5,51%		
	30			0,20%	0,82%	0,61%	1,63%		
	50	0,61%		0,61%	0,82%		2,04%		
	100		0,20%	0,20%	0,41%	0,20%	1,02%		
			<b>1,63%</b>	<b>1,43%</b>	<b>4,90%</b>	<b>6,12%</b>	<b>3,47%</b>	<b>17,55%</b>	
			<b>12,24%</b>	<b>10,00%</b>	<b>10,00%</b>	<b>8,78%</b>	<b>11,84%</b>	<b>52,86%</b>	

As Tabelas 4.23, 4.24, 4.25 e 4.26 apresentam os resultados por configuração de redes LVQ, por tipo de erro e em percentual representativo em relação ao total de 4725 processamentos de previsão.

A Tabela 4.23 apresenta os resultados por configuração de número de medidas para média.

Tabela 4.23 - Resultados por número de medidas por média, para previsão de 21 passos adiante, consecutivos.

Quantidade de Redes LVQ por Número de Medidas para Média - de 1900 a 2007

% erros	a03		a10		a30		Total	
0%	231	14,67%	0	0,00%	259	16,44%	490	10,37%
		47,14%		0,00%		52,86%		
		4,89%		0,00%		5,48%		
> 0% e < 17%	1344	85,33%	1388	88,13%	903	57,33%	3635	76,93%
		36,97%		38,18%		24,84%		
		28,44%		29,38%		19,11%		
>= 17%	0	0,00%	187	11,87%	413	26,22%	600	12,70%
		0,00%		0,00%		0,00%		
		0,00%		3,96%		8,74%		
Total	1575		1575		1575		4725	

A Tabela 4.24 apresenta os resultados através do número de entradas.

Tabela 4.24 - Resultados por número de entradas, para previsão de 21 passos adiante, consecutivos.

Quantidade de Redes LVQ por Número de Entradas - de 1900 a 2007

% erros	d04		d08		d12		Total	
0%	221	14,03%	148	9,40%	121	7,68%	490	10,37%
		45,10%		30,20%		24,69%		
		4,68%		3,13%		2,56%		
> 0% e < 17%	1303	82,73%	1132	71,87%	1200	76,19%	3635	76,93%
		35,85%		31,14%		33,01%		
		27,58%		23,96%		25,40%		
>= 17%	51	3,24%	295	18,73%	254	16,13%	600	12,70%
		0,00%		0,00%		0,00%		
		1,08%		6,24%		5,38%		
Total	1575		1575		1575		4725	



A Tabela 4.21 apresenta os resultados por número de padrões utilizados e a Tabela 4.26 por tamanho das redes LVQ, em neurônios x neurônios.

Tabela 4.25 - Resultados por número de padrões utilizados nas redes LVQ, para previsão de 21 passos adiante, consecutivos.

Quantidade de Redes LVQ por Padrões de Nível do Rio Paraguai - de 1900 a 2007

% erros	F010		F020		F030		F050		F100		Total					
0%	233	24,66%	171	18,10%	47	4,97%	28	2,96%	11	1,16%	490	10,37%				
		47,55%				34,90%				9,59%				5,71%		2,24%
		4,93%				3,62%				0,99%				0,59%		0,23%
> 0% e < 17%	683	72,28%	691	73,12%	736	77,88%	739	78,20%	787	83,28%	3636	76,95%				
		18,78%				19,00%				20,24%				20,32%		21,64%
		14,46%				14,62%				15,58%				15,64%		16,66%
≥ 17%	29	3,07%	83	8,78%	162	17,14%	178	18,84%	147	15,56%	599	12,68%				
		0,00%				0,00%				0,00%				0,00%		0,00%
		0,61%				1,76%				3,43%				3,77%		3,11%
Total	945		945		945		945		945		4725					

Tabela 4.26 - Resultados por tamanho das redes LVQ, para previsão de 21 passos adiante, consecutivos.

Quantidade de Redes LVQ por Tamanho - de 1900 a 2007

% erros	10x10 neurônios		20x20 neurônios		30x30 neurônios		50x50 neurônios		100x100 neurônios		Total					
0%	102	10,79%	91	9,63%	91	9,63%	85	8,99%	121	12,80%	490	10,37%				
		20,82%				18,57%				17,35%				24,69%		
		2,16%				1,93%				1,80%				2,56%		
> 0% e < 17%	706	74,71%	724	76,61%	737	77,99%	745	78,84%	724	76,61%	3636	76,95%				
		19,42%				19,91%				20,27%				20,49%		19,91%
		14,94%				15,32%				15,60%				15,77%		15,32%
≥ 17%	137	14,50%	130	13,76%	117	12,38%	115	12,17%	100	10,58%	599	12,68%				
		0,00%				0,00%				0,00%				0,00%		0,00%
		2,90%				2,75%				2,48%				2,43%		2,12%
Total	945		945		945		945		945		4725					

#### 4.4.2 Outras configurações para previsões de Agosto / 2007

Dos resultados obtidos, foram realizadas outras previsões com redes LVQ escolhidas por terem gerados resultados erro 0% em maior número de ocasiões, conforme pode ser observado na Tabela 4.21 e Tabela 4.22.

As previsões foram realizadas utilizando as redes treinadas com as referências históricas (medidas anteriores a 31 de julho de 2007) informadas através de arquivos de tipo texto.

A Tabela 4.27 é um exemplo de referências históricas utilizadas para as previsões.

Tabela 4.27 - Exemplo de arquivo com referências para previsão, com 8 medidas

	entrada-1	entrada-2	entrada-3	entrada-4	entrada-5	entrada-6	entrada-7	entrada-8	padrão estimado
linha 1	4.5	4.45	4.4	4.36	4.3	4.25	4.2	4.16	F1.00
linha 2	4.45	4.4	4.36	4.3	4.25	4.2	4.16	4.11	F1.00
linha 3	4.4	4.36	4.3	4.25	4.2	4.16	4.11	4.09	F1.00
linha 4	4.36	4.3	4.25	4.2	4.16	4.11	4.09	4.07	F1.00
linha 5	4.3	4.25	4.2	4.16	4.11	4.09	4.07	4.04	F1.00
linha 6	4.25	4.2	4.16	4.11	4.09	4.07	4.04	4.02	F1.00
linha 7	4.2	4.16	4.11	4.09	4.07	4.04	4.02	4.0	F1.00
linha 8	4.16	4.11	4.09	4.07	4.04	4.02	4.0	3.98	F1.00
linha 9	4.11	4.09	4.07	4.04	4.02	4.0	3.98	3.94	F1.00
linha 10	4.09	4.07	4.04	4.02	4.0	3.98	3.94	3.93	F1.00
linha 11	4.07	4.04	4.02	4.0	3.98	3.94	3.93	3.92	F1.00
linha 12	4.04	4.02	4.0	3.98	3.94	3.93	3.92	3.89	F1.00
linha 13	4.02	4.0	3.98	3.94	3.93	3.92	3.89	3.86	F1.00
linha 14	4.0	3.98	3.94	3.93	3.92	3.89	3.86	3.83	F1.00
linha 15	3.98	3.94	3.93	3.92	3.89	3.86	3.83	3.81	F1.00
linha 16	3.94	3.93	3.92	3.89	3.86	3.83	3.81	3.78	F1.00
linha 17	3.93	3.92	3.89	3.86	3.83	3.81	3.78	3.75	F1.00
linha 18	3.92	3.89	3.86	3.83	3.81	3.78	3.75	3.72	F1.00
linha 19	3.89	3.86	3.83	3.81	3.78	3.75	3.72	3.70	F1.00
linha 20	3.86	3.83	3.81	3.78	3.75	3.72	3.70	3.67	F1.00
linha 21	3.83	3.81	3.78	3.75	3.72	3.70	3.67	3.63	F1.00
linha 22	3.81	3.78	3.75	3.72	3.70	3.67	3.63	3.60	F1.00
linha 23	3.78	3.75	3.72	3.70	3.67	3.63	3.60	3.58	F1.00

As colunas entrada-1, entrada-2,..., entrada-8, da Tabela 4.27, informam os valores utilizados como referência para o processamento da previsão, e que são com as medidas anteriores ao objetivo a ser previsto. Cada linha dessa tabela corresponde a uma previsão a ser realizada. Este arquivo do exemplo foi utilizado para previsão de um passo adiante, e com as referências deslocadas em um passo adiante a cada linha, resultou-se em previsões para 23 passos consecutivos. O processamento das previsões dos arquivos listados acima foi executado dessa forma e os resultados apresentados em forma de gráficos.

Os gráficos das figuras foram gerados pelos valores das colunas nível inferior e nível superior, que limitam as faixas de medidas dos padrões previstos pelos

processamentos. Os limites inferiores são caracterizados pela linha inferior (em vermelho), os limites superiores pela linha superior (em azul) e os valores reais de nível do rio Paraguai dos dias 01 até 23 de agosto, são caracterizados pela linha ao centro e com marcações (linha em verde).

As tabelas que acompanham cada gráfico apresentam os valores dos resultados de cada passo da previsão, que são: padrão previsto, o nível inferior e o nível superior do padrão previsto, os valores observados (reais) de cada passo ou dia de agosto previsto. A coluna desvio da média é preenchida com os valores da diferença entre os valores da coluna valor médio de nível (do padrão previsto) e o valor da coluna medida real (observada).

$$DesvioDaMédia = ValorMédioDoNível Previsto - Valor Real Observado \quad (4.2)$$

A coluna “quadrado do desvio da média” apresenta o resultado da operação valor da coluna Desvio da Média elevado à segunda potência (quadrado).

$$QuadradoDoDesvioDaMédia = (DesvioDaMédia)^2 \quad (4.3)$$

A soma dos valores dessa coluna resulta no valor da célula Variância.

$$Variância = \sum (DesvioDaMédia)^2 \quad (4.4)$$

O valor da célula Desvio Padrão é obtido pelo cálculo da Raiz Quadrada do valor da Variância e que pode ser obtida pela fórmula (4.1).

$$DesvioPadrão = \sqrt{(Variância)} \quad (4.5)$$

O valor da célula Erro Padrão da Média é o quociente do valor do Desvio Padrão pelo valor da Raiz Quadrada do número de elementos da tabela, que tem 23 elementos, isto é, 23 previsões.

$$\text{ErroPadrãoDaMédia} = \frac{\text{DesvioPadrão}}{\sqrt{\text{NúmeroDeElementos}}} \quad (4.6)$$

A Tabela 4.28 e a Figura 4.2 apresentam as previsões e os valores reais. O padrão discretizado F20 utilizado permite margem ampla de acertos. Ao final previu corretamente o padrão do nível do rio, que pode variar até 0,39m.

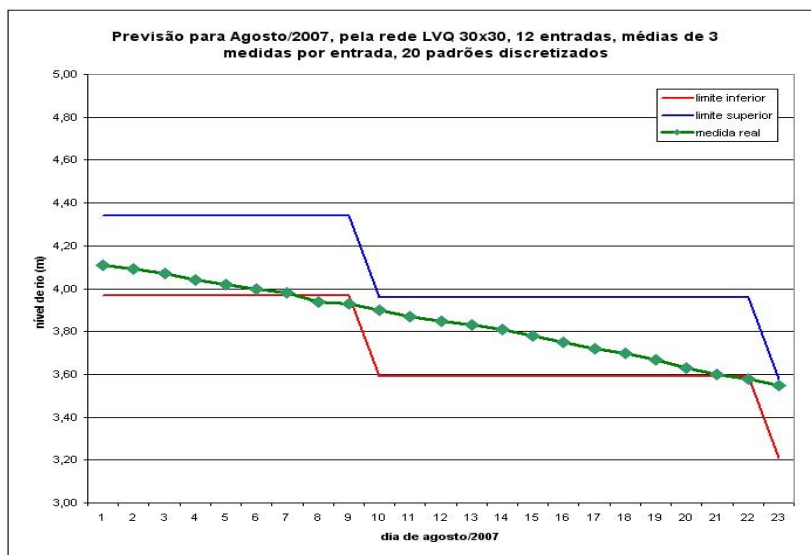


Figura 4.2 - Resultados da previsão da rede lvq\_a3\_d12\_m3030\_F20

Tabela 4.28 - Resultados da previsão da rede lvq\_a3\_d12\_m3030\_F20

dia	padrão	limite inferior	limite superior	nível médio	medida real	quadrado do desvio da média (medida real)	desvio da média (medida real)
1	F3.97	3,97	4,34	4,16	4,11	0,0020250	0,0450
2	F3.97	3,97	4,34	4,16	4,09	0,0042250	0,0650
3	F3.97	3,97	4,34	4,16	4,07	0,0072250	0,0850
4	F3.97	3,97	4,34	4,16	4,04	0,0132250	0,1150
5	F3.97	3,97	4,34	4,16	4,02	0,0182250	0,1350
6	F3.97	3,97	4,34	4,16	4,00	0,0240250	0,1550
7	F3.97	3,97	4,34	4,16	3,98	0,0306250	0,1750
8	F3.97	3,97	4,34	4,16	3,94	0,0462250	0,2150
9	F3.97	3,97	4,34	4,16	3,93	0,0506250	0,2250
10	F3.59	3,59	3,96	3,78	3,90	0,0156250	-0,1250
11	F3.59	3,59	3,96	3,78	3,87	0,0090250	-0,0950
12	F3.59	3,59	3,96	3,78	3,85	0,0056250	-0,0750
13	F3.59	3,59	3,96	3,78	3,83	0,0030250	-0,0550
14	F3.59	3,59	3,96	3,78	3,81	0,0012250	-0,0350
15	F3.59	3,59	3,96	3,78	3,78	0,0000250	-0,0050
16	F3.59	3,59	3,96	3,78	3,75	0,0006250	0,0250
17	F3.59	3,59	3,96	3,78	3,72	0,0030250	0,0550
18	F3.59	3,59	3,96	3,78	3,70	0,0056250	0,0750
19	F3.59	3,59	3,96	3,78	3,67	0,0110250	0,1050
20	F3.59	3,59	3,96	3,78	3,63	0,0210250	0,1450
21	F3.59	3,59	3,96	3,78	3,60	0,0306250	0,1750
22	F3.59	3,59	3,96	3,78	3,58	0,0380250	0,1950
23	F3.21	3,21	3,58	3,40	3,55	0,0240250	-0,1550
		DP 0,1260	Erro Padrão da Média 0,0263			Variância 0,0158685	

Nas Tabelas 4.29 e 4.30, com seus gráficos, os resultados das previsões foram semelhantes apesar das redes LVQ utilizadas terem configurações diferentes em quantidade de medidas por entrada, pois ainda utilizou-se 20 padrões discretizados.



Figura 4.3 - Resultados da previsão da rede lvq\_a3\_d8\_m3030\_F20

Tabela 4.29 - Resultados da previsão da rede lvq\_a3\_d8\_m3030\_F20

dia	padrão	limite inferior	limite superior	nível médio	medida real	quadrado do desvio da média (medida real)	desvio da média (medida real)
1	F3.97	3.97	4.34	4.16	4.11	0.0020250	0.0450
2	F3.97	3.97	4.34	4.16	4.09	0.0042250	0.0650
3	F3.97	3.97	4.34	4.16	4.07	0.0072250	0.0850
4	F3.97	3.97	4.34	4.16	4.04	0.0132250	0.1150
5	F3.97	3.97	4.34	4.16	4.02	0.0182250	0.1350
6	F3.97	3.97	4.34	4.16	4.00	0.0240250	0.1550
7	F3.97	3.97	4.34	4.16	3.98	0.0306250	0.1750
8	F3.59	3.59	3.96	3.78	3.94	0.0272250	-0.1650
9	F3.59	3.59	3.96	3.78	3.93	0.0240250	-0.1550
10	F3.59	3.59	3.96	3.78	3.90	0.0156250	-0.1250
11	F3.59	3.59	3.96	3.78	3.87	0.0090250	-0.0950
12	F3.59	3.59	3.96	3.78	3.85	0.0056250	-0.0750
13	F3.59	3.59	3.96	3.78	3.83	0.0030250	-0.0550
14	F3.59	3.59	3.96	3.78	3.81	0.0012250	-0.0350
15	F3.59	3.59	3.96	3.78	3.78	0.0000250	-0.0050
16	F3.59	3.59	3.96	3.78	3.75	0.0006250	0.0250
17	F3.59	3.59	3.96	3.78	3.72	0.0030250	0.0550
18	F3.59	3.59	3.96	3.78	3.70	0.0056250	0.0750
19	F3.59	3.59	3.96	3.78	3.67	0.0110250	0.1050
20	F3.59	3.59	3.96	3.78	3.63	0.0210250	0.1450
21	F3.59	3.59	3.96	3.78	3.60	0.0306250	0.1750
22	F3.21	3.21	3.58	3.40	3.58	0.0342250	-0.1850
23	F3.21	3.21	3.58	3.40	3.55	0.0240250	-0.1550
	DP	Erro Padrão da Média			Variância		
	0,1171	0,0244			0,0137207		

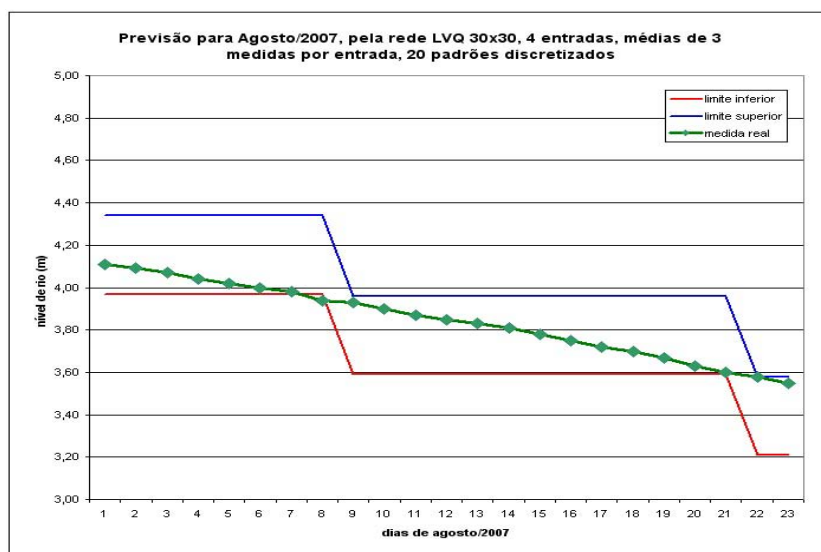


Figura 4.4 - Resultados da previsão da rede lvq\_a3\_d4\_m3030\_F20

Tabela 4.30 - Resultados da previsão da rede lvq\_a3\_d4\_m3030\_F20

dia	padrão	limite inferior	limite superior	nível médio	medida real	quadrado do desvio da média (medida real)	desvio da média (medida real)
1	F3.97	3,97	4,34	4,16	4,11	0,0020250	0,0450
2	F3.97	3,97	4,34	4,16	4,09	0,0042250	0,0650
3	F3.97	3,97	4,34	4,16	4,07	0,0072250	0,0850
4	F3.97	3,97	4,34	4,16	4,04	0,0132250	0,1150
5	F3.97	3,97	4,34	4,16	4,02	0,0182250	0,1350
6	F3.97	3,97	4,34	4,16	4,00	0,0240250	0,1550
7	F3.97	3,97	4,34	4,16	3,98	0,0306250	0,1750
8	F3.97	3,97	4,34	4,16	3,94	0,0462250	0,2150
9	F3.59	3,59	3,96	3,78	3,93	0,0240250	-0,1550
10	F3.59	3,59	3,96	3,78	3,90	0,0156250	-0,1250
11	F3.59	3,59	3,96	3,78	3,87	0,0090250	-0,0950
12	F3.59	3,59	3,96	3,78	3,85	0,0056250	-0,0750
13	F3.59	3,59	3,96	3,78	3,83	0,0030250	-0,0550
14	F3.59	3,59	3,96	3,78	3,81	0,0012250	-0,0350
15	F3.59	3,59	3,96	3,78	3,78	0,0000250	-0,0050
16	F3.59	3,59	3,96	3,78	3,75	0,0006250	0,0250
17	F3.59	3,59	3,96	3,78	3,72	0,0030250	0,0550
18	F3.59	3,59	3,96	3,78	3,70	0,0056250	0,0750
19	F3.59	3,59	3,96	3,78	3,67	0,0110250	0,1050
20	F3.59	3,59	3,96	3,78	3,63	0,0210250	0,1450
21	F3.59	3,59	3,96	3,78	3,60	0,0306250	0,1750
22	F3.21	3,21	3,58	3,40	3,58	0,0342250	-0,1850
23	F3.21	3,21	3,58	3,40	3,55	0,0240250	-0,1550
		DP 0,1206	Erro Padrão da Média 0,0251		Variância 0,0145467		

Diante dos resultados observados nos últimos gráficos, foram realizadas previsões para 29 dias e com o valor do parâmetro número de padrões maior: com 30, 50 e 100 padrões discretizados.

Os gráficos das figuras seguintes também foram gerados pelos valores das tabelas que as acompanham, pelas colunas nível inferior e nível superior, que limitam as faixas de medidas dos padrões previstos pelos processamentos. Também foi acrescentada a linha do valor médio de cada padrão discretizado em cada gráfico. Esse valor é obtido pela média entre o valor do nível inferior e nível superior e é representada pela linha pontilhada (de cor laranja) entre as linhas inferior e superior. Os limites inferiores são caracterizados pela linha inferior (em vermelho), os limites superiores pela linha superior (em verde) e os valores reais de nível do rio Paraguai dos dias 01 até 29 de agosto, são caracterizados pela linha ao centro e com marcações (linha em azul).

A Tabela 4.31 e sua Figura 4.5 apresentam os resultados em que a faixa com os valores do padrão discretizado utilizado é mais estreita, que representa 0,06m entre o nível inferior e o nível superior para 100 padrões discretizados. A previsão resultou em acerto do padrão do nível do rio, pois ao final dos 29 passos, a faixa com os limites inferior e superior acompanha o valor real observado. O mesmo ocorreu com a Tabela 4.32 e sua Figura 4.6, em que a previsão utilizou rede LVQ com quantidade de médias por entrada maior (8 médias por entrada), enquanto na previsão anterior utilizou-se rede LVQ com 4 médias por entrada.



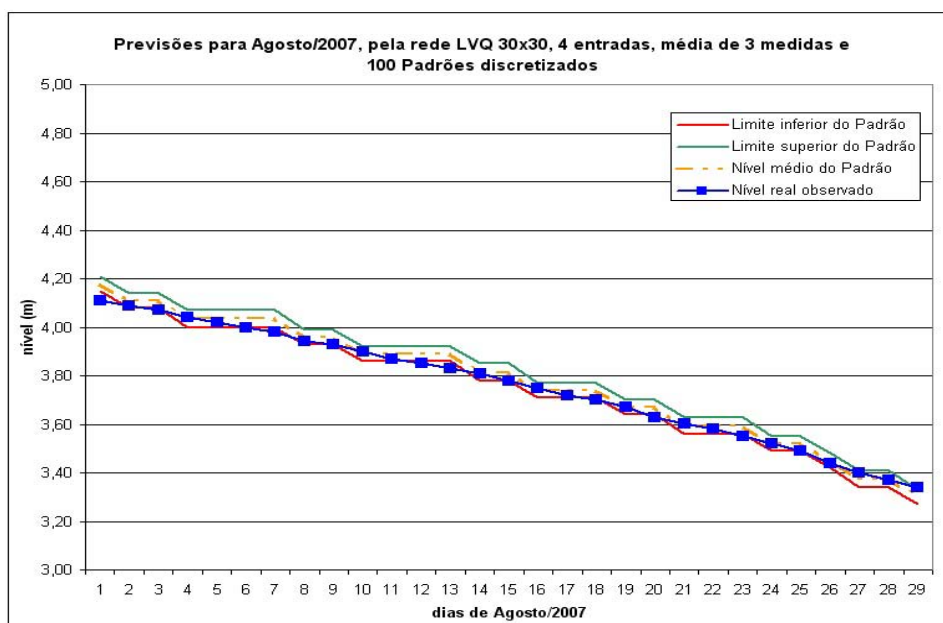


Figura 4.5 - Resultados da previsão da rede lvq\_a3\_d4\_m3030\_F100

Tabela 4.31 - Resultados da previsão da rede lvq\_a3\_d4\_m3030\_F100

dia	0,00	Limite inferior do Padrão	Limite superior do Padrão	nível médio	Nível real observado	quadrado do desvio da média (medida real)	desvio da média (medida real)
1	F4.15	4,15	4,21	4,18	4,11	0,0049000	0,070
2	F4.08	4,08	4,14	4,11	4,09	0,0004000	0,020
3	F4.08	4,08	4,14	4,11	4,07	0,0016000	0,040
4	F4.00	4,00	4,07	4,04	4,04	0,0000250	-0,005
5	F4.00	4,00	4,07	4,04	4,02	0,0002250	0,015
6	F4.00	4,00	4,07	4,04	4,00	0,0012250	0,035
7	F4.00	4,00	4,07	4,04	3,98	0,0030250	0,055
8	F3.93	3,93	3,99	3,96	3,94	0,0004000	0,020
9	F3.93	3,93	3,99	3,96	3,93	0,0009000	0,030
10	F3.86	3,86	3,92	3,89	3,90	0,0001000	-0,010
11	F3.86	3,86	3,92	3,89	3,87	0,0004000	0,020
12	F3.86	3,86	3,92	3,89	3,85	0,0016000	0,040
13	F3.86	3,86	3,92	3,89	3,83	0,0036000	0,060
14	F3.78	3,78	3,85	3,82	3,81	0,0000250	0,005
15	F3.78	3,78	3,85	3,82	3,78	0,0012250	0,035
16	F3.71	3,71	3,77	3,74	3,75	0,0001000	-0,010
17	F3.71	3,71	3,77	3,74	3,72	0,0004000	0,020
18	F3.71	3,71	3,77	3,74	3,70	0,0016000	0,040
19	F3.64	3,64	3,70	3,67	3,67	0,0000000	0,000
20	F3.64	3,64	3,70	3,67	3,63	0,0016000	0,040
21	F3.56	3,56	3,63	3,60	3,60	0,0000250	-0,005
22	F3.56	3,56	3,63	3,60	3,58	0,0002250	0,015
23	F3.56	3,56	3,63	3,60	3,55	0,0020250	0,045
24	F3.49	3,49	3,55	3,52	3,52	0,0000000	0,000
25	F3.49	3,49	3,55	3,52	3,49	0,0009000	0,030
26	F3.42	3,42	3,48	3,45	3,44	0,0001000	0,010
27	F3.34	3,34	3,41	3,38	3,40	0,0006250	-0,025
28	F3.34	3,34	3,41	3,38	3,37	0,0000250	0,005
29	F3.27	3,27	3,33	3,30	3,34	0,0016000	-0,040
		DP 0,0316	Erro Padrão da Média 0,0059			Variância 0,0009957	

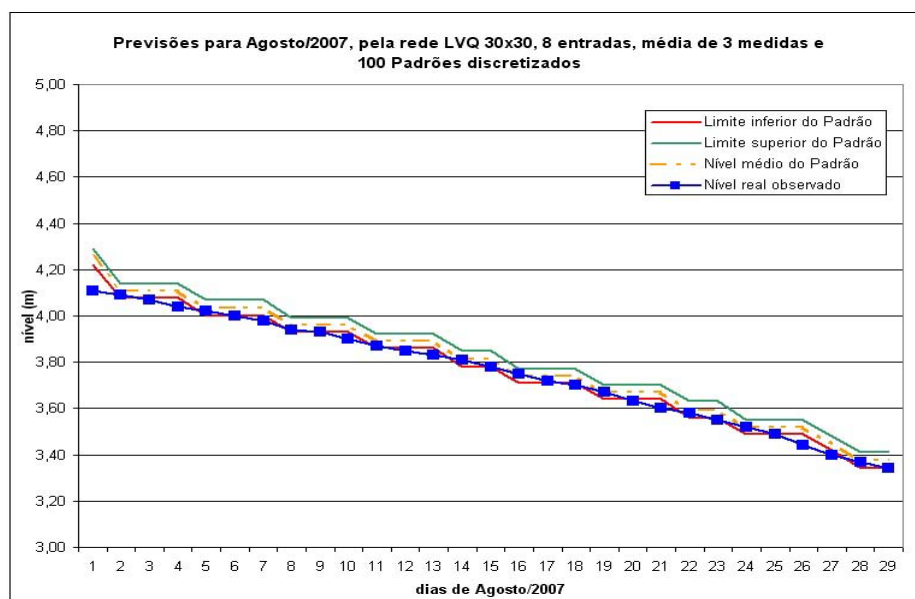


Figura 4.6 - Resultados da previsão da rede lvq\_a3\_d8\_m3030\_F100

Tabela 4.32 - Resultados da previsão da rede lvq\_a3\_d8\_m3030\_F100

dia	Padrão	Limite inferior do Padrão	Limite superior do Padrão	nível médio	Nível real observado	quadrado do desvio da média (medida real)	desvio da média (medida real)
1	F4.22	4,22	4,29	4,26	4,11	0,0210250	0,145
2	F4.08	4,08	4,14	4,11	4,09	0,0004000	0,020
3	F4.08	4,08	4,14	4,11	4,07	0,0016000	0,040
4	F4.08	4,08	4,14	4,11	4,04	0,0049000	0,070
5	F4.00	4,00	4,07	4,04	4,02	0,0002250	0,015
6	F4.00	4,00	4,07	4,04	4,00	0,0012250	0,035
7	F4.00	4,00	4,07	4,04	3,98	0,0030250	0,055
8	F3.93	3,93	3,99	3,96	3,94	0,0004000	0,020
9	F3.93	3,93	3,99	3,96	3,93	0,0009000	0,030
10	F3.93	3,93	3,99	3,96	3,90	0,0036000	0,060
11	F3.86	3,86	3,92	3,89	3,87	0,0004000	0,020
12	F3.86	3,86	3,92	3,89	3,85	0,0016000	0,040
13	F3.86	3,86	3,92	3,89	3,83	0,0036000	0,060
14	F3.78	3,78	3,85	3,82	3,81	0,0000250	0,005
15	F3.78	3,78	3,85	3,82	3,78	0,0012250	0,035
16	F3.71	3,71	3,77	3,74	3,75	0,0001000	-0,010
17	F3.71	3,71	3,77	3,74	3,72	0,0004000	0,020
18	F3.71	3,71	3,77	3,74	3,70	0,0016000	0,040
19	F3.64	3,64	3,70	3,67	3,67	0,0000000	0,000
20	F3.64	3,64	3,70	3,67	3,63	0,0016000	0,040
21	F3.64	3,64	3,70	3,67	3,60	0,0049000	0,070
22	F3.56	3,56	3,63	3,60	3,58	0,0002250	0,015
23	F3.56	3,56	3,63	3,60	3,55	0,0020250	0,045
24	F3.49	3,49	3,55	3,52	3,52	0,0000000	0,000
25	F3.49	3,49	3,55	3,52	3,49	0,0009000	0,030
26	F3.49	3,49	3,55	3,52	3,44	0,0064000	0,080
27	F3.42	3,42	3,48	3,45	3,40	0,0025000	0,050
28	F3.34	3,34	3,41	3,38	3,37	0,0000250	0,005
29	F3.34	3,34	3,41	3,38	3,34	0,0012250	0,035
		DP 0,0477	Erro Padrão da Média 0,0089			Variancia 0,0022776	

Com os resultados obtidos até este ponto, optou-se em processar previsões com tamanhos maiores de redes, com 50x50 e 100x100 neurônios, apesar da

Tabela 4.21 e da Tabela 4.22 não indicarem essas configurações entre os resultados com erros 0%. Por esses processamentos, os gráficos resultantes demonstraram resultados mais positivos. Convém observar que as Tabelas 4.21 e 4.22 apresentaram resultados resumidos da posição do processamento de previsão do passo 21, e que por isso essas configurações não constaram como erros 0%.

A Figura 4.7 e a Tabela 4.33 apresentam as previsões processadas utilizando rede LVQ de tamanho 100 x 100 neurônios e 8 elementos por entrada. Nesta configuração, os resultados são semelhantes às previsões anteriores. Na Tabela 4.34 e o gráfico da Figura 4.8 apresentam pequena diferença, principalmente nas previsões dos passos 1 a 3, mas não refletiu no resultado para o passo 29 que previu o padrão com os limites das faixas de nível.

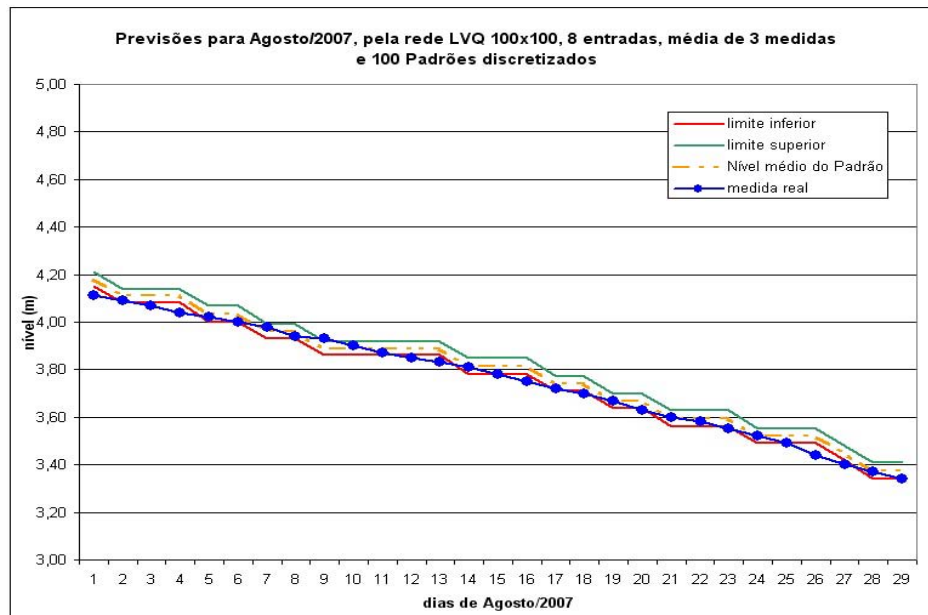


Figura 4.7 - Resultados da previsão da rede lvq\_a3\_d8\_m100100\_F100

Tabela 4.33 - Resultados da previsão da rede lvq\_a3\_d8\_m100100\_F100

dia	Padrão	Limite inferior do Padrão	Limite superior do Padrão	nível médio	Nível real observado	quadrado do desvio da média (medida real)	desvio da média (medida real)
1	F4.15	4,15	4,21	4,18	4,11	0,0049000	0,070
2	F4.08	4,08	4,14	4,11	4,09	0,0004000	0,020
3	F4.08	4,08	4,14	4,11	4,07	0,0016000	0,040
4	F4.08	4,08	4,14	4,11	4,04	0,0049000	0,070
5	F4.00	4,00	4,07	4,04	4,02	0,0002250	0,015
6	F4.00	4,00	4,07	4,04	4,00	0,0012250	0,035
7	F3.93	3,93	3,99	3,96	3,98	0,0004000	-0,020
8	F3.93	3,93	3,99	3,96	3,94	0,0004000	0,020
9	F3.86	3,86	3,92	3,89	3,93	0,0016000	-0,040
10	F3.86	3,86	3,92	3,89	3,90	0,0001000	-0,010
11	F3.86	3,86	3,92	3,89	3,87	0,0004000	0,020
12	F3.86	3,86	3,92	3,89	3,85	0,0016000	0,040
13	F3.86	3,86	3,92	3,89	3,83	0,0036000	0,060
14	F3.78	3,78	3,85	3,82	3,81	0,0000250	0,005
15	F3.78	3,78	3,85	3,82	3,78	0,0012250	0,035
16	F3.78	3,78	3,85	3,82	3,75	0,0042250	0,065
17	F3.71	3,71	3,77	3,74	3,72	0,0004000	0,020
18	F3.71	3,71	3,77	3,74	3,70	0,0016000	0,040
19	F3.64	3,64	3,70	3,67	3,67	0,0000000	0,000
20	F3.64	3,64	3,70	3,67	3,63	0,0016000	0,040
21	F3.56	3,56	3,63	3,60	3,60	0,0000250	-0,005
22	F3.56	3,56	3,63	3,60	3,58	0,0002250	0,015
23	F3.56	3,56	3,63	3,60	3,55	0,0020250	0,045
24	F3.49	3,49	3,55	3,52	3,52	0,0000000	0,000
25	F3.49	3,49	3,55	3,52	3,49	0,0009000	0,030
26	F3.49	3,49	3,55	3,52	3,44	0,0064000	0,080
27	F3.42	3,42	3,48	3,45	3,40	0,0025000	0,050
28	F3.34	3,34	3,41	3,38	3,37	0,0000250	0,005
29	F3.34	3,34	3,41	3,38	3,34	0,0012250	0,035
		DP 0,0388	Erro Padrão da Média 0,0072			Variância 0,0015086	

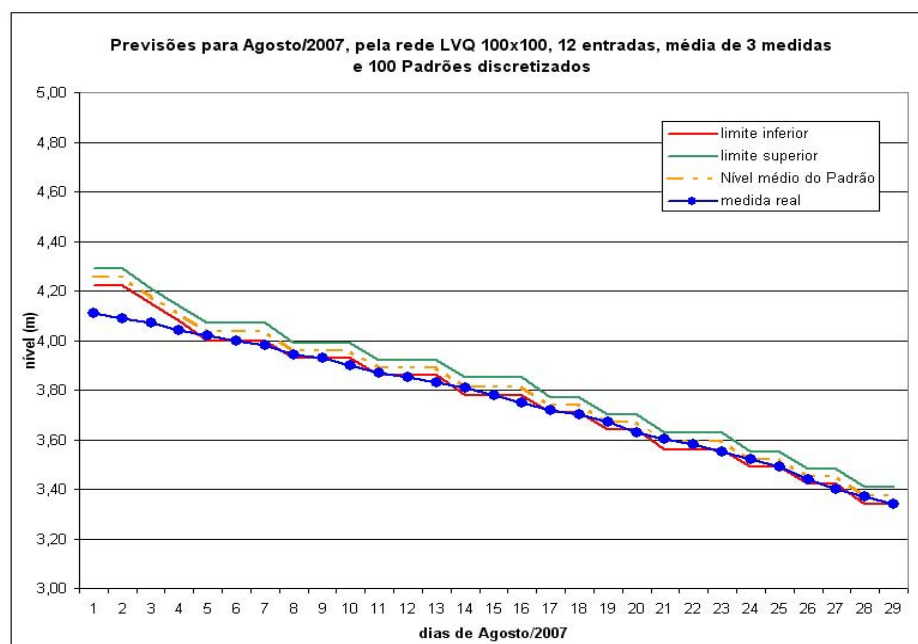


Figura 4.8 - Resultados da previsão da rede lvq\_a3\_d12\_m100100\_F100

Tabela 4.34 - Resultados da previsão da rede lvq\_a3\_d12\_m100100\_F100

dia	Padrão	Limite inferior do Padrão	Limite superior do Padrão	nível médio	Nível real observado	quadrado do desvio da média (medida real)	desvio da média (medida real)
1	F4.22	4,22	4,29	4,26	4,11	0,0210250	0,145
2	F4.22	4,22	4,29	4,26	4,09	0,0272250	0,165
3	F4.15	4,15	4,21	4,18	4,07	0,0121000	0,110
4	F4.08	4,08	4,14	4,11	4,04	0,0049000	0,070
5	F4.00	4,00	4,07	4,04	4,02	0,0002250	0,015
6	F4.00	4,00	4,07	4,04	4,00	0,0012250	0,035
7	F4.00	4,00	4,07	4,04	3,98	0,0030250	0,055
8	F3.93	3,93	3,99	3,96	3,94	0,0004000	0,020
9	F3.93	3,93	3,99	3,96	3,93	0,0009000	0,030
10	F3.93	3,93	3,99	3,96	3,90	0,0036000	0,060
11	F3.86	3,86	3,92	3,89	3,87	0,0004000	0,020
12	F3.86	3,86	3,92	3,89	3,85	0,0016000	0,040
13	F3.86	3,86	3,92	3,89	3,83	0,0036000	0,060
14	F3.78	3,78	3,85	3,82	3,81	0,0000250	0,005
15	F3.78	3,78	3,85	3,82	3,78	0,0012250	0,035
16	F3.78	3,78	3,85	3,82	3,75	0,0042250	0,065
17	F3.71	3,71	3,77	3,74	3,72	0,0004000	0,020
18	F3.71	3,71	3,77	3,74	3,70	0,0016000	0,040
19	F3.64	3,64	3,70	3,67	3,67	0,0000000	0,000
20	F3.64	3,64	3,70	3,67	3,63	0,0016000	0,040
21	F3.56	3,56	3,63	3,60	3,60	0,0000250	-0,005
22	F3.56	3,56	3,63	3,60	3,58	0,0002250	0,015
23	F3.56	3,56	3,63	3,60	3,55	0,0020250	0,045
24	F3.49	3,49	3,55	3,52	3,52	0,0000000	0,000
25	F3.49	3,49	3,55	3,52	3,49	0,0009000	0,030
26	F3.42	3,42	3,48	3,45	3,44	0,0001000	0,010
27	F3.42	3,42	3,48	3,45	3,40	0,0025000	0,050
28	F3.34	3,34	3,41	3,38	3,37	0,0000250	0,005
29	F3.34	3,34	3,41	3,38	3,34	0,0012250	0,035
		DP 0,0576	Erro Padrão da Média 0,0107			Variância 0,0033216	

A rede LVQ que gerou os valores do gráfico da Figura 4.8 foi processada para gerar a imagem do mapa dos neurônios dessa rede por cores que representam os padrões discretizados e por soma dos seus valores-peso, conforme as respectivas legendas. Também é mostrado o neurônio indicado pelo resultado da previsão do dia 29 (passo 29), em forma de gráfico de linha, das suas características, isto é, seus valores peso.

A Tabela 4.35 apresenta os resultados dessas previsões em outra forma de apresentação, na qual os neurônios dos resultados das previsões apresentam sua localização, em linha por coluna, no mapa da rede LVQ. A primeira coluna mostra o dia da previsão. A segunda coluna apresenta o neurônio que forneceu o resultado da previsão e sua posição no mapa da rede LVQ, em linha (L:yy) e coluna (C:xx). Neste exemplo, a rede nomeada como a3\_d12\_m100100\_F100 é uma rede LVQ de 100x100 neurônios, 100 padrões discretizados, 12 entradas e foi gerado por 3 medidas para média de cada entrada. A terceira coluna mostra o padrão previsto pela rede, a quarta e quinta colunas apresentam respectivamente os valores inferior e superior, da faixa de níveis que o padrão representa. A quinta coluna apresenta o nível médio do padrão e a última coluna mostra o nível real observado no respectivo dia da previsão.

Tabela 4.35 - Resultados da previsão da rede lvq\_a3\_d12\_m100100\_F100, com os neurônios e suas localizações

Resultados das previsões pela Rede LVQ a3\_d12\_m100100\_F100

	<b>Neurônio do resultado</b>	<b>Padrão</b>	<b>Limite inferior (m)</b>	<b>Limite superior (m)</b>	<b>Nível médio do Padrão (m)</b>	<b>Medida real (m)</b>
dia 01:	LVQ ( L:16 C:35 )	Padrão previsto= F4.22	de: 4,22	até: 4,29	4,26	4,11
dia 02:	LVQ ( L:15 C:36 )	Padrão previsto= F4.22	de: 4,22	até: 4,29	4,26	4,09
dia 03:	LVQ ( L:18 C:37 )	Padrão previsto= F4.15	de: 4,15	até: 4,21	4,18	4,07
dia 04:	LVQ ( L:16 C:38 )	Padrão previsto= F4.08	de: 4,08	até: 4,14	4,11	4,04
dia 05:	LVQ ( L:17 C:39 )	Padrão previsto= F4.00	de: 4,00	até: 4,07	4,04	4,02
dia 06:	LVQ ( L:21 C:37 )	Padrão previsto= F4.00	de: 4,00	até: 4,07	4,04	4,00
dia 07:	LVQ ( L:18 C:40 )	Padrão previsto= F4.00	de: 4,00	até: 4,07	4,04	3,98
dia 08:	LVQ ( L:20 C:40 )	Padrão previsto= F3.93	de: 3,93	até: 3,99	3,96	3,94
dia 09:	LVQ ( L:21 C:39 )	Padrão previsto= F3.93	de: 3,93	até: 3,99	3,96	3,93
dia 10:	LVQ ( L:23 C:39 )	Padrão previsto= F3.93	de: 3,93	até: 3,99	3,96	3,90
dia 11:	LVQ ( L:29 C:32 )	Padrão previsto= F3.86	de: 3,86	até: 3,92	3,89	3,87
dia 12:	LVQ ( L:27 C:37 )	Padrão previsto= F3.86	de: 3,86	até: 3,92	3,89	3,85
dia 13:	LVQ ( L:29 C:37 )	Padrão previsto= F3.86	de: 3,86	até: 3,92	3,89	3,83
dia 14:	LVQ ( L:22 C:42 )	Padrão previsto= F3.78	de: 3,78	até: 3,85	3,82	3,81
dia 15:	LVQ ( L:21 C:44 )	Padrão previsto= F3.78	de: 3,78	até: 3,85	3,82	3,78
dia 16:	LVQ ( L:31 C:34 )	Padrão previsto= F3.78	de: 3,78	até: 3,85	3,82	3,75
dia 17:	LVQ ( L:25 C:42 )	Padrão previsto= F3.71	de: 3,71	até: 3,77	3,74	3,72
dia 18:	LVQ ( L:34 C:35 )	Padrão previsto= F3.71	de: 3,71	até: 3,77	3,74	3,70
dia 19:	LVQ ( L:33 C:37 )	Padrão previsto= F3.64	de: 3,64	até: 3,70	3,67	3,67
dia 20:	LVQ ( L:35 C:36 )	Padrão previsto= F3.64	de: 3,64	até: 3,70	3,67	3,63
dia 21:	LVQ ( L:37 C:34 )	Padrão previsto= F3.56	de: 3,56	até: 3,63	3,60	3,60
dia 22:	LVQ ( L:33 C:40 )	Padrão previsto= F3.56	de: 3,56	até: 3,63	3,60	3,58
dia 23:	LVQ ( L:37 C:38 )	Padrão previsto= F3.56	de: 3,56	até: 3,63	3,60	3,55
dia 24:	LVQ ( L:40 C:37 )	Padrão previsto= F3.49	de: 3,49	até: 3,55	3,52	3,52
dia 25:	LVQ ( L:43 C:30 )	Padrão previsto= F3.49	de: 3,49	até: 3,55	3,52	3,49
dia 26:	LVQ ( L:45 C:33 )	Padrão previsto= F3.42	de: 3,42	até: 3,48	3,45	3,44
dia 27:	LVQ ( L:46 C:29 )	Padrão previsto= F3.42	de: 3,42	até: 3,48	3,45	3,40
dia 28:	LVQ ( L:46 C:32 )	Padrão previsto= F3.34	de: 3,34	até: 3,41	3,38	3,37
dia 29:	LVQ ( L:48 C:32 )	Padrão previsto= F3.34	de: 3,34	até: 3,41	3,38	3,34

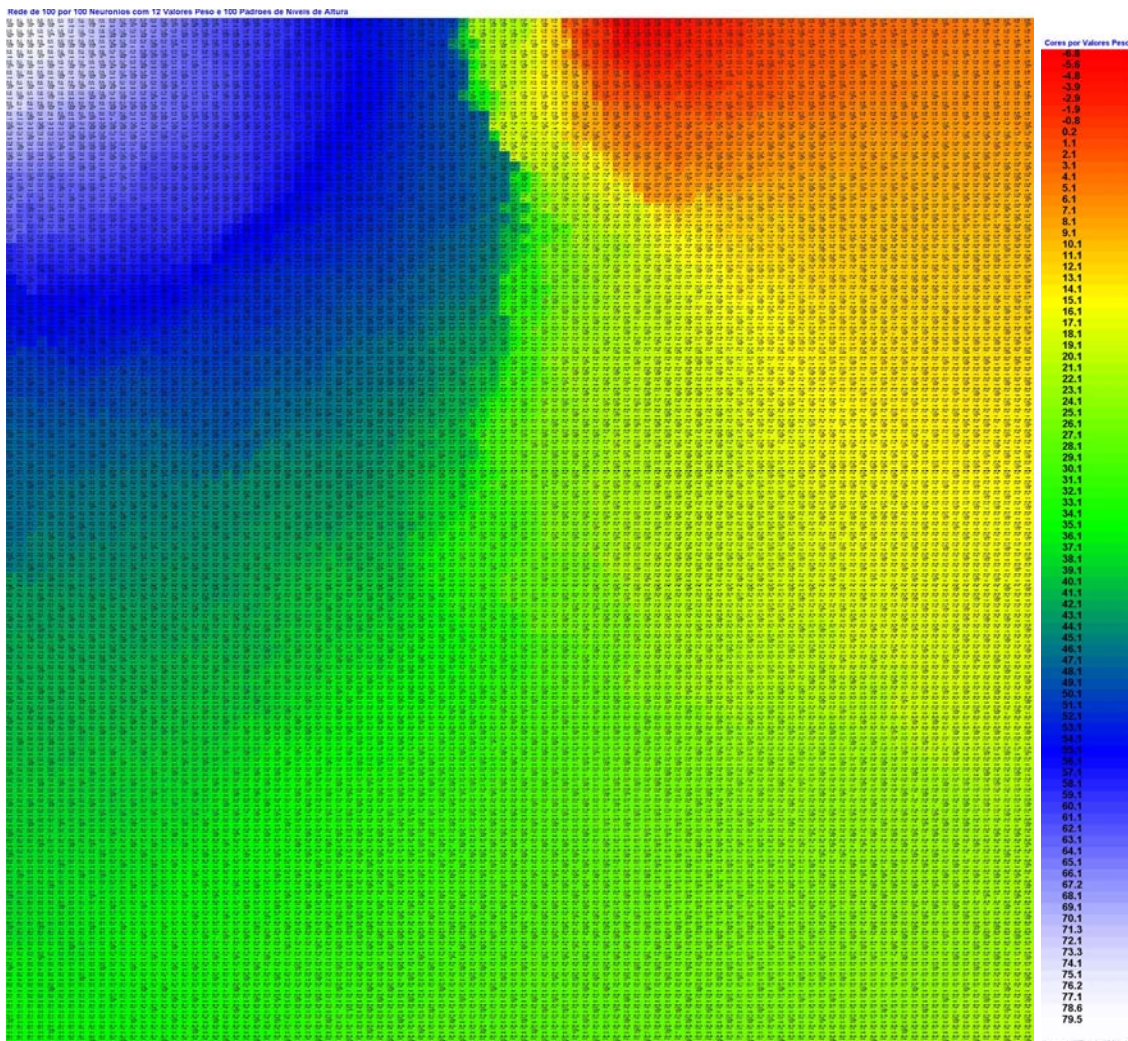


Figura 4.9 - Rede LVQ 100x100 neurônios, de 12 pesos por neurônio e 100 padrões discretizados e sua legenda de cores por soma dos valores dos pesos de cada neurônio.

A Figura 4.9 é a representação desta rede LVQ em um mapa de cores. Cada neurônio desse mapa da rede LVQ apresenta uma cor que representa o valor da soma dos seus valores peso, ou características do neurônio, conforme pode ser observado na legenda que acompanha a figura. A análise do mapa através das cores indica que os neurônios com a cor vermelha têm seus pesos com valores baixos, isto é, características de níveis baixos do rio. A cor azul ou branca indica valores altos dos pesos dos neurônios, ou características de níveis altos. E a cor verde, indicam valores médios dos pesos dos neurônios.



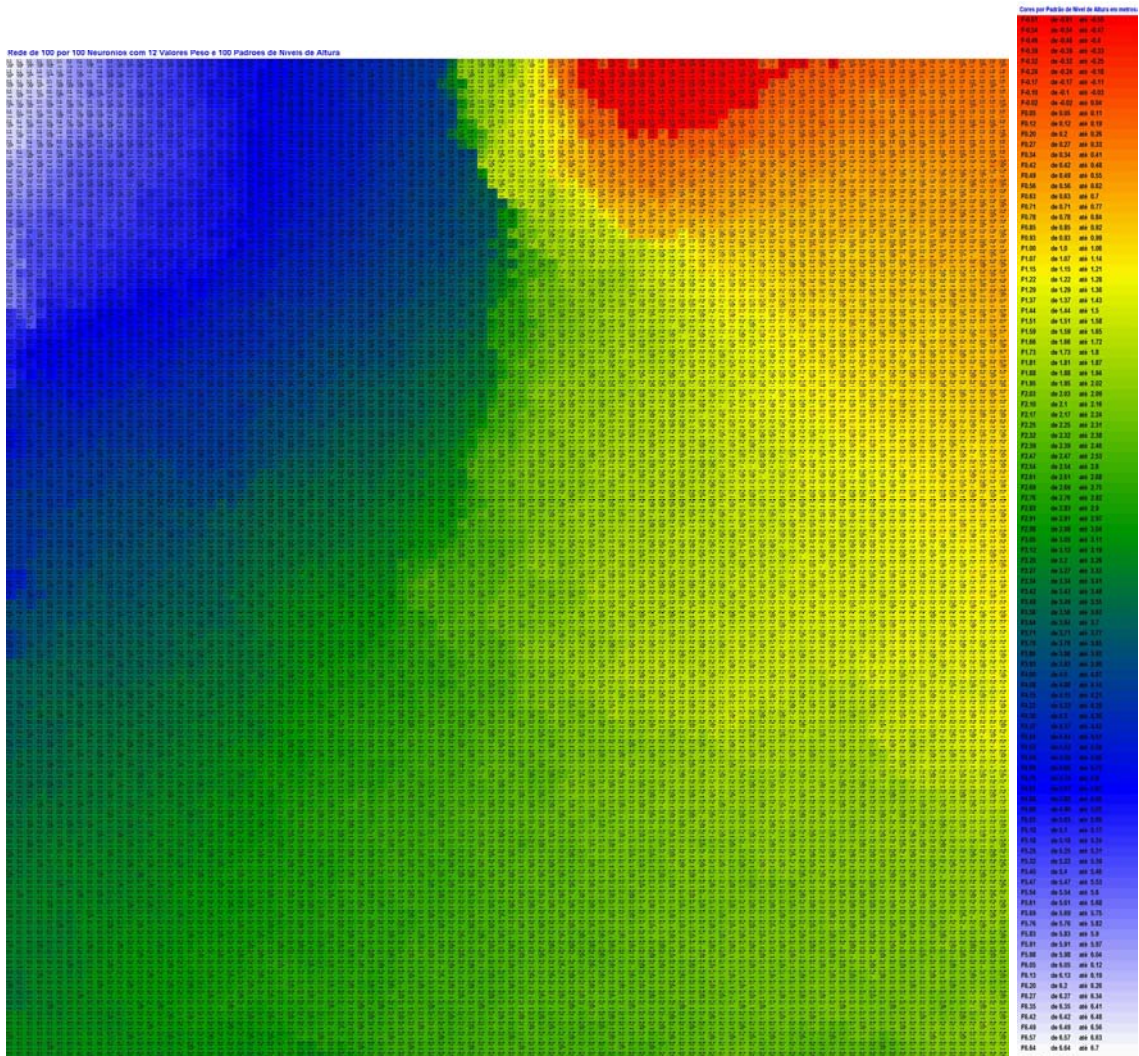


Figura 4.10 – Rede LVQ 100x100 neurônios, 12 pesos por neurônio e 100 padrões discretizados e sua legenda de cores por padrão discretizado

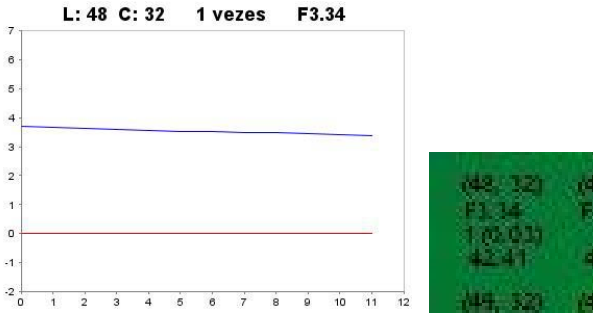


Figura 4.11 - Neurônio do resultado da previsão para dia 29, da rede LVQ 100x100, de 12 pesos por neurônio e 100 padrões discretizados

A Figura 4.10 apresenta a representação da mesma rede LVQ, mas utilizando as cores para mostrar o padrão discretizado que cada neurônio representa e reconhece. Neste mapa de cores, as regiões em azul e branco representam padrões de nível do rio Paraguai mais elevados, enquanto as regiões em vermelho são representações de padrões de nível do rio mais baixo.

Na Figura 4.11, é apresentada uma imagem ampliada do neurônio do resultado da previsão, com a cor que representa o padrão da previsão. As informações que são mostradas em cada neurônio deste mapa da rede LVQ, referem-se às características desse neurônio. Nesta previsão, o neurônio do resultado está localizado no quadrante superior esquerdo (linha 48, coluna 32), quase ao centro do mapa de cores, em região de transição entre azul e verde, com predominância da cor verde. A informação F3.34 indica o padrão do neurônio representando níveis de rio de 3,34 m até 3,41m. O valor 1 (0,03) informa o número de ocorrências de reconhecimento deste neurônio no treinamento da rede LVQ e o valor 42,41 informa a soma dos seus valores peso.

Pela Figura 4.9 pode-se entender como características dos neurônios ou comportamento do rio, pois as cores representam a soma dos valores-peso de cada neurônio. E pela Figura 10 pode-se entender como a situação do nível do rio. Por exemplo, a possibilidade ou não de se transportar cargas e/ou passageiros, pode ser analisada pela cor do neurônio resultado da previsão: vermelho para impossibilidade devido a baixo nível do rio, azul ou branco para nível ótimo para navegação e verde para possibilidade de navegação normal.

Tabela 4.36 - Resultados da previsão da rede lvq\_a3\_d8\_m100100\_F100, com os neurônios e suas localizações

Resultados das previsões pela Rede LVQ a3\_d8\_m100100\_F100

	<b>Neurônio do resultado</b>	<b>Padrão</b>	<b>Limite inferior (m)</b>	<b>Limite superior (m)</b>	<b>Nível médio do Padrão (m)</b>	<b>Medida real (m)</b>
dia 01:	LVQ ( L:66 C:73 )	Padrão sugerido= F4.15	de: 4,15	até: 4,21 m	4,18	4,11
dia 02:	LVQ ( L:64 C:71 )	Padrão sugerido= F4.08	de: 4,08	até: 4,14 m	4,11	4,09
dia 03:	LVQ ( L:64 C:71 )	Padrão sugerido= F4.08	de: 4,08	até: 4,14 m	4,11	4,07
dia 04:	LVQ ( L:61 C:75 )	Padrão sugerido= F4.08	de: 4,08	até: 4,14 m	4,11	4,04
dia 05:	LVQ ( L:63 C:69 )	Padrão sugerido= F4.00	de: 4,00	até: 4,07 m	4,04	4,02
dia 06:	LVQ ( L:60 C:74 )	Padrão sugerido= F4.00	de: 4,00	até: 4,07 m	4,04	4,00
dia 07:	LVQ ( L:60 C:71 )	Padrão sugerido= F3.93	de: 3,93	até: 3,99 m	3,96	3,98
dia 08:	LVQ ( L:56 C:75 )	Padrão sugerido= F3.93	de: 3,93	até: 3,99 m	3,96	3,94
dia 09:	LVQ ( L:60 C:69 )	Padrão sugerido= F3.86	de: 3,86	até: 3,92 m	3,89	3,93
dia 10:	LVQ ( L:54 C:76 )	Padrão sugerido= F3.86	de: 3,86	até: 3,92 m	3,89	3,90
dia 11:	LVQ ( L:53 C:78 )	Padrão sugerido= F3.86	de: 3,86	até: 3,92 m	3,89	3,87
dia 12:	LVQ ( L:54 C:75 )	Padrão sugerido= F3.86	de: 3,86	até: 3,92 m	3,89	3,85
dia 13:	LVQ ( L:52 C:76 )	Padrão sugerido= F3.86	de: 3,86	até: 3,92 m	3,89	3,83
dia 14:	LVQ ( L:56 C:69 )	Padrão sugerido= F3.78	de: 3,78	até: 3,85 m	3,82	3,81
dia 15:	LVQ ( L:56 C:67 )	Padrão sugerido= F3.78	de: 3,78	até: 3,85 m	3,82	3,78
dia 16:	LVQ ( L:52 C:72 )	Padrão sugerido= F3.78	de: 3,78	até: 3,85 m	3,82	3,75
dia 17:	LVQ ( L:48 C:78 )	Padrão sugerido= F3.71	de: 3,71	até: 3,77 m	3,74	3,72
dia 18:	LVQ ( L:52 C:68 )	Padrão sugerido= F3.71	de: 3,71	até: 3,77 m	3,74	3,70
dia 19:	LVQ ( L:47 C:77 )	Padrão sugerido= F3.64	de: 3,64	até: 3,70 m	3,67	3,67
dia 20:	LVQ ( L:50 C:72 )	Padrão sugerido= F3.64	de: 3,64	até: 3,70 m	3,67	3,63
dia 21:	LVQ ( L:49 C:72 )	Padrão sugerido= F3.56	de: 3,56	até: 3,63 m	3,60	3,60
dia 22:	LVQ ( L:48 C:72 )	Padrão sugerido= F3.56	de: 3,56	até: 3,63 m	3,60	3,58
dia 23:	LVQ ( L:47 C:70 )	Padrão sugerido= F3.56	de: 3,56	até: 3,63 m	3,60	3,55
dia 24:	LVQ ( L:45 C:73 )	Padrão sugerido= F3.49	de: 3,49	até: 3,55 m	3,52	3,52
dia 25:	LVQ ( L:40 C:75 )	Padrão sugerido= F3.49	de: 3,49	até: 3,55 m	3,52	3,49
dia 26:	LVQ ( L:45 C:70 )	Padrão sugerido= F3.49	de: 3,49	até: 3,55 m	3,52	3,44
dia 27:	LVQ ( L:36 C:77 )	Padrão sugerido= F3.42	de: 3,42	até: 3,48 m	3,45	3,40
dia 28:	LVQ ( L:38 C:75 )	Padrão sugerido= F3.34	de: 3,34	até: 3,41 m	3,38	3,37
dia 29:	LVQ ( L:34 C:77 )	Padrão sugerido= F3.34	de: 3,34	até: 3,41 m	3,38	3,34

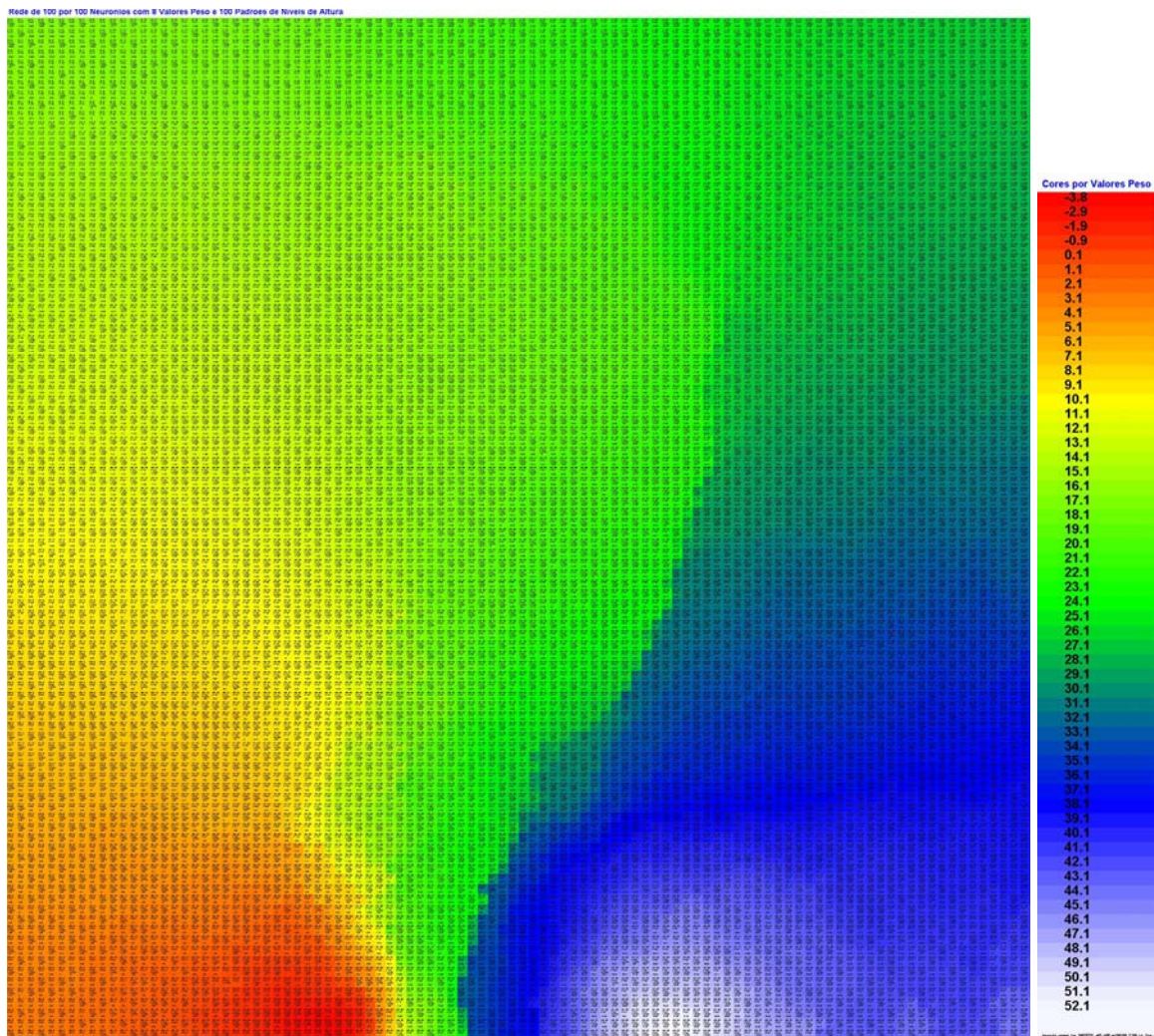


Figura 4.12 - Rede LVQ 100x100 neurônios, 8 pesos por neurônio e 100 padrões padrões discretizados e sua legenda de cores por soma dos valores peso de cada neurônio.

Na Figura 4.14, é apresentada a cor que representa o padrão do resultado da previsão, fornecido pelo neurônio indicado pelo processamento. Outras informações também são mostradas em cada neurônio deste mapa da rede LVQ. Na primeira linha é apresentada a posição do neurônio, na segunda linha o padrão discretizado que é representado pelo neurônio, na terceira linha está a quantidade de ocorrências que o neurônio reconheceu na fase de treinamento e na última linha está o valor da soma dos pesos do neurônio. Nesta previsão, o neurônio do resultado está localizado no quadrante superior direito (linha 34, coluna 77), em região de predominância da cor verde.

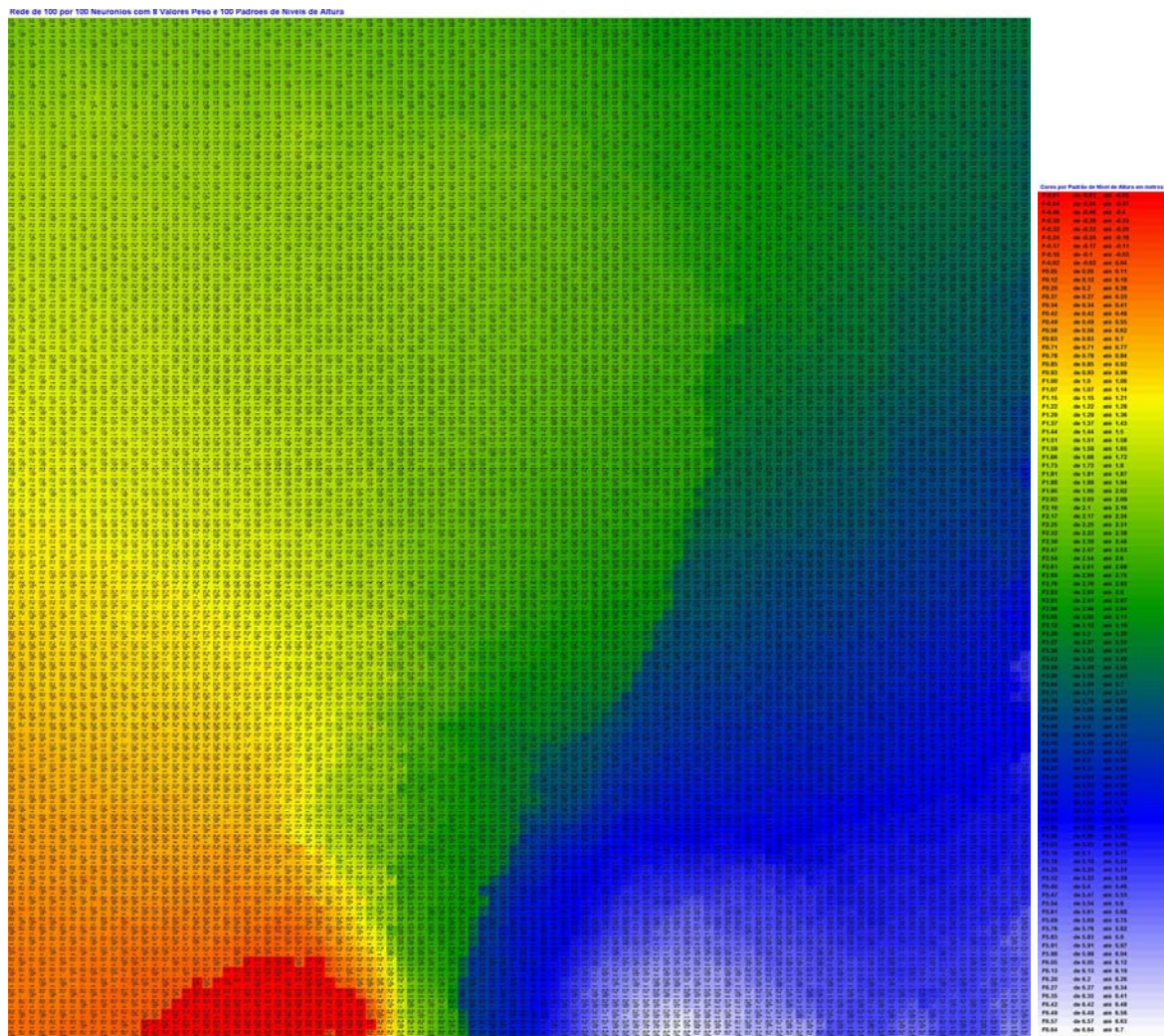


Figura 4.13 - Rede LVQ 100x100 neurônios, 8 pesos por neurônio e 100 padrões discretizados e sua legenda de cores por padrão discretizado.

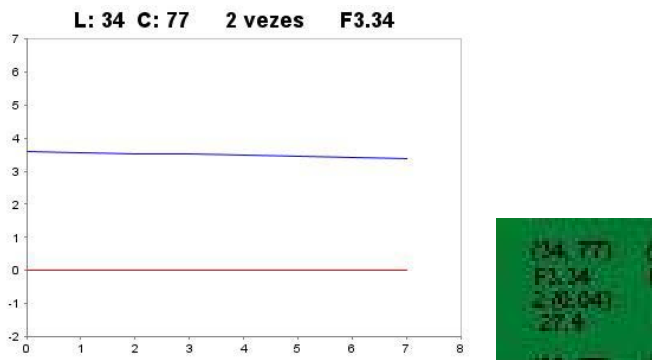


Figura 4.14 - Neurônio do resultado para previsão do dia 29, da rede LVQ 100x100 de 8 pesos por neurônio e 100 padrões discretizados

Tabela 4.37 - Resultados da previsão da rede lvq\_a3\_d8\_m3030\_F100, com os neurônios e suas localizações

Resultados das previsões pela Rede LVQ a3\_d8\_m3030\_F100

	Neurônio do resultado	Padrão	Limite inferior (m)	Limite superior (m)	Nível médio do Padrão (m)	Medida real (m)
dia 01:	LVQ ( L:15 C:28 )	Padrão sugerido= F4.22	de: 4,22	até: 4,29 m	4,26	4,11
dia 02:	LVQ ( L:14 C:29 )	Padrão sugerido= F4.08	de: 4,08	até: 4,14 m	4,11	4,09
dia 03:	LVQ ( L:14 C:29 )	Padrão sugerido= F4.08	de: 4,08	até: 4,14 m	4,11	4,07
dia 04:	LVQ ( L:13 C:28 )	Padrão sugerido= F4.08	de: 4,08	até: 4,14 m	4,11	4,04
dia 05:	LVQ ( L:13 C:29 )	Padrão sugerido= F4.00	de: 4,00	até: 4,07 m	4,04	4,02
dia 06:	LVQ ( L:12 C:29 )	Padrão sugerido= F4.00	de: 4,00	até: 4,07 m	4,04	4,00
dia 07:	LVQ ( L:12 C:29 )	Padrão sugerido= F4.00	de: 4,00	até: 4,07 m	4,04	3,98
dia 08:	LVQ ( L:12 C:26 )	Padrão sugerido= F3.93	de: 3,93	até: 3,99 m	3,96	3,94
dia 09:	LVQ ( L:11 C:28 )	Padrão sugerido= F3.93	de: 3,93	até: 3,99 m	3,96	3,93
dia 10:	LVQ ( L:11 C:28 )	Padrão sugerido= F3.93	de: 3,93	até: 3,99 m	3,96	3,90
dia 11:	LVQ ( L:11 C:29 )	Padrão sugerido= F3.86	de: 3,86	até: 3,92 m	3,89	3,87
dia 12:	LVQ ( L:11 C:26 )	Padrão sugerido= F3.86	de: 3,86	até: 3,92 m	3,89	3,85
dia 13:	LVQ ( L:10 C:28 )	Padrão sugerido= F3.86	de: 3,86	até: 3,92 m	3,89	3,83
dia 14:	LVQ ( L:10 C:27 )	Padrão sugerido= F3.78	de: 3,78	até: 3,85 m	3,82	3,81
dia 15:	LVQ ( L:9 C:28 )	Padrão sugerido= F3.78	de: 3,78	até: 3,85 m	3,82	3,78
dia 16:	LVQ ( L:9 C:27 )	Padrão sugerido= F3.71	de: 3,71	até: 3,77 m	3,74	3,75
dia 17:	LVQ ( L:8 C:28 )	Padrão sugerido= F3.71	de: 3,71	até: 3,77 m	3,74	3,72
dia 18:	LVQ ( L:8 C:27 )	Padrão sugerido= F3.71	de: 3,71	até: 3,77 m	3,74	3,70
dia 19:	LVQ ( L:8 C:26 )	Padrão sugerido= F3.64	de: 3,64	até: 3,70 m	3,67	3,67
dia 20:	LVQ ( L:7 C:27 )	Padrão sugerido= F3.64	de: 3,64	até: 3,70 m	3,67	3,63
dia 21:	LVQ ( L:7 C:29 )	Padrão sugerido= F3.64	de: 3,64	até: 3,70 m	3,67	3,60
dia 22:	LVQ ( L:7 C:28 )	Padrão sugerido= F3.56	de: 3,56	até: 3,63 m	3,60	3,58
dia 23:	LVQ ( L:6 C:27 )	Padrão sugerido= F3.56	de: 3,56	até: 3,63 m	3,60	3,55
dia 24:	LVQ ( L:5 C:29 )	Padrão sugerido= F3.49	de: 3,49	até: 3,55 m	3,52	3,52
dia 25:	LVQ ( L:5 C:28 )	Padrão sugerido= F3.49	de: 3,49	até: 3,55 m	3,52	3,49
dia 26:	LVQ ( L:5 C:27 )	Padrão sugerido= F3.49	de: 3,49	até: 3,55 m	3,52	3,44
dia 27:	LVQ ( L:3 C:29 )	Padrão sugerido= F3.42	de: 3,42	até: 3,48 m	3,45	3,40
dia 28:	LVQ ( L:3 C:28 )	Padrão sugerido= F3.34	de: 3,34	até: 3,41 m	3,38	3,37
dia 29:	LVQ ( L:3 C:28 )	Padrão sugerido= F3.34	de: 3,34	até: 3,41 m	3,38	3,34

### Rede de 30 por 30 Neurônios com 8 Valores Peso e 100 Padrões de Níveis de Altura

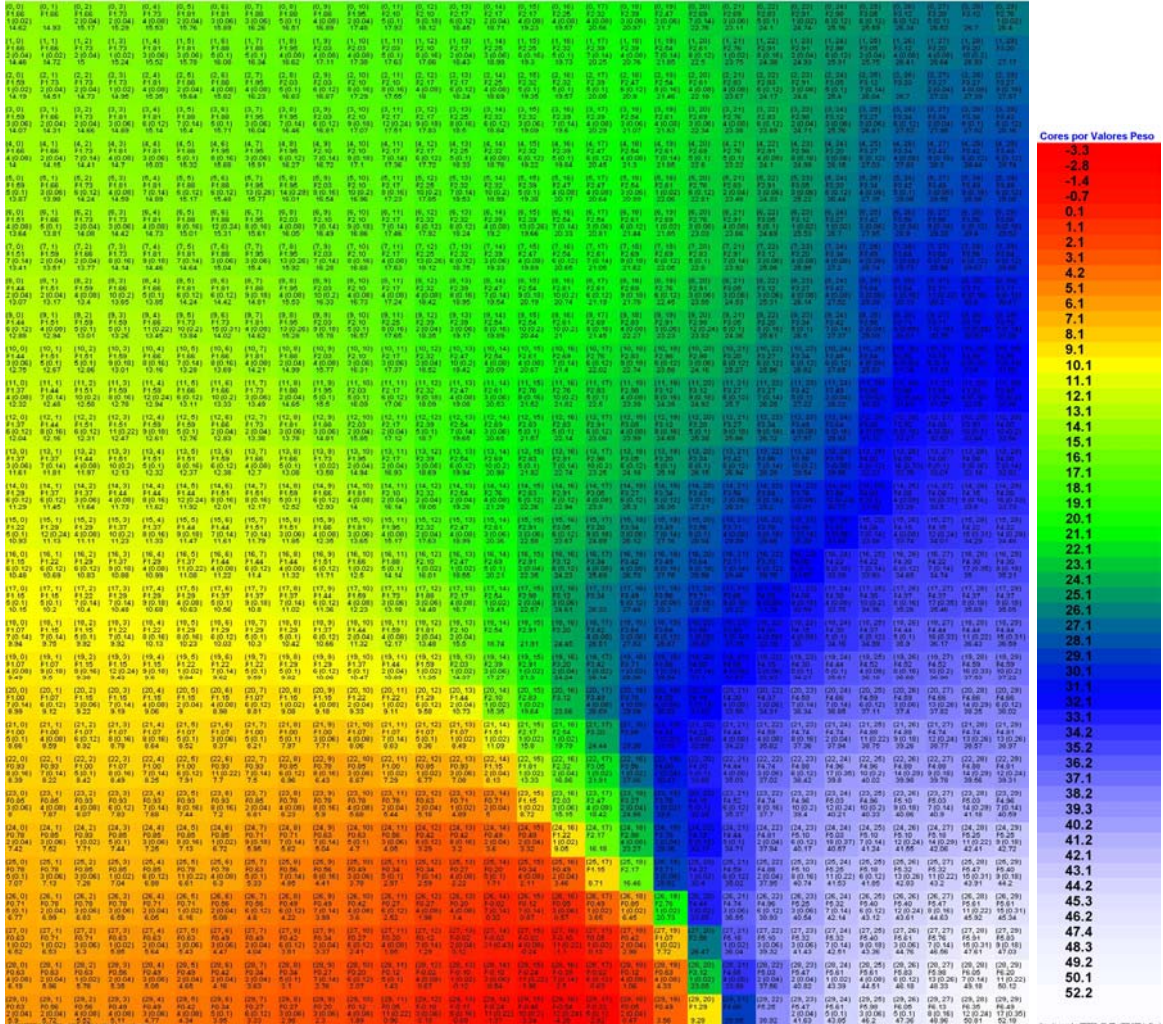


Figura 4.15 - Rede LVQ 30x30 neurônios, 8 pesos e 100 padrões discretizados de altura e sua legenda de cores por soma dos valores peso de cada neurônio.

Na Figura 4.17, é apresentada a cor do padrão do resultado da previsão. Nesta configuração da rede LVQ, o neurônio do resultado também está localizado no quadrante superior direito (linha 3, coluna 28), em região com predominância da cor verde.

Rede de 30 por 30 Neurônios com 8 Valores Peso e 100 Padrões de Níveis de Altura

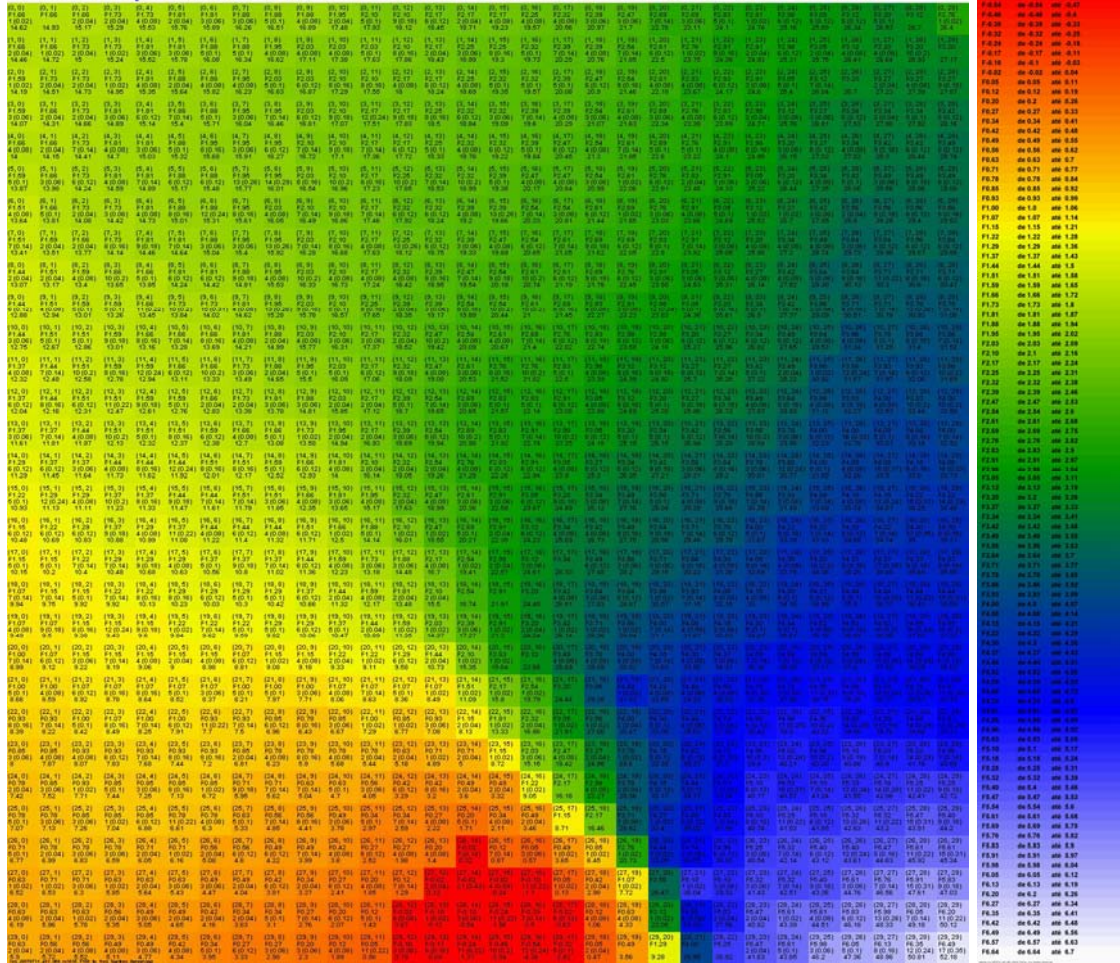


Figura 4.16 - Rede LVQ 30x30 neurônios, 8 pesos e 100 padrões discretizados e sua legenda de cores por padrão discretizado

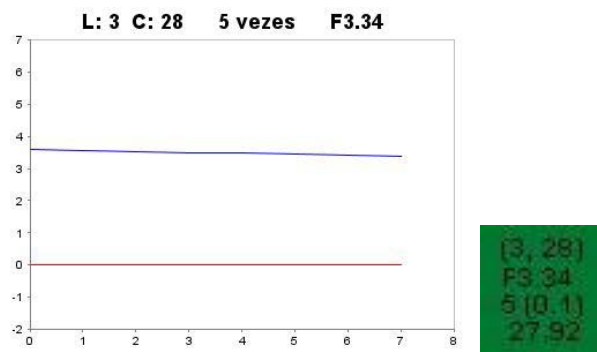


Figura 4.17 - Neurônio do resultado para previsão do dia 29, da rede LVQ 30x30, 8 pesos por neurônio e 100 padrões discretizados



#### 4.4.3 Previsões mais recentes

Foram realizadas previsões de um passo adiante para os dias dos meses de outubro e novembro de 2007, utilizando como referência 8 medidas anteriores ao dia previsto. A configuração da rede LVQ escolhida foi rede com 8 pesos por neurônio, tamanho de rede em 30x30 neurônios e 100 padrões discretizados, treinada por dados de gerados pela média de 3 medidas, identificada como a3\_d8\_m3030\_F100. Foram realizados também previsões dos mesmos dias e mesmas referências, através de redes LVQ com 8 pesos por neurônio, tamanho da rede em 100x100 neurônios e 100 padrões discretizados. Essa rede também foi treinada por dados gerados pela média de 3 medidas e foi identificada como a3\_d8\_m100100\_F100.

A Figura 4.18 apresenta o gráfico das previsões processadas pela rede a3\_d8\_m3030\_F100. A Tabela 4.38 apresenta os resultados das previsões processadas.

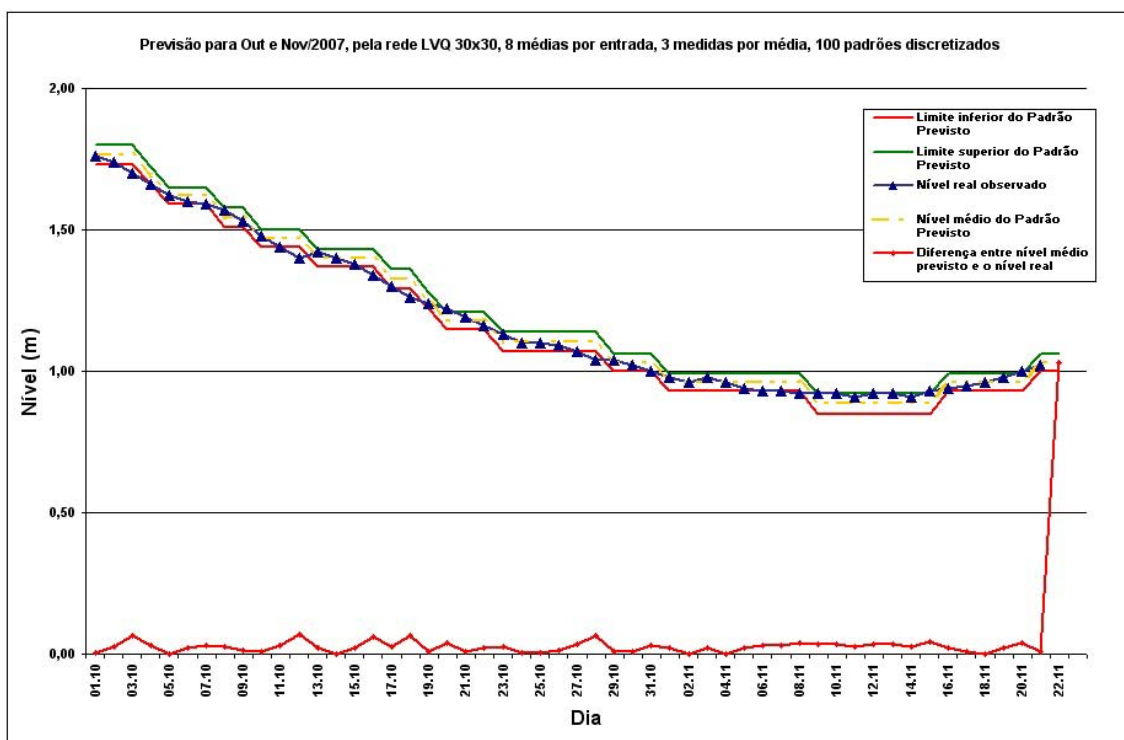


Figura 4.18 – Previsões para out. e nov./2007 – rede LVQ a3\_d8\_m3030\_F100

Tabela 4.38 - Previsões de out. e nov./2007 – rede LVQ a3\_d8\_m3030\_F100

a3\_d8\_m3030\_F100  
8 medidas anteriores

Dia	Nível real	Neurônio na rede LVQ 30x30	Padrão	Nível médio do Padrão	Limite inferior do Padrão	Limite superior do Padrão	Diferença
01.10	1,76	LVQ (L:0 C:4)	F1.73	1,77	1,73	1,80	0,01
02.10	1,74	LVQ (L:1 C:2)	F1.73	1,77	1,73	1,80	0,03
03.10	1,70	LVQ (L:1 C:2)	F1.73	1,77	1,73	1,80	0,07
04.10	1,66	LVQ (L:1 C:0)	F1.66	1,69	1,66	1,72	0,03
05.10	1,62	LVQ (L:2 C:0)	F1.59	1,62	1,59	1,65	0,00
06.10	1,60	LVQ (L:5 C:0)	F1.59	1,62	1,59	1,65	0,02
07.10	1,59	LVQ (L:7 C:1)	F1.59	1,62	1,59	1,65	0,03
08.10	1,57	LVQ (L:8 C:1)	F1.51	1,55	1,51	1,58	0,03
09.10	1,53	LVQ (L:8 C:1)	F1.51	1,55	1,51	1,58	0,01
10.10	1,48	LVQ (L:10 C:0)	F1.44	1,47	1,44	1,50	0,01
11.10	1,44	LVQ (L:11 C:1)	F1.44	1,47	1,44	1,50	0,03
12.10	1,40	LVQ (L:12 C:1)	F1.44	1,47	1,44	1,50	0,07
13.10	1,42	LVQ (L:12 C:0)	F1.37	1,40	1,37	1,43	0,02
14.10	1,40	LVQ (L:13 C:1)	F1.37	1,40	1,37	1,43	0,00
15.10	1,38	LVQ (L:13 C:0)	F1.37	1,40	1,37	1,43	0,02
16.10	1,34	LVQ (L:13 C:0)	F1.37	1,40	1,37	1,43	0,06
17.10	1,30	LVQ (L:15 C:1)	F1.29	1,33	1,29	1,36	0,03
18.10	1,26	LVQ (L:15 C:1)	F1.29	1,33	1,29	1,36	0,07
19.10	1,24	LVQ (L:16 C:1)	F1.22	1,25	1,22	1,28	0,01
20.10	1,22	LVQ (L:16 C:0)	F1.15	1,18	1,15	1,21	0,04
21.10	1,19	LVQ (L:16 C:0)	F1.15	1,18	1,15	1,21	0,01
22.10	1,16	LVQ (L:17 C:0)	F1.15	1,18	1,15	1,21	0,02
23.10	1,13	LVQ (L:18 C:0)	F1.07	1,11	1,07	1,14	0,02
24.10	1,10	LVQ (L:19 C:0)	F1.07	1,11	1,07	1,14	0,00
25.10	1,10	LVQ (L:19 C:0)	F1.07	1,11	1,07	1,14	0,00
26.10	1,09	LVQ (L:20 C:1)	F1.07	1,11	1,07	1,14	0,01
27.10	1,07	LVQ (L:20 C:1)	F1.07	1,11	1,07	1,14	0,03
28.10	1,04	LVQ (L:21 C:2)	F1.07	1,11	1,07	1,14	0,06
29.10	1,04	LVQ (L:21 C:1)	F1.00	1,03	1,00	1,06	0,01
30.10	1,02	LVQ (L:21 C:1)	F1.00	1,03	1,00	1,06	0,01
31.10	1,00	LVQ (L:21 C:1)	F1.00	1,03	1,00	1,06	0,03
01.11	0,98	LVQ (L:22 C:1)	F0.93	0,96	0,93	0,99	0,02
02.11	0,96	LVQ (L:22 C:1)	F0.93	0,96	0,93	0,99	0,00
03.11	0,98	LVQ (L:22 C:1)	F0.93	0,96	0,93	0,99	0,02
04.11	0,96	LVQ (L:23 C:2)	F0.93	0,96	0,93	0,99	0,00
05.11	0,94	LVQ (L:23 C:2)	F0.93	0,96	0,93	0,99	0,02
06.11	0,93	LVQ (L:24 C:2)	F0.93	0,96	0,93	0,99	0,03
07.11	0,93	LVQ (L:24 C:2)	F0.93	0,96	0,93	0,99	0,03
08.11	0,92	LVQ (L:24 C:2)	F0.93	0,96	0,93	0,99	0,04
09.11	0,92	LVQ (L:24 C:3)	F0.85	0,89	0,85	0,92	0,04
10.11	0,92	LVQ (L:24 C:3)	F0.85	0,89	0,85	0,92	0,04
11.11	0,91	LVQ (L:24 C:3)	F0.85	0,89	0,85	0,92	0,03
12.11	0,92	LVQ (L:24 C:4)	F0.85	0,89	0,85	0,92	0,04
13.11	0,92	LVQ (L:24 C:4)	F0.85	0,89	0,85	0,92	0,04
14.11	0,91	LVQ (L:24 C:4)	F0.85	0,89	0,85	0,92	0,03
15.11	0,93	LVQ (L:24 C:4)	F0.85	0,89	0,85	0,92	0,05
16.11	0,94	LVQ (L:23 C:5)	F0.93	0,96	0,93	0,99	0,02
17.11	0,95	LVQ (L:23 C:5)	F0.93	0,96	0,93	0,99	0,01
18.11	0,96	LVQ (L:23 C:5)	F0.93	0,96	0,93	0,99	0,00
19.11	0,98	LVQ (L:23 C:5)	F0.93	0,96	0,93	0,99	0,02
20.11	1,00	LVQ (L:22 C:7)	F0.93	0,96	0,93	0,99	0,04
21.11	1,02	LVQ (L:21 C:9)	F1.00	1,03	1,00	1,06	0,01
22.11		LVQ (L:21 C:9)	F1.00	1,03	1,00	1,06	1,03
23.11							
Todas unidades estão em metro						Total	2,32

Na Tabela 4.38, a coluna Dia mostra a data a ser prevista. A coluna Nível real apresenta o valor real do nível do rio Paraguai medido no dia da previsão. A coluna Neurônio na Rede LVQ mostra a posição do neurônio que forneceu a previsão, em linha e coluna do mapa da rede LVQ. A coluna Padrão informa o padrão discretizado obtido na previsão. A coluna Nível médio do Padrão informa o valor médio da faixa de níveis do rio que o padrão representa. Esse valor é obtido pelo cálculo da média entre o valor da coluna Limite inferior do Padrão e o valor da coluna Limite superior do padrão. Estas duas últimas colunas são os limites do padrão de níveis do rio discretizado. A última coluna apresenta a diferença do valor médio do padrão discretizado previsto e o valor real do nível do rio medido no dia previsto. A Figura 4.19 apresenta o gráfico com os resultados da previsão de outubro e novembro de 2007, gerados pela rede LVQ a3\_d8\_m100100\_F100. A Tabela 4.39 apresenta os resultados destas previsões processadas.

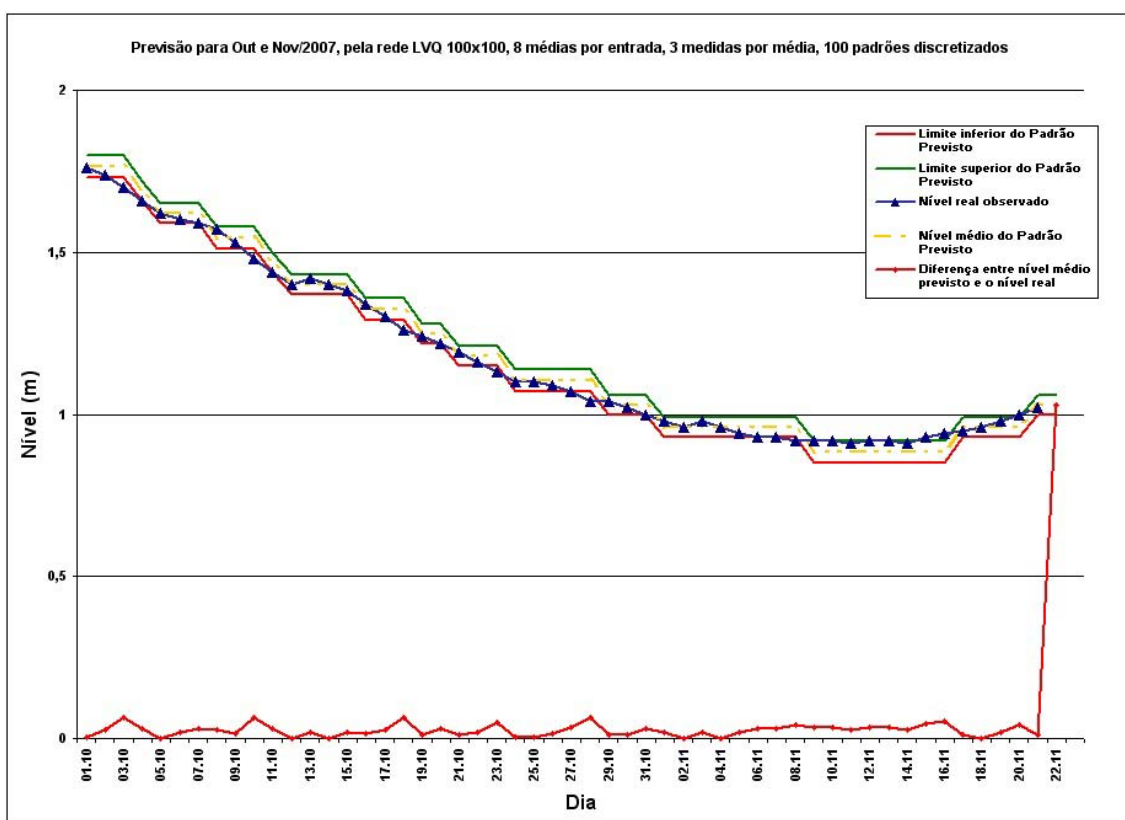


Figura 4.19 - Previsões para out. e nov./2007 – rede LVQ a3\_d8\_m100100\_F100

Tabela 4.39 - Previsões de out. e nov./2007 – rede LVQ a3\_d8\_m100100\_F100

a3\_d8\_m100100\_F100  
8 medidas anteriores

Dia	Nível real	Neurônio na rede LVQ	Padrão	Nível médio do Padrão	Limite inferior do Padrão	Limite superior do Padrão	Diferença
01.10	1,76	LVQ (L:13 C:3)	F1.73	1,765	1,73	1,8	0,01
02.10	1,74	LVQ (L:14 C:5)	F1.73	1,765	1,73	1,8	0,03
03.10	1,70	LVQ (L:15 C:7)	F1.73	1,765	1,73	1,8	0,07
04.10	1,66	LVQ (L:17 C:6)	F1.66	1,69	1,66	1,72	0,03
05.10	1,62	LVQ (L:22 C:0)	F1.59	1,62	1,59	1,65	0,00
06.10	1,60	LVQ (L:23 C:4)	F1.59	1,62	1,59	1,65	0,02
07.10	1,59	LVQ (L:25 C:6)	F1.59	1,62	1,59	1,65	0,03
08.10	1,57	LVQ (L:29 C:4)	F1.51	1,545	1,51	1,58	0,03
09.10	1,53	LVQ (L:28 C:7)	F1.51	1,545	1,51	1,58	0,01
10.10	1,48	LVQ (L:30 C:5)	F1.51	1,545	1,51	1,58	0,06
11.10	1,44	LVQ (L:33 C:6)	F1.44	1,47	1,44	1,5	0,03
12.10	1,40	LVQ (L:35 C:3)	F1.37	1,4	1,37	1,43	0,00
13.10	1,42	LVQ (L:38 C:3)	F1.37	1,4	1,37	1,43	0,02
14.10	1,40	LVQ (L:38 C:5)	F1.37	1,4	1,37	1,43	0,00
15.10	1,38	LVQ (L:39 C:5)	F1.37	1,4	1,37	1,43	0,02
16.10	1,34	LVQ (L:42 C:6)	F1.29	1,325	1,29	1,36	0,01
17.10	1,30	LVQ (L:44 C:4)	F1.29	1,325	1,29	1,36	0,03
18.10	1,26	LVQ (L:45 C:3)	F1.29	1,325	1,29	1,36	0,07
19.10	1,24	LVQ (L:48 C:6)	F1.22	1,25	1,22	1,28	0,01
20.10	1,22	LVQ (L:48 C:1)	F1.22	1,25	1,22	1,28	0,03
21.10	1,19	LVQ (L:49 C:2)	F1.15	1,18	1,15	1,21	0,01
22.10	1,16	LVQ (L:49 C:2)	F1.15	1,18	1,15	1,21	0,02
23.10	1,13	LVQ (L:53 C:3)	F1.15	1,18	1,15	1,21	0,05
24.10	1,10	LVQ (L:53 C:5)	F1.07	1,105	1,07	1,14	0,00
25.10	1,10	LVQ (L:56 C:5)	F1.07	1,105	1,07	1,14	0,00
26.10	1,09	LVQ (L:57 C:6)	F1.07	1,105	1,07	1,14	0,01
27.10	1,07	LVQ (L:58 C:7)	F1.07	1,105	1,07	1,14	0,03
28.10	1,04	LVQ (L:58 C:7)	F1.07	1,105	1,07	1,14	0,06
29.10	1,04	LVQ (L:60 C:7)	F1.00	1,03	1,00	1,06	0,01
30.10	1,02	LVQ (L:60 C:6)	F1.00	1,03	1,00	1,06	0,01
31.10	1,00	LVQ (L:60 C:6)	F1.00	1,03	1,00	1,06	0,03
01.11	0,98	LVQ (L:62 C:7)	F0.93	0,96	0,93	0,99	0,02
02.11	0,96	LVQ (L:64 C:5)	F0.93	0,96	0,93	0,99	0,00
03.11	0,98	LVQ (L:64 C:6)	F0.93	0,96	0,93	0,99	0,02
04.11	0,96	LVQ (L:64 C:9)	F0.93	0,96	0,93	0,99	0,00
05.11	0,94	LVQ (L:65 C:9)	F0.93	0,96	0,93	0,99	0,02
06.11	0,93	LVQ (L:65 C:7)	F0.93	0,96	0,93	0,99	0,03
07.11	0,93	LVQ (L:65 C:7)	F0.93	0,96	0,93	0,99	0,03
08.11	0,92	LVQ (L:68 C:13)	F0.93	0,96	0,93	0,99	0,04
09.11	0,92	LVQ (L:66 C:8)	F0.85	0,885	0,85	0,92	0,04
10.11	0,92	LVQ (L:66 C:8)	F0.85	0,885	0,85	0,92	0,04
11.11	0,91	LVQ (L:66 C:8)	F0.85	0,885	0,85	0,92	0,03
12.11	0,92	LVQ (L:68 C:8)	F0.85	0,885	0,85	0,92	0,04
13.11	0,92	LVQ (L:69 C:14)	F0.85	0,885	0,85	0,92	0,04
14.11	0,91	LVQ (L:69 C:14)	F0.85	0,885	0,85	0,92	0,03
15.11	0,93	LVQ (L:68 C:8)	F0.85	0,885	0,85	0,92	0,05
16.11	0,94	LVQ (L:69 C:14)	F0.85	0,885	0,85	0,92	0,05
17.11	0,95	LVQ (L:70 C:17)	F0.93	0,96	0,93	0,99	0,01
18.11	0,96	LVQ (L:70 C:17)	F0.93	0,96	0,93	0,99	0,00
19.11	0,98	LVQ (L:70 C:19)	F0.93	0,96	0,93	0,99	0,02
20.11	1,00	LVQ (L:70 C:19)	F0.93	0,96	0,93	0,99	0,04
21.11	1,02	LVQ (L:71 C:22)	F1.00	1,03	1,00	1,06	0,01
22.11		LVQ (L:71 C:22)	F1.00	1,03	1,00	1,06	1,03
23.11							
Todas unidades estão em metro						Diferença Total	2,31

Nos gráficos das Figuras 4.18 e 4.19 são mostrados os padrões resultados das previsões através das linhas inferior e superior. Estas linhas representam os limites inferior e superior da faixa de níveis do padrão discretizado. É apresentado também o valor médio do padrão previsto através de linha tracejada, o valor do nível real medido no dia e linha que mostra a diferença entre o nível médio da previsão e o nível real do rio Paraguai.

Em ambos gráficos, a linha da diferença encontra com a linha do valor médio do padrão porque no momento da finalização deste trabalho, não havia sido informado o valor real medido para o dia 22 de novembro.

#### **4.5 Análise dos Resultados**

O treinamento da rede BPNN mencionado por Li et al (1998) utilizaram entre 30 a 60 minutos para processar 274 valores médios mensais, dependendo da estrutura da rede neural utilizada. Neste trabalho, foram processadas 39295 medidas de nível do rio Paraguai a cada experimento completo desta metodologia, conforme o diagrama da Figura 4.1, em que o tempo de processamento variou de 1 minuto em redes com 10x10 neurônios até 110 minutos em redes com 100x100 neurônios, sendo a média de 15 minutos para rede LVQ com 50x50 neurônios. Observa-se que a comparação através do tempo de processamento é relativa, pois os ambientes computacionais são diferentes.

Os experimentos no treinamento de redes LVQ de configurações diferentes foram separados em 2 tipos distintos. O primeiro tipo foram os experimentos que geraram previsões de 7 períodos selecionados para comparar com os resultados das previsões realizadas por rede BPNN mencionados no artigo de Li et al (1998). O segundo tipo de experimentos foram os realizados pelas redes LVQ geradas pelo conjunto de medidas de 1900 até julho de 2007, para

realizar previsões de alguns dias de agosto de 2007. Os resultados e suas análises são apresentados a seguir.

#### **4.5.1 Sobre resultados das previsões utilizando 7 períodos**

Os resultados para previsão de um passo adiante apresentados na Tabela 4.8 e na Tabela 4.9, mostram que, das 301 redes LVQ de 7 períodos específicos com 0% erros, 65% dos resultados foram obtidos pelas redes geradas na configuração de 3 medidas para obter médias, isto é 196 redes. Desse total, 67 redes são da configuração de 4 entradas (22,26%), 63 de 8 entradas (20,93%) e 65 da configuração de 12 entradas (21,93%), ou seja, não existem diferenças significativas na escolha da rede LVQ sobre a configuração do número de entradas.

Sobre o número de padrões utilizados pelas redes na previsão de um passo, o melhor resultado foi obtido pela configuração de 10 padrões com 4 entradas (4,31%). Considerando que esses 10 padrões têm uma diferença que chega a 0,8m entre o limite inferior e o limite superior de cada padrão, essa configuração dá resultados sem muita precisão. O segundo melhor resultado é obtido por 50 padrões com 12 entradas que tem 5,65% do total de redes com erros 0% (17 de um total de 301). Essa configuração é mais interessante que a anterior, pois 50 padrões têm precisão maior, isto é, cada padrão representa uma faixa de 0,13m, mais preciso que 0,8m para cada faixa dos 10 padrões. Em relação ao tamanho, as redes LVQ com 50x50 e 100x100 neurônios ficaram com 22,92% e 23,26% do total de resultados com erros 0%, isto é 69 e 70 redes LVQ respectivamente, de um total de 301 redes LVQ com erros 0%. Novamente a diferença é mínima, e a pequena vantagem das redes 100x100 é obtida na configuração com 10 padrões. Ao desconsiderar todas as previsões com 10 padrões, pelas respostas serem muito amplas, a situação entre as redes de 50x50 e 100x100 se equivale plenamente (16,61% para ambas as configurações).

Para 1575 previsões de um passo adiante utilizando redes LVQ gerados com 7 períodos diferentes e em diversas configurações, os melhores resultados com erros 0% foram com redes geradas pelos parâmetros: 3 medidas por média, 12 entradas, 50 padrões e tamanho 50x50.

Os resultados de previsão de cada um dos 7 períodos, as Tabelas 4.10, 4.11, 4.12 e 4.13, indicam que a rede LVQ gerada pelas medidas de janeiro de 1900 até dezembro de 1947 forneceu o maior número de previsões de um passo com 0% de erro (123 previsões). Isto significa 7,81% do total de processamentos (1575) e 40,86% das previsões com erros 0% (301), sendo que a configuração com melhor número de previsões foi de 3 medidas para média (73 redes), 4 entradas (67 redes), 10 padrões discretizados (29 redes) e tamanho 20x20 neurônios (27 redes). Porém nesta configuração os resultados são pouco precisos pelo padrão utilizado. A segunda configuração, mais adequada, é a que utiliza 20 padrões (26 redes) e 30 padrões (25 redes), com preferência ao de 30 padrões por permitir maior precisão na previsão. Em 30 padrões, cada padrão tem uma diferença entre o nível mais baixo e o mais alto de 0,22m, enquanto em 20 padrões, essa diferença aumenta para 0,28m. Uma opção interessante pode ser a utilização de 50 padrões porque a variação entre os 5 padrões em relação à soma total de 123 redes com erros 0% desse período é pequena, e a precisão obtida é melhor que a perda em percentual de redes com erros 0%. Por exemplo: para 10 padrões: 23,58%; para 20 padrões: 21,14%; 30 padrões: 20,35%; para 50 padrões: 18,70% e para 100 padrões: 16,26%.

Os piores resultados foram das redes dos períodos de janeiro de 1900 até setembro de 1930 e de abril de 1983 até fevereiro de 1995, em ambos obtiveram resultados erro 0% em apenas 8 redes, sendo 4 para cada período.

Pela discrepância dos resultados, identificou-se que o número de medidas foi importante para um treinamento eficaz, como ficou demonstrado nas tabelas acima, isto é, a quantidade de informações históricas.

Porém, mesmo para um conjunto grande de medidas, houve resultados de previsão insatisfatórios, pois os “alvos” ou dias a serem previstos escolhidos apresentaram resultados altos já na apresentação do artigo. Esses resultados podem advir da inexistência de situações históricas semelhantes e por isso não houve o aprendizado necessário pela rede, impossibilitando uma previsão com maior qualidade para esses “alvos”. Nesse caso, a rede identifica a situação mais semelhante, que geralmente resulta em uma previsão incorreta.

Uma observação deve ser feita para explicar que os treinamentos das redes LVQ foram realizados utilizando as medidas dos períodos selecionados, desconsiderando qualquer influência alheia às medidas de nível do rio. Isto significa que foram utilizadas somente as medidas de nível do rio Paraguai, pré-processadas, sem outras informações ambientais como as geológicas e climatológicas entre outros.

Os resultados da Tabela 4.14 demonstram a possibilidade e a capacidade desta metodologia de realizar previsões, com resultados satisfatórios. Em 14 das 28 previsões da tabela, obteve resultados 0% de erros, e mais outras 3 previsões ficaram abaixo de 17% de erros, totalizando 17 previsões abaixo de 17%. Isso significa que 60,71% dos resultados foram satisfatórios.

#### **4.5.2 Sobre resultados das previsões dos dias de agosto.**

Dos processamentos de previsão de 1 passo adiante utilizando redes LVQ geradas pelo treinamento com o conjunto total de medidas (desde 1900 até 31 de julho de 2007) obtivemos valores apresentados nas Tabelas 4.15, 4.16, 4.17, 4.18, 4.19 e 4.20. O total de redes LVQ com resultado 0% erros somou



46 redes, de um total de 225 redes para os processamentos de previsão, correspondendo a 20,44% de redes com 0% de erro. Pela Tabela 4.16 percebe-se um desempenho melhor das redes LVQ com configuração de 30 medidas por média (54,35%) em relação ao de 3 medidas por média (45,65% sobre as 46 redes com 0% erros), principalmente porque nessa configuração são computados os resultados das previsões feitas por redes que utilizaram 10 padrões.

Para esses processamentos para previsão de um passo adiante, observa-se que os melhores resultados de 0% erros foram obtidos com as redes LVQ geradas por 30 medidas para média (Tabela 4.17), 4 elementos por entrada (Tabela 4.18), 20 padrões discretizados (Tabela 4.19) e 100x100 neurônios (Tabela 4.20). Notamos que o número de 20 padrões discretizados forneceu maior quantidade de redes LVQ com 0% erros, com 54,35% em relação ao total de 46 redes de 0% e 55,56% em relação ao total de 45 redes LVQ com 20 padrões. Por esses números a escolha tende a ser 20 padrões, porém com 100 padrões discretizados todas as 45 redes LVQ desse padrão forneceram previsões com erros entre 0% e 17%, portanto ainda aceitável pela melhor precisão que as previsões de 20 padrões podem fornecer.

Os processamentos de previsão para 21 passos adiante utilizando essas mesmas redes LVQ, pelo volume de resultados, permitem uma visão melhor do comportamento dessa metodologia. Esses resultados são apresentados resumidamente na Tabela 4.21 e na Tabela 4.22, que contemplam 490 previsões com erro 0%, do total de 4725 processamentos. Como na previsão de um passo adiante, os resultados confirmaram que as configurações de redes LVQ com 30 medidas para média fornecem respostas mais numerosas, isto é 52,86% contra 47,14% dos resultados para 3 medidas para média.

Dentre essa configuração de 30 medidas para média, as redes com 4 elementos por entrada forneceram melhores resultados porque além de ter o

maior número de redes com 0% de erro, poucas foram gerados por 10 padrões (3,88%). Isto indica previsões mais precisas, com maior qualidade. Ao desconsiderar todos os resultados de 10 padrões discretizados, os resultados percentuais das redes com 30 medidas para média e 4 elementos por entrada passa de 16,53% para 12,65%, enquanto para 8 entradas os resultados alteram-se de 18,78% a 8,16% e com 12 elementos por entrada, passam de 17,55% para 10,20%. Essas diferenças indicam a influência dos resultados de 10 padrões.

E é também das redes de 30 medidas, 4 elementos por entrada e 30 padrões discretizados, que tem os resultados de erro 0% mais numerosos, seguido de 30 medidas, 8 elementos por entrada e 20 padrões discretizados, sempre sem considerar os resultados de redes com 10 padrões.

Sobre o tamanho das redes LVQ desses processamentos, a Tabela 4.22 indica que 100x100 neurônios tem os resultados com 0% de erros mais quantitativos (11,84%), enquanto 30x30 tem 10,00%. Novamente ignorando os resultados para 10 padrões, temos então as redes de 30x30 com 7,96% das previsões com 0% de erros. Essa situação é mais condizente ao apresentado em toda a Tabela 4.22, pois a configuração 30x30 apresenta maiores quantidades de previsões com 0% de erros.

A análise dos resultados da Tabela 4.23 até a Tabela 4.26, permite confirmar as considerações realizadas para a Tabela 4.21 e Tabela 4.22, pois são apresentados os resultados percentuais de todas as 4725 previsões realizadas, desde 1 passo adiante até 21 passos adiante. Por exemplo, confirma-se que para erros 0%, as redes LVQ geradas com 30 medidas para média dão os melhores resultados, e também redes com 4 elementos por entrada. E o mesmo ocorre com redes de 100x100 neurônios. Sobre o número de padrões, a configuração de 20 padrões forneceu resultados melhores (34,90% das 490 previsões) que os de 30 padrões com 30 medidas para média e 4 elementos

por entrada, sugerido pelas Tabelas 4.21 e 4.22. Porém esse resultado refere-se a todas as previsões (4725), sem especificar uma determinada configuração de número de medidas para média, número de elementos por entrada e número de padrões utilizados. Por isso, é mais plausível a utilização da configuração sugerida anteriormente, que é 30 medidas para média, 4 elementos por entrada e 30 padrões.

De todos os processamentos de previsão realizados foram selecionadas algumas redes utilizadas para previsão de 23 dias, 1 passo adiante por vez, com as referências utilizadas obtidas das medidas reais, isto é, cada previsão foi processada para o dia seguinte. Nas Figuras 4.2, 4.3 e 4.4 observamos resultados em que a linha que representa os valores reais aparece entre as linhas dos limites inferior e superior dos padrões previstos pelos processamentos, indicando previsões corretas. As Tabelas 4.28, 4.29 e 4.30 demonstram os valores utilizados para gerar os gráficos das Figuras 4.2, 4.3 e 4.4, em que os processamentos dessas previsões utilizaram redes LVQ geradas por 3 medidas para cálculo de média. É uma indicação de que os treinamentos das redes foram adequados.

Como os resultados mostraram-se satisfatórios, outras previsões foram realizadas com configurações de redes LVQ mais detalhistas, isto é, maior número de padrões, maior número de neurônios e também maior número de elementos por entrada. Esses processamentos são representados pelas Figuras 4.5 até 4.8 e pelos seus dados, das Tabelas 4.31 até 4.34.

Notamos que quanto maiores são os números de padrões utilizados mais preciso serão as previsões. O mesmo ocorre com os números de neurônios ou tamanho das redes LVQ e com os números de elementos por entrada.

Sobre a quantidade de medidas utilizadas para o cálculo de média que compõe cada elemento da entrada, a precisão e o número de acertos (erros 0%) são afetados pelo aumento desse número. A Figura 4.7 mostra os resultados da

previsão gerados pela rede LVQ de tamanho 100x100 neurônios, 8 pesos por neurônio e 100 padrões discretizados, treinada por dados reduzidos de 3 medidas por média. Esta rede obteve um dos menores valores de desvio padrão (0,0388), valor da variância (0,0015086) e valor de Erro padrão da média (0,0072). Valores menores e melhores ocorreram com a rede LVQ de 30x30 neurônios, 4 pesos por neurônio e 100 padrões discretizados, treinada por dados reduzidos de 3 medidas para cálculo da média. Porém não pode ser considerada a indicada por ser uma rede menor (30x30 neurônios). Uma rede LVQ de tamanho reduzido, ao atualizado por novos treinamentos, poderá gerar alguma generalização por não ter neurônios suficientes que reconheçam melhor as características destes treinamentos. Por isso, a rede LVQ lvq\_a3\_d8\_m100100\_F100 pode classificar melhor, e realizar previsões com resultados mais próximos aos níveis reais.

As Figuras 4.9 e 4.10 são imagens da rede LVQ 100x100 neurônios, de 3 medidas para média, 12 entradas e 100 padrões, com cores por padrão discretizado e por soma de valores peso de cada neurônio. As Figuras 4.12 e 4.13 são imagens da rede LVQ 100x100 neurônios de 3 medidas para média, 8 entradas e 100 padrões, com cores por padrão discretizado e por soma de valores peso de cada neurônio. E as Figuras 4.15 e 4.16 são imagens da rede LVQ 30x30 neurônios, de 3 medidas para média, 8 entradas e 100 padrões, com cores por padrão discretizado e por soma de valores peso de cada neurônio, respectivamente.

Todas essas figuras apresentam regiões com padrões de níveis semelhantes, que acompanham os valores peso dos neurônios. Visualmente fica fácil a interpretação do resultado da previsão. Por exemplo, a previsão para 29 de agosto de 2007 processada pela rede LVQ 30x30 (da Figura 4.15 e da Figura 4.16), foi padrão F3.34, neurônio L3:C28, isto é, linha 3 e coluna 28. No mapa da rede (Figuras 4.15 e 4.16) localizamos esse neurônio no quadrante superior direito, em uma região com predominância da cor verde, com níveis entre 3,34

até 3,41m. É uma região mais próxima do azul, de níveis de rio mais alto e ainda distante da região em vermelho, com níveis baixos. As Figuras 4.11, 4.14 e 4.17, são as representações dos valores peso dos neurônios que resultaram nas previsões. Através delas se podem visualizar as características dos padrões previstos, e possivelmente, de como os valores anteriores influenciaram nos resultados das previsões.

#### **4.6 Conclusões dos resultados obtidos**

As redes LVQ geradas pela média de 30 dias (mensal), apresentaram resultados insatisfatórios. Isso porque a redução do número de medidas através de médias provocou perda de valores para uma previsão eficaz, indicando o compromisso entre a precisão da previsão e a redução dos dados utilizados.

A utilização de um número pequeno de padrões (10 faixas) mostrou muitas previsões dentro dos limites desses padrões, mas como esses limites são muito amplos, ficam comprometidas as praticidades dos resultados desses valores.

Redes LVQ com tamanho 10x10 neurônios não foram considerados neste trabalho por não representar corretamente os padrões que existem no conjunto de medidas de nível do rio, apesar de terem sido processadas e exibidas nas tabelas e gráficos. Esta opção foi motivada porque a representatividade fica comprometida em uma rede LVQ pequena.

Os melhores resultados foram obtidos nos processamentos com média de 3 medidas, 4 ou 8 pesos por neurônio, tamanho de rede com 30x30 ou 50x50 neurônios e 50 padrões de nível de rio, considerando o conjunto de todos os experimentos. Observa-se que os processamentos mais recentes foram feitos utilizando rede LVQ de 100x100 neurônios e 100 padrões discretizados, com 8

pesos por neurônio, e mostrou bons resultados conforme pode ser observado nas Figuras 4.18 e 4.19.

Os tamanhos das redes LVQ geradas tiveram influência no resultado, porém não foram garantia de qualidade da previsão. Mas permitiram uma aprendizagem melhor quando houve mais dados para os treinamentos.

A precisão foi obtida pelo número de padrões discretizados e pela qualidade e quantidade de informações utilizados no treinamento das redes LVQ.

Portanto, com uma série temporal com qualidade e em boa quantidade, foi possível obter previsões do nível do rio em um passo adiante ou mais, através desta metodologia que utilizou rede neural artificial de tipo LVQ.



## **5 CONCLUSÕES E TRABALHOS FUTUROS**

Neste Capítulo são apresentadas as conclusões obtidas neste trabalho e também os possíveis trabalhos futuros para o prosseguimento desta iniciativa.

### **5.1 Conclusões**

Os resultados das previsões processadas dos 7 períodos foram positivos em comparação aos obtidos no artigo de Li et al (1998), pois as previsões apresentaram em várias ocasiões valores de erros relativos menores, inclusive com erros 0%.

As previsões para agosto de 2007 mostraram uma qualidade boa nos resultados, por apresentar os valores previstos em 100 padrões discretizados, pois cada padrão representa, 0,07m de faixa de nível do rio. O desvio padrão de 0,0316 desses resultados (Tabela 4.31) significa menor dispersão dos valores, portanto, resultados mais próximos das medidas reais. Nos experimentos foram constatados que a quantidade de informações utilizadas para os treinamentos da rede LVQ tem influência na qualidade das previsões, e que na impossibilidade dessa quantidade, pode-se recorrer em utilizar um número menor de padrões discretizados de nível de rio, 10 padrões por exemplo.

O parâmetro número de padrões discretizados influencia a precisão da previsão como pode ser observado nos gráficos do Capítulo 4, e sua utilização em quantidade baixa é justificável diante de alguma deficiência nos dados, como por exemplo, a pouca quantidade de medidas. Portanto quanto menor o número de padrões discretizados, maiores acertos por generalizar cada resultado e menor precisão. Recomenda-se então, utilizar um número intermediário, entre 30 a 50 padrões discretizados.



Na preparação dos dados para treinar a rede LVQ, é importante observar que se utilizar algum tipo de redução dos dados originais de forma acentuada, como exemplo uma redução anual, corre-se o risco de ocultar eventos específicos e/ou característicos, isto é, que caracterizam algum período da série temporal. Portanto pode-se reduzir, porém, com cuidados e estudos.

A possibilidade de gerar previsões a médio e longo prazo permite servir como alerta para situações em que as autoridades civis necessitem tomar medidas preventivas, antes que se tornem corretivas.

A metodologia propõe utilizar as medidas do nível do rio, anotadas e armazenadas periodicamente, processar este conjunto de medidas para extrair padrões de entrada para o treinamento da rede LVQ, e efetuar as previsões do nível a partir de medidas existentes.

A vantagem principal da metodologia apresentada neste trabalho é a possibilidade de implementação de um sistema computacional simples, que utiliza uma única fonte de dados da qual é possível se extrair informações de comportamento do rio.

Também, o uso do algoritmo LVQ permite uma calibração extremamente fácil do sistema quando novas situações surgirem.

Assim, a contribuição principal deste trabalho é o desenvolvimento de uma metodologia para o desenvolvimento de um sistema predictor baseado em modelo de rede neural artificial, que apresenta a capacidade de visualização de padrões para análises qualitativas e quantitativas.

## 5.2 Trabalhos Futuros

Sugere-se estudar a aplicação desta metodologia em outros campos do conhecimento, que processem algum tipo de fluxo ou movimentação. A metodologia utiliza medidas históricas, obtidas regularmente, para treinar uma rede LVQ que pode ser usada para previsão de medidas futuras.

Por exemplo, no controle/administração do trafego de pessoas como em (Lopes, 2003), (Souza, 2005), (Portal, 2007), (Sztrajtman, 2007) ou pode ser utilizado como em (Hallenbeck e Weinblatt, 2004) e (Capo, 2005) para auxiliar a administração do tráfego de veículos ou de cargas. Em resumo, é possível obter medidas históricas, como quantidade de pessoas em trânsito, unidades de veículos em movimento ou ainda, a quantidade de carga movimentada em algum ponto de medição, obtidas sistematicamente para treinar redes LVQ e usar essas medidas para previsões de comportamento futuro.

Também pode ser estudada a utilização da metodologia em aplicações que necessitem de respostas para previsões mais curtas, de dias para horas ou minutos, através da utilização de medidas obtidas em intervalos mais curtos no treinamento da rede LVQ.

Tornar essa metodologia em uma aplicação que tenha maior interação e interface mais amigável aos usuários pode ser considerado como extensão ao presente trabalho.



## REFERÊNCIAS BIBLIOGRÁFICAS

- ADRIAANS, P.; ZANTINGE, D. **Datamining**, Harlow: Addison Wesley Longman, 1996. 176 p.
- ANTENEODO, C. Processos estocásticos. In: Escola do CBPF, 5., 2004, Rio de Janeiro. **Anais...**, Rio de Janeiro: SBM, 2004. Disponível em: <<http://mesonpi.cat.cbpf.br/e2004/docs/PG2-cban.pdf>>. Acesso em: 19 jan. 2005.
- ATIYA, A. F.; EL-SHOURA, S.M.; SHAHEEN, S. I.; EL-SHERIF, M. S. A comparison between neural-network forecasting techniques – Case Study: River Flow Forecasting, **IEEE Transactions on Neural Networks**, v.10, n.2, mar. 1999.
- BALESTRASSI, P. P., **Identificação de padrões em gráficos de controle estatístico de processos, em tempo real, utilizando séries temporais e redes neurais artificiais**, 170f. Tese de Doutorado em Engenharia de Produção – Programa de Pós-Graduação em Engenharia de Produção, UFSC, 2000.
- BARATTI, R.; CANNAS, B. River flow forecasting for reservoir management through neural networks, **Neurocomputing**, v. 55, 3-4, p. 421-437, 2003.
- BEALE, R.; JACKSON, T., **Neural computing: an introduction**, IOP Publishing Ltd, Bristol. 1990. 240p. ISBN 0-85274-262-2
- BEM, R. A.; BOTELHO, S. S. C.; MATA, M. M. Aplicação de redes neurais artificiais na análise de padrões de variabilidade de mesoescala. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. **Anais...** São José dos Campos: INPE, 2003. p. 2007 - 2014. CD-ROM, Online. ISBN 85-17-00017-X. Disponível em: <<http://urlib.net/ltid.inpe.br/sbsr/2002/11.16.12.39>>. Acesso em: 23 nov. 2007.
- BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control**. San Francisco: Holden-Day, 1976.
- CAPO, J. M. **Gerenciamento de projetos aplicado ao transporte de cargas especiais indivisíveis**. 134 f. Dissertação de Mestrado em Gestão e Desenvolvimento Regional – Universidade de Taubaté, Departamento de Economia, Contabilidade e Administração, Taubaté, 2005.
- CARVALHO, L. M. V. **Análise de séries temporais, climatologia estatística: Notas de Aula**. Disponível em: <[www.icesb.ucesb.edu/gem/modulo\\_1.htm](http://www.icesb.ucesb.edu/gem/modulo_1.htm)>. Acesso em 23 jun. 2005 - 14h10.

CORTEZ, P. A. R., **Algoritmos genéticos e redes neuronais na previsão de séries temporais**. Dissertação de Mestrado, Universidade de Minho, Braga, Portugal, 1997.

CRISTINA, L. **Pantanal vive situação atípica de seca e cheia simultâneas**, Radiobrás, 2003. Disponível em: <<http://www.radiobras.gov.br/materia.phtml?materia=145116&editoria=CT>>. Acesso em 21 nov.2007.

EMBRAPA; GALDINO, S.; RESENDE, E. K. **Previsão de cheias e secas da Embrapa auxilia pantaneiros**. Disponível em : <<http://www.cpap.embrapa.br/publicacoes/online/ADM004.pdf>>. Acesso em 26 jul. 2005.

EMBRAPA; GALDINO, S. **Classificação e ano de cheia ou de seca**. Disponível em: <<http://tinyurl.com/33dja9>>. Acesso em 23 nov. 2007.

FAUSETT, L., **Fundamentals of neural networks – architectures, algorithms, and applications**, Upper Saddle River: Prentice Hall International, 1994. 461p. ISBN 0-13-334186-0.

FERRARI, H.; HOTO, R.; MACULAN, N.; OLIVEIRA, C.; BRUNETTO, M. A. O. Uma comparação entre redes neurais Wavelet, LMS, MLP e RBF para classificação de DPOC, In: Congresso de Matemática e suas Aplicações, Foz do Iguaçu, 2006. **Anais...** p.1-3. Disponível em: <<http://www.simulab.uel.br/hoto/publica/ferrari-hoto-foz2006.pdf>>. Acesso em: 23 nov. 2007.

FONSECA, J. S.; MARTINS, G. A.; TOLEDO, G. L. **Estatística aplicada**. São Paulo: Editora Atlas, 1988. 267p.

GOOGLE MAPS, **Ladário – MS**. Disponível em: <<http://tinyurl.com/2muvyf>>. Acesso em 23 nov. 2007.

HALLENBECK, M.; WEINBLATT, H. **National Cooperative Highway Research Program (NCHRP) Report 509: Equipment for Collecting Traffic Load Data**. Transportation Research Board, Washington D.C., 2004. 68 p. Disponível em: <[http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp\\_rpt\\_509.pdf](http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_509.pdf)>. Acesso em: 28 dez.2007.

HARVEY, A. C. **Time series models**, Oxford: Philip Allan Publishers Ltd. 1981. 229p.

Instituto Brasileiro de Geografia e Estatística/Cidades (IBGE), **Estimativa Censo 2007**. Disponível em: <<http://www.ibge.gov.br/cidadesat/default.php>>. Acesso em 23 nov.2007.

KEOGH, E. J.; PAZZANI, M. J., **Relevance feedback retrieval of time series data**. University of California, 1999. Disponível em: <[http://www.cs.ucr.edu/~eamonn/sigir99\\_keogh\\_pazzani.pdf](http://www.cs.ucr.edu/~eamonn/sigir99_keogh_pazzani.pdf)>. Acesso em: 23 nov. 2007.

KASMIER, L. J. **Estatística aplicada à economia e administração**. São Paulo: McGraw-Hill, 1982. 760p.

KISI, Ö. **Daily river flow forecasting using artificial neural networks and auto-regressive models**. Istanbul-Turkey: Istanbul Technical University, Faculty of Civil Engineering, 2003.

KOHONEN, T. **Self-organizing and associative memory**. 3. ed. Berlin: Springer-Verlag, 1989a.

KOHONEN, T. A self-learning musical grammar or 'associative memory of the second kind'. In: International Joint Conference on Neural Networks. Washington, DC. **Proceedings**...Washington DC: IEEE, 1989b. p. 1-5.

KOHONEN, T. Improved versions of learning vector quantization. In: International Joint Conference on Neural Networks, 1990a, San Diego, CA. **Proceedings**... San Diego: IEEE, 1990. p. 545-550.

KOHONEN, T. The self-organizing map. **Proceedings of the IEEE**, v.78, n.9, p.1464-1480, 1990b.

LI, W.; SÁ, L. A.; GALVÃO, G. P.; BEVILAQUA, R. M. Prediction of the level of Paraguay River using Neural Networks. **Pesquisa Agropecuária Brasileira**, v. 33, n. ESPECIAL, 1791-1797p., 1998.

LIPSCHUTZ, S., **Probabilidade**. 4. ed. São Paulo: Makron Books, 1993. 280p.

Laboratório Nacional de Computação Científica (LNCC), **Ilustração de um neurônio**. Disponível em: <[http://www.lncc.br/~labinfo/tutorialRN/frm1\\_inspiracaoBiologica.htm](http://www.lncc.br/~labinfo/tutorialRN/frm1_inspiracaoBiologica.htm)>. Acesso em 29 dez.2007.

LOPES FILHO, J. I. de O. **Pós-Análise da Previsão de Demanda por Transportes no Município de Fortaleza**. 179 fl. Dissertação de Mestrado - Programa de Mestrado de Engenharia de Transportes, Universidade Federal do Ceará, Fortaleza, 2003.

MIGON, H., **Análise de séries temporais**. Rio de Janeiro: IM/DME, UFRJ, 2004.

MINISTÉRIO DOS TRANSPORTES, **Bacia do Alto Paraguai**. Disponível em: <<http://tinyurl.com/2um53r>>. Acesso em 27 de jul. 2005.

MINISTÉRIO DOS TRANSPORTES. **Perfil longitudinal do Rio Paraguai**. Disponível em: <<http://tinyurl.com/3xve9w>>. Acesso em 02 out. 2006.

MORETIN, P. A., TOLOI, C. M. C. **Modelos para séries temporais**. Rio de Janeiro: Instituto de Matemática Pura e Aplicada. Rio de Janeiro, 1981.

MULLER, M.; FILL, H. D., Redes neurais aplicadas na propagação de vazões, In: XV Simpósio Brasileiro de Recursos Hídricos, 15, 2003, Curitiba.

**Anais**....Disponível em

<[http://www.lactec.org.br/OInstituto/downloads/Biblioteca/2003/065\\_2003.pdf](http://www.lactec.org.br/OInstituto/downloads/Biblioteca/2003/065_2003.pdf)>. Acesso em 20 nov. 2007.

NETO, L.B., COELHO, P.H.G., CHIGANER, L., MEZA, L.A., MONTEIRO, L.P, Previsão da vazão afluente mensal do Rio São Francisco usando-se redes de Elman, **Revista Engevista**, v.7, n.1, p.63-72, abr. 2005.

NORDEMANN, D. J. R. Periodicidades e tendências nas medias mensais do nível do Rio Paraguai em Corumbá de 1900-1995. In: Encontro sobre Sensoriamento Remoto aplicado a estudos no Pantanal, 1995, Corumbá.

**Anais eletrônicos**... 1995a. p. 69-72. (INPE-6255-pre/2345). Disponível em: <<http://urlib.net/sid.inpe.br/iris@1912/2005/07.20.07.30>>. Acesso em: 27 nov. 2007.

NORDEMANN, D. J. R. Previsão baseada no estudo das medias anuais do nível do rio Paraguai em Corumbá de 1900 a 1995. In: Encontro sobre Sensoriamento Remoto aplicado a estudos no Pantanal, 1995, Corumbá.

**Anais eletrônicos**... 1995b. p. 73-76. (INPE-6256-PRE/2346). Disponível: <<http://urlib.net/sid.inpe.br/iris@1912/2005/07.20.07.31>>. Acesso em: 27 nov. 2007.

NORDEMANN, D.J.R.; Periodicidades, tendências e previsão a partir da análise espectral dinâmica da série dos níveis do Rio Paraguai, em Ladário (1900/1995). **EMBRAPA**, v. 33, n. Especial, 1787-1790p., out. 1998.

OLIVEIRA, M.M.; EBECKEN, N.F.F.; SANTOS, I. A.; NEVES, C.F.; CALOBA, L.P.; OLIVEIRA, J.L.F. Modelagem da maré meteorológica utilizando redes neurais artificiais: uma aplicação para a Baía de Paranaguá – PR, Parte 2: Dados Meteorológicos de Reanálise do NCEP/NCAR, **Revista Brasileira de Meteorologia**, v.22, n.1, 53-62p., 2007.

OLIVO, A. A., **Modelos matemáticos para a previsão de cheias fluviais**. 2004. 151f. Tese de Doutorado em Ciência - Instituto Tecnológico da Aeronáutica, São José dos Campos, 2004.

PHIEN, H. N.; DANH, N. T.; GUPTA, A. D. Neural network models for river flow forecasting, **Water SA**, v.25, n. 1, 33-40p., 1999. Disponível em: <[http://www.wrc.org.za/archives/watersa%20archive/1999/January/jan99\\_p33.pdf](http://www.wrc.org.za/archives/watersa%20archive/1999/January/jan99_p33.pdf)>. Acesso em 21 jun. 2005.

PORTAL DA PROPAGANDA, **Alta elevada no fluxo de pessoas no Natal**. Pixtron Comunicação Digital. Bonilha Comunicação e Marketing, Curitiba, 2007. Disponível em: <[http://www.portaldapropaganda.com.br/portal/index.php?option=com\\_frontpage&Itemid=1](http://www.portaldapropaganda.com.br/portal/index.php?option=com_frontpage&Itemid=1)>. Acesso em 27 dez.2007.

SILVA, J. D. S, **Neurocomputação**. INPE, São José dos Campos, 2005. Notas de aula, 2 ago.2004.

SIMON, G.; LENDASSE, A.; COTTRELL, M.; FORT, J.C.; VERLEYSEN, M. Time series forecasting: obtaining long term trends with self-organizing maps, pattern recognition letters. **Artificial Neural Networks in Pattern Recognition**, v. 26, n. 12, p.1795-1808, 2005.

SOARES, J. F.; FARIAS, A. A.; CESAR, C. C. **Introdução à estatística**, Rio de Janeiro: Guanabara Koogan, 1991.

SOUZA, S. G. de. **Um Algoritmo Descentralizado para Controle de Tráfego Urbano em Tempo Real**. 2005. 112f. Tese de Doutorado em Engenharia Elétrica. Universidade Federal de Santa Catarina, Florianópolis, 2005. Disponível em:<<http://www.tede.ufsc.br/teses/PEEL1031.pdf>>. Acesso em 27 dez.2007.

SZTRAJTMAN, M. J. **Medição de tráfego em shopping centers**. Disponível em: <[http://www.smartcount.com.br/documentos/Shopping\\_center\\_medicao.pdf](http://www.smartcount.com.br/documentos/Shopping_center_medicao.pdf)>. Acesso em: 27 dez.2007.

SPIEGEL, Murray R., **Estatística**, Tradução e revisão técnica Pedro Consentino – 3. ed. São Paulo: Makron Books, 1993.

TUCCI, C.E.M. **Modelos hidrológicos**. Porto Alegre, RS: UFRGS, 1998. 669p. ISBN 85-7025-445-8.