



UNIVERSIDADE FEDERAL DA PARAÍBA

**CENTRO DE CIÊNCIAS E TECNOLOGIA
CURSO DE PÓS GRADUAÇÃO EM ENGENHARIA CIVIL
CAMPUS II - CAMPINA GRANDE**

**DESAGREGAÇÃO ESPACIAL DA PRECIPITAÇÃO SIMULADA
POR MODELOS ATMOSFÉRICOS NO NORDESTE DO BRASIL**

ALESSANDRA MARIA RAMOS

CAMPINA GRANDE

DEZEMBRO - 2000

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
CURSO DE PÓS – GRADUAÇÃO EM ENGENHARIA CIVIL

DESAGREGAÇÃO ESPACIAL DA PRECIPITAÇÃO SIMULADA
POR MODELOS ATMOSFÉRICOS NO NORDESTE DO BRASIL

ALESSANDRA MARIA RAMOS

CAMPINA GRANDE – PB

DEZEMBRO – 2000

ALESSANDRA MARIA RAMOS

**DESAGREGAÇÃO ESPACIAL DA PRECIPITAÇÃO SIMULADA POR
MODELOS ATMOSFÉRICOS NO NORDESTE DO BRASIL**

**Dissertação apresentada ao Curso de Pós –
Graduação em Engenharia Civil, na área de
Engenharia de Recursos Hídricos, em
cumprimento às exigências para obtenção
do Grau de Mestre**

ÁREA DE CONCENTRAÇÃO: RECURSOS HÍDRICOS

ORIENTADOR: CARLOS OLIVEIRA GALVÃO

Campina Grande – PB

Dezembro – 2000

R175d Ramos, Alessandra Maria.
Desagregação espacial da precipitação simulada por modelos atmosféricos no nordeste do Brasil / Alessandra Maria Ramos. - Campina Grande, 2000.
93 f.

Dissertação (Mestrado em Engenharia Civil) - Universidade Federal da Paraíba, Centro de Ciências e Tecnologia, 2000.
"Orientação : Prof. Dr. Carlos de Oliveira Galvão".
Referências.

1. Engenharia Hidráulica - Nordeste - Brasil. 2. Precipitação Simulada. 3. Recursos Hídricos. 4. Modelo Atmosférico. 5. Dissertação - Engenharia Civil. I. Galvão, Carlos de Oliveira. II. Universidade Federal da Paraíba - Campina Grande (PB). III. Título

CDU 626(812/814)(043)

ALESSANDRA MARIA RAMOS

**DESAGREGAÇÃO ESPACIAL DA PRECIPITAÇÃO SIMULADA
POR MODELOS ATMOSFÉRICOS NO NORDESTE DO BRASIL**

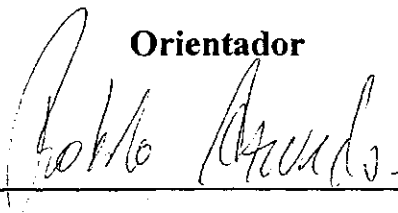
Aprovada em 12 de Dezembro de 2000

COMISSÃO EXAMINADORA



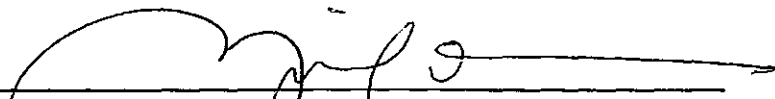
Prof. Dr. Carlos de Oliveira Galvão

Orientador



Dr. José Roberto Gonçalves de Azevedo

Examinador



Dr Manoel Francisco Gomes Filho

Examinador

*A meus pais,
Sandra e Hermilio*

AGRADECIMENTOS

Aos professores Tarciso Cabral e Alain Passerat, com os quais conheci a área de Recursos Hídricos e muito aprendi.

Ao professor Heber Pimentel que, com seu convite e entusiasmo, trouxe-me ao mestrado em Campina Grande.

Ao professor Sérgio Góis, que me acolheu em seu ambiente de trabalho, enquanto aluna especial.

Aos meus colegas Paula Freire e Dimitri Pinto, com os quais compartilhei dúvidas e soluções.

A Carlos Galvão, que além de grande orientador, soube ser um amigo essencial para a minha formação.

Ao professor Manoel Gomes Filho que, sempre prestativo e atencioso, contribuiu para a minha pesquisa desde o início.

A todos os funcionários, professores e alunos da área de Recursos Hídricos da UFPB em Campina Grande, que ajudaram direta ou indiretamente no meu trajeto, e pelos quais adquiri grande carinho.

À minha família e a André Pineli pela compreensão e apoio constantes.

RESUMO

A importância de previsões de longo prazo de precipitação para o gerenciamento dos recursos hídricos de uma região, e a comprovação da boa previsibilidade de um modelo de circulação global (MCG) para uma região do Nordeste do Brasil, motivaram o desenvolvimento desta dissertação. Para o aproveitamento dessas previsões em recursos hídricos, o problema da diferença de escalas do modelo atmosférico e das bacias hidrográficas, espacial e temporal, vem sendo solucionado através da utilização de técnicas de desagregação espacial. Esta dissertação se propôs a, numa bacia hidrográfica do semi-árido nordestino, melhorar a desagregação espacial de precipitação sazonal e mensal, utilizando-se para isto, o método das Redes Neurais Artificiais (RNAs). Constatou-se que a desagregação da precipitação sazonal na bacia do Rio Piancó apresentou melhoria, em comparação a um estudo anterior, e que, tanto no nível sazonal quanto no mensal, obteve-se altas correlações entre os dados observados e os valores obtidos na modelagem, indicando que o método utilizado é uma técnica adequada de desagregação.

ABSTRACT

This dissertation was motivated by the importance of long-range precipitation forecasting for regional water resources management, and by recent research that showed good precipitation predictability by an Atmospheric General Circulation Model for a region in north-eastern Brazil. For applying these forecasts in water resources, the problem of difference of scales, spatial and temporal, between the atmospheric model and the drainage basins has been dealt with the so-called downscaling techniques. This work aimed to improve downscaling of seasonal and monthly precipitation, for a drainage basin in the semiarid north-eastern Brazil, using an Artificial Neural Network (ANN) technique. The seasonal downscaling was improved in comparison with a previous study for the same basin and, both in seasonal and monthly time scales, high correlations between modelled and observed data were obtained, thus indicating that the ANN technique is adequate for downscaling.

SUMÁRIO

DEDICATÓRIA.....	i
AGRADECIMENTOS.....	ii
RESUMO.....	iii
ABSTRACT.....	iv
SUMÁRIO.....	v
LISTA DE TABELAS.....	vii
LISTA DE FIGURAS.....	viii
1. INTRODUÇÃO.....	1
2. REVISÃO DE LITERATURA.....	4
2.1. CARACTERIZAÇÃO DO CLIMA DO NORDESTE DO BRASIL.....	4
2.2. PREVISÃO DE LONGO PRAZO DE PRECIPITAÇÃO.....	8
2.3. DESAGREGAÇÃO ESPACIAL DE PRECIPITAÇÃO.....	12
2.4. REDES NEURAS ARTIFICIAIS.....	14
3. METODOLOGIA.....	18
3.1. DELIMITAÇÃO DA REGIÃO DE ESTUDO.....	18
3.2. COLETA DE DADOS PLUVIOMÉTRICOS.....	22
3.3. DIAGNÓSTICO DAS RELAÇÕES DA PRECIPITAÇÃO ENTRE AS DUAS ESCALAS.....	24
3.4. SELEÇÃO DO MÉTODO DE DESAGREGAÇÃO.....	24
3.5. CALIBRAÇÃO E VALIDAÇÃO DOS MODELOS.....	25
3.5.1 <i>Regressão Linear Simples</i>	25
3.5.2 <i>Redes Neurais Artificiais</i>	26
4. RESULTADOS.....	35
4.1. DIAGNÓSTICO DAS RELAÇÕES DA PRECIPITAÇÃO ENTRE AS DUAS ESCALAS.....	35
4.1.1. <i>Nível Sazonal</i>	35
4.1.2. <i>Nível Mensal</i>	38
4.2. DESAGREGAÇÃO VIA REGRESSÃO LINEAR.....	45
4.2.1. <i>Nível Sazonal</i>	45
4.2.2. <i>Nível Mensal</i>	47

4.3. DESAGREGAÇÃO VIA REDES NEURAI ARTIFICIAIS	52
4.3.1. <i>Redes Neurais Artificiais como Regressão Linear</i>	52
4.3.2. <i>Pré-processamento dos dados de entrada de uma Rede Neural</i>	54
4.3.3. <i>Modelagem das Redes Neurais</i>	58
4.3.4. <i>Nível Sazonal</i>	58
4.3.5. <i>Nível Mensal</i>	63
4.3.6. <i>Importância da análise gráfica</i>	79
4.3.7. <i>Ocorrência de Superajustamento</i>	80
4.3.8. <i>Comentários finais sobre o método de redes neurais</i>	84
4.4. COMPARAÇÃO ENTRE OS MÉTODOS ADOTADOS PARA A DESAGREGAÇÃO	84
5. CONCLUSÃO E RECOMENDAÇÕES.....	86
REFERÊNCIAS BIBLIOGRÁFICAS.....	88

LISTA DE TABELAS

Tabela 3.1 – Postos pluviométricos da bacia do Rio Piancó.....	22
Tabela 3.2 – Postos pluviométricos da célula do MCG sobre a bacia do Rio Piancó	23
Tabela 4.1 - Coeficientes de correlação obtidos entre as séries mensais para os períodos de 1911-88 e 1939-88.....	43
Tabela 4.2 – Coeficientes de correlação obtidos entre séries observadas e calculadas pela regressão linear.....	49
Tabela 4.3 – Coeficientes de correlação obtidos nas simulações das estruturas de rede selecionadas para a verificação da estabilidade do algoritmo de treinamento, para o nível sazonal.....	59
Tabela 4.4 – Valores pelos quais as séries mensais foram normalizadas	63
Tabela 4.5 – Coeficientes de correlação obtidos nas simulações das estruturas de rede selecionadas, para os meses de fevereiro a maio.....	63
Tabela 4.6 – Coeficientes de correlação obtidos das redes julgadas mais adequadas para a desagregação, nos meses de fevereiro a maio	65
Tabela 4.7 – Coeficientes de correlação obtidos entre a precipitação observada da bacia do Piancó e os resultados da Regressão linear e das Redes Neurais, nos períodos de treinamento e validação	85

LISTA DE FIGURAS

Figura 2.1 – Precipitação anual média.....	5
Figura 2.2 – Distribuição espacial do percentual da precipitação para o quadrimestre chuvoso (FMAM), em relação ao total anual, sobre a região do semi-árido nordestino.....	6
Figura 2.3 - Sub-regiões climáticas do Nordeste do Brasil	8
Figura 2.4 – Exemplo da estrutura de uma RNA multicamada com três camadas	15
Figura 2.5 – Rede de Elman.....	16
Figura 3.1 - Malha do modelo do ECMWF	19
Figura 3.2 - Mapas esquemáticos (fora de escala) mostrando a célula do MCG sobre a Bacia do Rio Piancó, e os respectivos postos pluviométricos	19
Figura 3.3 – Comparação entre o total precipitado de março-junho observado e previsto pelo MCG do ECMWF para a célula sobre a bacia do Piancó	20
Figura 3.4 – Possibilidades de redução de escala para a região de estudo.....	21
Figura 3.5 – Reta de ajuste da Regressão Linear obtida entre duas variáveis.....	25
Figura 3.6 – Fluxograma do processo de treinamento e validação de uma RNA com aprendizado supervisionado.....	27
Figura 3.7 – Configuração da rede	31
Figura 3.8 – Fluxograma do processo de seleção da rede neural para a desagregação espacial	32
Figura 3.9 – Algoritmo, gráfico e símbolo utilizados pela função TANSIG	34
Figura 3.10 – Algoritmo, gráfico e símbolo utilizados pela função PURELIN	34
Figura 4.1 – Séries temporais de precipitação sazonal (fevereiro-maio) da célula e da bacia do Rio Piancó (1911-88).....	35
Figura 4.2 – Relação entre dados sazonais (fevereiro-maio) de precipitação da célula e da bacia, para o período de 1911 a 1988	36
Figura 4.3 – Coeficientes de correlação entre as séries de precipitação sazonal da bacia e da célula, para uma janela móvel de 30 anos.....	37
Figura 4.4 – Coeficientes de correlação entre as séries sazonais de precipitação da bacia e da célula, para uma janela móvel de 30 anos, iniciando entre 1936 e 1941	37
Figura 4.5 – Relação entre dados sazonais (fevereiro-maio) de precipitação da célula e da bacia, para o período de 1939 a 1988	38
Figura 4.6 – Séries temporais de precipitação da célula e da bacia do Rio Piancó, para os meses de: a) fevereiro; b) março; c) abril; e d) maio	39

Figura 4.7 – Relação entre dados mensais de precipitação da célula e da bacia, para o período de 1911 a 1988, nos meses de: a) fevereiro; b) março; c) abril; e d) maio	41
Figura 4.8 – Coeficientes de correlação entre as séries mensais de precipitação da bacia e da célula, para uma janela móvel de 30 anos	42
Figura 4.9 – Coeficientes de correlação entre as séries mensais de precipitação da bacia e da célula, para uma janela móvel de 30 anos, iniciando entre 1936 e 1941	43
Figura 4.10 – Relação entre dados mensais de precipitação da célula e da bacia, para o período de 1939 a 1988, nos meses de: a) fevereiro; b) março; c) abril; d) maio.....	44
Figura 4.11 – Calibração do modelo de Regressão Linear entre dados de precipitação sazonal, para o período de 1939 a 1968.....	45
Figura 4.12 – Validação modelo de Regressão Linear entre dados de precipitação sazonal, para o período de 1969 a 1988	46
Figura 4.13 – Comparação entre o total precipitado sazonal observado e calculado pelo modelo de Regressão Linear para a bacia do Piancó.....	46
Figura 4.14 – Calibração do modelo de Regressão Linear entre dados de precipitação mensal, para o período de 1939 a 1968, nos meses de: a) fevereiro; b) março; c) abril; d) maio.....	48
Figura 4.15 – Validação do modelo de Regressão Linear entre dados de precipitação mensal, para o período de 1969 a 1988, nos meses de: a) fevereiro; b) março; c) abril; d) maio.....	49
Figura 4.16 – Comparação entre o total precipitado observado e calculado pelo modelo de Regressão Linear, para a bacia do Piancó, nos meses de: a) fevereiro; b) março; c) abril; d) maio.....	50
Figura 4.17 – Configuração da rede composta por neurônios lineares.....	52
Figura 4.18 – Resultados obtidos em 10 simulações de uma rede neural com 5 neurônios sigmóides na camada escondida e 1 neurônio linear na camada de saída, em totais precipitados.....	56
Figura 4.19 – Relações entre as séries normalizadas sazonais da célula e da bacia, para o conjunto de dez redes com 4 neurônios na camada escondida, no período de a) treinamento; e de b) validação	60
Figura 4.20 – Relações resultantes da Rede 11 entre as séries normalizadas sazonais da célula e da bacia, no período de a) treinamento; e de b) validação.....	61
Figura 4.21 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, nos períodos de: a) treinamento e b) validação.....	62

Figura 4.22 – Relações entre as séries normalizadas da célula e da bacia para o conjunto de dez redes, no período de treinamento (1939-68), referentes aos meses de a) fevereiro; b) março; c) abril; d) maio.....	67
Figura 4.23 – Relações entre as séries normalizadas da célula e da bacia para o conjunto de dez redes, no período de validação (1969-88), referentes aos meses de a) fevereiro; b) março; c) abril; d) maio.....	69
Figura 4.24 – Relações resultantes da Rede 13 entre as séries normalizadas da célula e da bacia, referentes ao mês de Fevereiro, no período de a) treinamento; e de b) validação	71
Figura 4.25 – Relações resultantes da Rede 26 entre as séries normalizadas da célula e da bacia, referentes ao mês de Março, no período de a) treinamento; e de b) validação	72
Figura 4.26 – Relações resultantes da Rede 14 entre as séries normalizadas da célula e da bacia, referentes ao mês de Abril, no período de a) treinamento; e de b) validação.....	73
Figura 4.27 – Relações resultantes da Rede 38 entre as séries normalizadas da célula e da bacia, referentes ao mês de Maio, no período de a) treinamento; e de b) validação.....	74
Figura 4.28 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, no mês de Fevereiro, nos períodos de: a) treinamento e b) validação	75
Figura 4.29 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, no mês de Março, nos períodos de: a) treinamento e b) validação.....	76
Figura 4.30 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, no mês de Abril, nos períodos de: a) treinamento e b) validação.....	77
Figura 4.31 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, no mês de Maio, nos períodos de: a) treinamento e b) validação.....	78
Figura 4.32 – Relações resultantes da Rede 37 entre as séries normalizadas da célula e da bacia, referentes ao mês de Maio, no período de a) treinamento; e de b) validação	79
Figura 4.33 – Validação de redes neurais com 9 neurônios na camada escondida correspondentes a treinamentos que obtiveram coeficientes de correlação de: a) 0,9846; b) 0,9849; e c) 0,9850	80
Figura 4.34 – Fluxograma do processo de treinamento e validação de uma RNA com aprendizado supervisionado.....	83

1. INTRODUÇÃO

Nos últimos anos, o aumento da população, a necessidade de alimentos e energia, e o crescimento da atividade industrial têm aumentado significativamente a demanda de água, e o atendimento a esta demanda constitui, atualmente, um dos problemas mais importantes em muitas partes do mundo.

Segundo levantamentos realizados pela Organização Meteorológica Mundial das Nações Unidas (WMO, 1997 *apud* Freitas e Santos, 1999), um terço da população mundial vive em regiões de moderado a alto estresse hídrico, apresentando um nível de consumo superior a 20% de sua disponibilidade d'água. Neste cenário, o Brasil possui a maior disponibilidade hídrica do planeta, correspondendo a aproximadamente 14% do deflúvio médio mundial (Freitas e Santos, 1999). No entanto, de acordo com informações coletadas pelo PROÁGUA, programa do Governo Federal sob a coordenação da Secretaria de Recursos Hídricos do Ministério de Meio Ambiente, do percentual de disponibilidade hídrica detido pelo Brasil, o Nordeste do país possui apenas 3% do total (MMA, 2000). Acrescendo-se a isto o fato de que se trata da segunda região mais populosa do país, constata-se então um quadro preocupante, principalmente na região semi-árida nordestina, que apresenta índices pluviométricos inferiores a 800 mm/ano e elevadas taxas de evapotranspiração, além de possuir solos rasos sobre embasamento cristalino, que não favorecem a um aproveitamento de recursos hídricos subterrâneos nesta região, visto que, em tal caso, a existência de água subterrânea se verifica ou em pequenas manchas sedimentares ou nas fraturas das rochas, que em sua grande maioria, abrigam águas com alto teor de salinidade (Costa e Costa, 1997).

Agravando a situação, os mecanismos climáticos que produzem as precipitações do Nordeste são complexos e, aliados a fatores como a orografia da região e a proximidade do mar, resultam numa forte variabilidade espaço-temporal das precipitações, que pode gerar eventos extremos como grandes cheias ou secas devastadoras. Felizmente, hoje já existe a possibilidade de se prever, com alguma antecedência, as probabilidades de ocorrência dos

eventos extremos e os períodos em que provavelmente ocorrerão, informações que podem ser usadas para minimizar os efeitos negativos desses eventos.

Vários estudos têm sido feitos e, apesar de toda a complexidade, os mecanismos oceânicos e atmosféricos que determinam a maior parte da variabilidade climática nas regiões tropicais, já estão bem caracterizados, possibilitando a realização nessas regiões, entre as quais o Nordeste do Brasil (NEB), de previsões de precipitação de longo prazo (uma semana a alguns meses de antecedência). Esses estudos têm sido motivados pela observação de que a precipitação no NEB é influenciada pelos eventos El Niño-Oscilação Sul (ENOS), pelo gradiente de temperatura da superfície do mar (TSM) entre Oceano Atlântico sul e norte, e pela migração latitudinal da Zona de Convergência Intertropical (ZCIT) (Uvo, 1998).

A importância de previsões de longo prazo de precipitação para o gerenciamento dos recursos hídricos da região, e a comprovação da viabilidade de se empregar previsões fornecidas por modelos de circulação global atmosférica (MCGs) em escalas menores, como a de uma bacia hidrográfica no Nordeste do Brasil (Galvão, 1999), motivaram o desenvolvimento desta dissertação. Entre os principais problemas para a efetiva aplicação da previsão meteorológica em recursos hídricos está a diferença entre a escala para a qual são realizadas as previsões, e a escala da região onde se deseja empregá-las. Esta questão tem sido contornada utilizando-se técnicas de transferência da informação meteorológica para as escalas menores, as chamadas técnicas de *downscaling* ou desagregação espacial.

Para reduzir o erro associado às previsões, que torna-se maior quando a escala temporal de interpretação é reduzida, os resultados dos modelos climáticos, normalmente fornecidos em nível diário, têm sido agrupados em escalas temporais maiores, geralmente para prever o comportamento da precipitação para a estação chuvosa do ano, denominada de precipitação sazonal. Por conseguinte, a desagregação também tem sido efetuada em escalas maiores, como a sazonal. Contudo, é provável que os MCGs também simulem razoavelmente bem a precipitação mensal. Galvão (1999) mostrou que, no nível sazonal, para uma região no Nordeste do Brasil, as simulações de precipitação por MCGs e a sua desagregação para escalas menores têm boa precisão. No entanto, ainda assim o autor sugere que a desagregação espacial deve ser aprimorada, considerando modelos mais apropriados que possam levar a uma melhoria das previsões.

Diante dos problemas aqui citados, para uma região do Nordeste brasileiro, os objetivos desta dissertação são: (i) melhorar a desagregação espacial de precipitação

sazonal, (ii) realizar a desagregação espacial de precipitação mensal, referente aos meses da estação chuvosa e (iii) avaliar o comportamento do método de Redes Neurais Artificiais como técnica de desagregação espacial.

2. REVISÃO DE LITERATURA

2.1. CARACTERIZAÇÃO DO CLIMA DO NORDESTE DO BRASIL

A enorme extensão territorial da Região Nordeste (1.540.827 km²), o relevo – constituído por amplas planícies (baixadas litorâneas), por vales baixos, geralmente inferiores a 500 m, entre superfícies que se alçam, muitas vezes, a cotas de 800 m na Borborema, Araripe, Ibiapaba e de 1.200 m na Diamantina – somado à conjugação de diferentes sistemas de circulação atmosférica, tornam a climatologia desta região uma das mais complexas do mundo (Nimer, 1989). Esta complexidade se traduz numa grande variabilidade espacial pluviométrica, com médias anuais que decrescem da costa para o interior da região (Figura 2.1). No litoral oriental nordestino, áreas possuem precipitação média anual em torno de 1.500 mm/ano (de Pernambuco a Sergipe, e do Recôncavo Baiano ao extremo sul da Bahia). No sertão, por sua vez, os índices são inferiores a 1.000 mm/ano, havendo locais (raso da Catarina, Cariri e depressão de Patos, na Paraíba) onde o índice pluviométrico cai para menos de 500 mm/ano.

A variabilidade espacial da precipitação no NEB delimita três áreas distintas na região: a zona da mata, o agreste, e o semi-árido. A zona da mata corresponde a uma estreita faixa costeira que se estende do Rio Grande do Norte ao sul da Bahia, e onde o clima é tropical quente e úmido e a pluviometria varia de 1.200 a 2.500 mm por ano. O agreste abrange a região imediatamente a oeste da zona da mata, e é considerado uma área de transição por apresentar características intermediárias às que estão presentes nas outras duas regiões delimitadas. O setor que constitui o semi-árido é o mais complexo do NEB, apresentando características que limitam seu potencial produtivo, como a evapotranspiração elevada, chuvas escassas, e solos rasos com embasamento rochoso aflorante e alta salinidade.

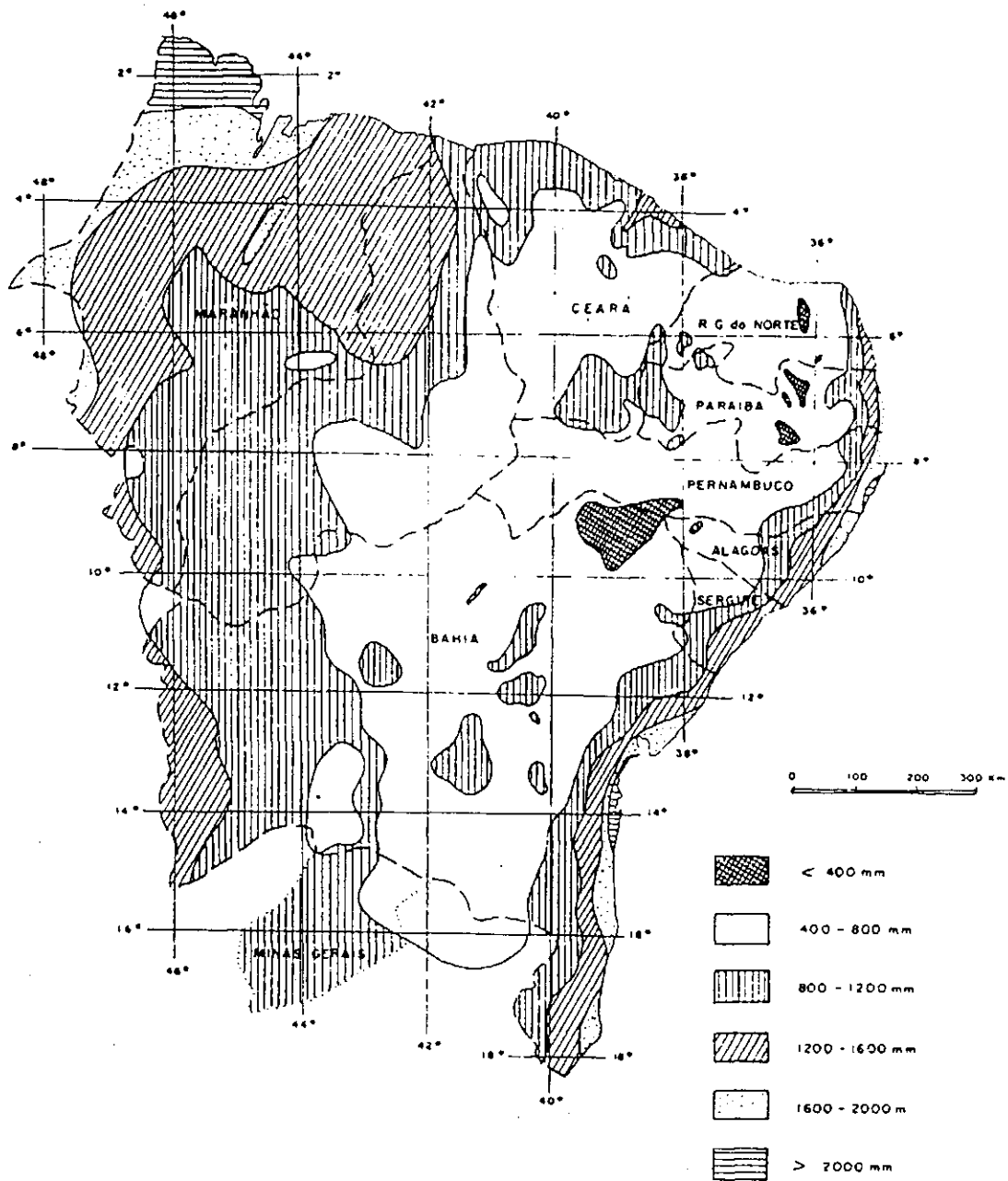


Figura 2.1 – Precipitação anual média (Cadier, 1994)

A maior parte do total pluviométrico anual observado sobre o semi-árido nordestino concentra-se em quatro meses do ano, fevereiro a maio (FMAM), compreendendo a estação chuvosa dessa região. Souza *et al.* (1998) mostraram que, com relação ao total

anual, o percentual de precipitação observado no período de FMAM é em torno de 40% a 60% na parte sul do semi-árido nordestino, e de 60% a 85% no setor centro-norte da região (Figura 2.2). Entre os meses de fevereiro e abril, a Zona de Convergência Intertropical (ZCIT), que constitui o principal sistema meteorológico indutor de chuvas no semi-árido nordestino, atinge sua posição mais ao sul, próxima da região, aumentando a instabilidade atmosférica e ocasionando as chuvas da estação. Esse sistema tem sua origem no encontro dos ventos alísios provenientes dos hemisférios norte e sul, e sua migração latitudinal e intensidade são consideradas os principais fatores que ditam o comportamento da estação chuvosa na região (Roucou *et al.*, 1996). Segundo Uvo (1998), a posição da ZCIT - que é um dos sistemas que mais interferem na precipitação sobre o Nordeste - é altamente variável, e é esta variabilidade a causa direta da distribuição espacial e temporal da precipitação sobre a região. Conforme o referido autor, a posição da ZCIT nos meses de fevereiro a abril, é determinada por um grupo de fatores que incluem os gradientes de TSM e de pressão ao nível do mar entre o Atlântico norte e sul, as TSMs sobre o Pacífico tropical e a Oscilação do Sul. Mostrou-se, ainda, que existem significantes relações intrasazonais entre as TSMs do Pacífico e do Atlântico e a precipitação do Nordeste.

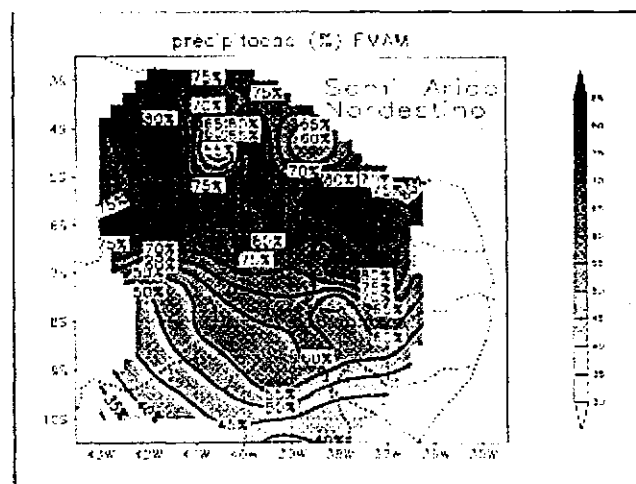


Figura 2.2 – Distribuição espacial do percentual da precipitação para o quadrimestre chuvoso (FMAM), em relação ao total anual, sobre a região do semi-árido nordestino (Souza *et al.*, 1998)

Além da variabilidade espacial da precipitação, o NEB apresenta uma grande variabilidade interanual desta, com flutuações atingindo até 40% da média (Uvo *et al.*, 1998). Em regiões semi-áridas do NEB, estas flutuações acontecem com maior frequência e maior intensidade. Cadier (1994) cita a bacia experimental de Sumé, localizada no semi-árido paraibano, como exemplo dessa variabilidade: enquanto no ano de 1983 precipitou apenas um total de 248 mm, em 1985 a precipitação anual foi de 1.438 mm.

A característica mais negativa da precipitação no NEB, no entanto, não reside nos seus totais, mas na sua irregular distribuição anual, onde se constata, inclusive, uma nítida variabilidade intrasazonal. De acordo com Kousky (1985), Ramos (1975) observou que um número pequeno de eventos de precipitação pode ocorrer em poucos dias numa estação chuvosa, ficando o restante do período sem nenhuma ocorrência de precipitação. Isto se deve à atuação de diferentes sistemas de precipitação (como a ZCIT, por exemplo) sobre a região nos meses da estação chuvosa.

Diferentes regimes de precipitação podem ser identificados no NEB. Para fins de análise climática, a região tem sido dividida em quatro sub-regiões (Galvão, 1999): o norte, o leste, o sul do NEB e a pré-Amazônia (Figura 2.3). O norte do Nordeste (norte-NEB) é composto pela região semi-árida dos Estados do Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas e norte da Bahia. Possui como quadra chuvosa os meses de fevereiro a maio. É a região que apresenta, no NEB, a maior variabilidade interanual de precipitação, a maior escassez hídrica, e tem sido a mais estudada climática e hidrologicamente (Brito *et al.*, 1991; Alves e Repelli, 1992; Alves *et al.*, 1994). A parte leste do NEB compreende a faixa litorânea dos Estados do Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe e Bahia, equivalendo, aproximadamente, à região da zona da mata. Possui um regime de chuvas satisfatório, em termos de volume e regularidade, porém ainda presencia a ocorrência de secas e enchentes em alguns anos. Tem sua estação chuvosa entre maio e agosto, e apesar de menos estudada do que o norte-NEB, dispõe de alguns trabalhos realizados sobre a mesma (Rao *et al.*, 1993; Lima e Rao, 1994; Herdies *et al.*, 1994). O sul do Nordeste (sul-NEB) tem sua precipitação máxima entre novembro e dezembro e, semelhante à sub-região pré-Amazônia, carece de estudos para uma melhor compreensão sobre os mecanismos que provocam a variabilidade interanual da precipitação na sub-região.

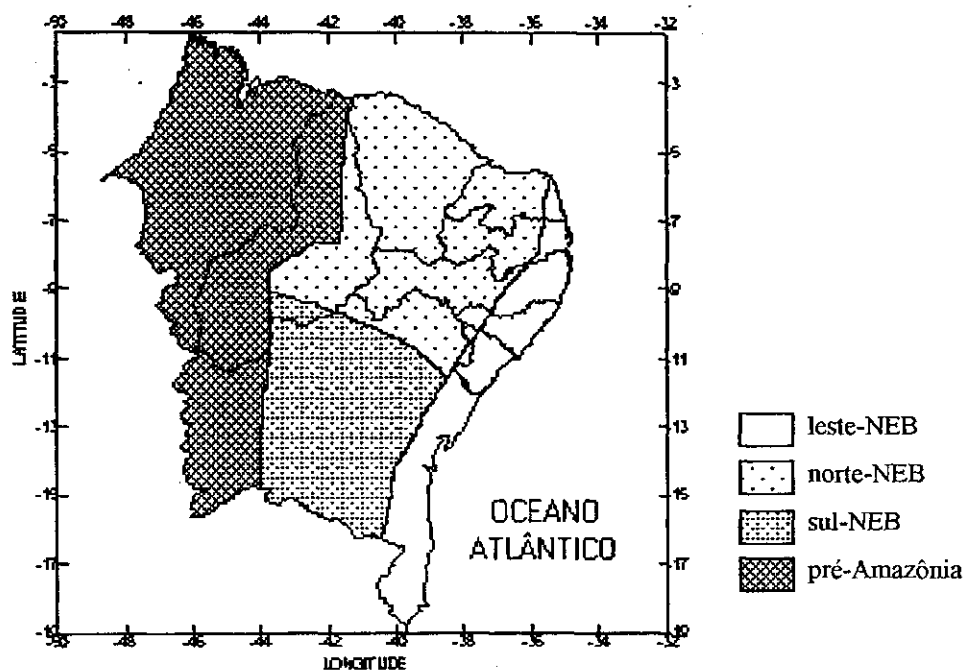


Figura 2.3 - Sub-regiões climáticas do Nordeste do Brasil (adaptada de Nobre e Molion, 1988)

2.2. PREVISÃO DE LONGO PRAZO DE PRECIPITAÇÃO

A previsão climática de longo prazo é uma estimativa do comportamento médio da atmosfera com antecedência de um a seis meses, que pode ser produzida em escalas temporais diferentes, como a diária, mensal ou sazonal. A disponibilização desse tipo de informação é de grande importância nos dias de hoje, pois pode levar à preparação de medidas preventivas contra eventos extremos, como cheias e secas. O Nordeste do Brasil tem recebido considerável atenção de pesquisadores da área de previsão meteorológica devido a sua grande variabilidade interanual de precipitação (Kousky, 1985), aliada a uma alta variabilidade espaço-temporal da mesma em cada estação chuvosa (Uvo *et al.*, 1998).

Segundo Brankovic *et al.* (1994), as regiões tropicais apresentam maior índice de acerto a nível mundial, quando se trata de previsões sazonais, pois os fatores que determinam os fenômenos meteorológicos são diretamente influenciados pelas condições à superfície, tais como a temperatura da superfície do mar (TSM) e a umidade do solo nos continentes, que variam lentamente no tempo, e hoje já estão razoavelmente

caracterizados. De acordo com experimentos do programa de pesquisa sobre previsibilidade sazonal PROVOST – *Prediction of climate variations on seasonal and interannual timescales*, conduzidos por quatro centros de pesquisa europeus, a área sobre o globo terrestre que possui maior previsibilidade se estende por grande parte do leste do Brasil – com destaque para o nordeste do país – e atinge a Bacia Amazônica, a Guiana Francesa, o Suriname, a Guiana e parte do sudeste do Caribe (Evans e Graham, 1999). Outras regiões, como o leste da África (Brankovic e Palmer, 1998), a Índia (Sperber e Palmer, 1996) e o Sahel africano (Folland et al., 1991), estreita faixa de terra situada ao sul do deserto do Saara, também se mostraram áreas de boa capacidade preditiva para precipitação sazonal.

Uma vasta literatura existente sobre o assunto (Harzallah *et al.*, 1996; Sperber e Palmer, 1996; Uvo *et al.*, 1998; entre outros) comprova que muito da variabilidade atmosférica verificada em escala sazonal é originada por anomalias da TSM. A mais importante fonte de variabilidade relevante da TSM é o *El Niño*, que consiste no aquecimento anômalo das águas superficiais do Oceano Pacífico Equatorial. Fenômeno de escala global, é cíclico porém não possui período regular. Quando ocorre, afeta a circulação atmosférica e, associado ao fenômeno Oscilação Sul (OS) – mais importante fenômeno atmosférico de grande escala relacionado à variabilidade interanual nos trópicos, conforme Kayano e Kousky (1994) – que recebe a denominação de ENOS (*El Niño/Oscilação Sul*), determina perturbações no padrão de variabilidade da temperatura do ar e, principalmente, da precipitação, em diversas regiões do Globo, ocasionando enchentes em algumas regiões, e chuvas abaixo da média, e até mesmo grandes secas, em outras. Segundo Alves e Repelli (1992), tal resposta deve-se às características específicas de cada região, principalmente a sua posição geográfica, seu quadrimestre mais chuvoso e os principais sistemas atmosféricos responsáveis pelas chuvas sobre as mesmas. No Brasil, o fenômeno apresenta conseqüências opostas, dada a extensão do território, determinando excesso de precipitação no sul do país, e a ocorrência de precipitação abaixo da média no Nordeste brasileiro (Fontana e Berlato, 1997). Kane (1997) mostrou que, para o Nordeste do país, o *El Niño* é associado a cerca de 50% dos eventos de seca analisados ocorridos na região, apesar da precipitação no NEB estar mais fortemente relacionada a anomalias da TSM do Oceano Atlântico, cujo sistema acoplado oceano-atmosfera tem se mostrado o principal indutor de chuvas neste local. De acordo com Ward e Folland (1991), os mais

importantes preditores para uma previsão sazonal no NEB são obtidos no Oceano Atlântico.

Recentemente, a capacidade de se prever anomalias da TSM e o entendimento sobre as conseqüências destas a nível global tiveram uma grande evolução, e isto tem aumentado o interesse de centros meteorológicos em previsões sazonais. Duas outras razões também reforçam esse interesse (ECMWF, 2000): o desenvolvimento de um sistema de observação oceânico relativamente vasto no Pacífico Equatorial, de onde se obtém temperaturas medidas da superfície a cerca de 500 m de profundidade no Oceano, além de outros dados fornecidos por satélites europeus, americanos e japoneses, que podem possibilitar as previsões, como dados de vento, temperatura e umidade do ar. O segundo fato que estimula a produção de previsões sazonais é o aperfeiçoamento de modelos numéricos atmosféricos e oceânicos, que são modelos computacionais complexos que descrevem as condições climáticas da Terra, baseando-se em códigos numéricos de formas aproximadas das leis físicas que regem os movimentos na atmosfera ou nos oceanos, e as interações entre a atmosfera e a superfície.

Além da utilização dos modelos numéricos (também chamados “dinâmicos”) para previsões sazonais, outros métodos já bastante empregados para esse tipo de previsão são os estatísticos. Eles são baseados em fórmulas estatísticas que relacionam variáveis a serem previstas com variáveis do estado da atmosfera e dos oceanos. Folland *et al.* (1991) e Colman *et al.* (1998) fizeram uso de dois métodos estatísticos, regressão linear múltipla e análise discriminante linear, para prever a precipitação no Sahel africano e no Nordeste do Brasil (respectivamente), usando TSMs como preditores.

Os modelos dinâmicos mais largamente usados para as previsões de longo prazo são os modelos de circulação global atmosférica (MCGs). Esses MCGs apresentam uma malha horizontal com resolução de 2 a 4 graus de latitude-longitude, e são divididos verticalmente em 10 a 30 camadas. Como variáveis de saída, eles fornecem, entre outras, temperatura do ar, velocidade do vento em várias direções, cobertura de nuvens, precipitação, umidade do solo, cobertura de neve, radiação solar, pressão e temperatura na superfície do mar (Becker, 1997). No caso das regiões tropicais, as estações são melhor definidas como estações secas e úmidas, uma vez que a variação anual da temperatura do ar é bastante pequena nesta área do globo, o que resulta numa especial atenção para a precipitação como variável de principal interesse nas previsões, visto que ela define, em grande parte, o clima, a hidrologia e a biologia de uma região (Uvo, 1998).

Devido à baixa resolução horizontal da malha desses MCGs (equivalente a 2 a 5 centenas de quilômetros em regiões tropicais), variáveis relevantes às escalas local (até alguns milhares de m^2) ou da bacia hidrográfica (até cerca de 10 mil km^2) são mal representadas, ou mesmo ignoradas. Este é um dos principais problemas para se realizar a aplicação da previsão meteorológica em Recursos Hídricos: a diferença entre a escala para a qual são realizadas as previsões, e a escala da região onde se deseja empregá-las. A baixa resolução desses modelos deve-se ao poder computacional exigido para a execução dos mesmos, pois quanto maior a resolução espacial requerida, maior será o tempo de computação necessário.

Esses MCGs fornecem seus resultados em nível diário, apesar de geralmente não simularem bem a precipitação nessa escala temporal. Bons resultados da simulação desses modelos são normalmente associados apenas a previsões sazonais, visto que não se tem conhecimento de estudos que avaliem o desempenho de MCGs na escala mensal, o que os tornaria adequados para uso com modelos hidrológicos. O NEB é uma região cuja precipitação sazonal já está sendo prevista de modo experimental usando esses MCGs (Cavalcanti *et al.*, 1999; Evans e Graham, 1999).

Os resultados produzidos por um MCG podem ser considerados de natureza pontual ou espacial. Na hipótese pontual, supõe-se que o MCG simula dados referentes à localização de cada nó, encontrando respaldo no esquema numérico do modelo, que baseia-se em métodos que só produzem valores correspondentes a cada ponto da malha. Na abordagem espacial, os dados gerados nos nós da malha do modelo são tomados como médias de áreas cujos centros geométricos correspondem aos respectivos nós, encontrando suporte na concepção e parametrização do modelo, que usa como valores de entrada dados calculados como médias espaciais. Apesar de não chegarem a um consenso sobre qual é a natureza dos dados fornecidos pelos MCGs, Skelly e Henderson-Sellers (1996) sugerem que a abordagem espacial pode ser mais útil e intuitiva em aplicações hidrológicas.

Aliados aos métodos de previsão sazonal descritos anteriormente, prognósticos de consenso elaborados por meteorologistas das agências de previsão do Nordeste são de extrema valia e devem ser considerados. Esses prognósticos são baseados na análise conceitual subjetiva dos técnicos sobre padrões das principais variáveis atmosféricas e oceânicas em grande escala (Repelli e Alves, 1994), normalmente realizados no mês que antecede o início da estação chuvosa.

Um dos modelos que fornecem atualmente previsões meteorológicas sazonais para o NEB é o modelo de circulação global atmosférico do *European Centre for Medium-Range Weather Forecasts* – ECMWF, um centro de estudos climáticos da União Européia, localizado no Reino Unido (Becker, 1997). Os resultados da previsão são fornecidos em nós de uma malha de resolução de $2,5^{\circ} \times 2,5^{\circ}$ de latitude-longitude. Outros modelos, como o do Centro de Previsão de Tempo e Estudos Climáticos do Instituto Nacional de Pesquisas Espaciais – CPTEC/INPE (Cavalcanti *et al.*, 1999), e do *Hadley Centre for Climate Prediction and Research* do serviço meteorológico britânico (Evans e Graham, 1999), também já fazem este tipo de previsão para o Nordeste do Brasil.

2.3. DESAGREGAÇÃO ESPACIAL DE PRECIPITAÇÃO

Como ferramentas para produção de previsões sazonais, os MCGs conseguem representar bem os padrões climáticos de grande escala. Todavia, devido à baixa resolução do modelo, características locais (acidentes topográficos, linhas de costa, uso do solo,...) que influenciam o clima numa determinada região podem ser ignoradas, conduzindo a resultados irrealistas se aplicadas diretamente em pequena escala. Para solucionar o problema da diferença entre a escala para a qual são realizadas as previsões, e a esc

ala da região onde se deseja empregá-las, faz-se uso de técnicas de transferência da informação meteorológica para as escalas menores, as chamadas técnicas de *downscaling* ou desagregação espacial (Wilby *et al.*, 1998). Além da importância destas para as previsões meteorológicas de longo prazo, essas técnicas são também úteis nas previsões de tempo fornecidas por modelos dinâmicos semelhantes aos MCGs, que embora possuam uma maior resolução, ainda precisam ter suas informações desagregadas para as escalas menores, e ainda na desagregação de cenários decorrentes de possíveis alterações climáticas (como o aumento do efeito estufa com a emissão de gases e aerossóis), também simulados por MCGs. A maioria dos modelos que avaliam impactos produzidos por mudanças climáticas operam numa escala espacial de 1 a 100 km de resolução (mesoescala). Portanto, as informações sobre possíveis alterações climáticas têm de ser transferidas para a mesma escala de resolução, a fim de que possam ser usadas como dados de entrada por esses modelos (Semenov e Barrow, 1997; Schubert, 1998).

Duas formas de desagregação espacial podem ser usadas para solucionar o problema da diferença entre escalas (Hewitson e Crane, 1996): (i) a utilização de métodos empíricos, que desagregam estatisticamente a previsão meteorológica fornecida pelo MCG para a escala da bacia; (ii) e através do aninhamento de modelos numéricos semelhantes aos MCGs, porém com resolução mais fina, utilizando condições iniciais e de contorno advindos de um MCG.

A desagregação de variáveis meteorológicas através de métodos estatísticos parte da suposição de que relações empíricas estáveis podem ser estabelecidas entre processos atmosféricos que ocorrem em escalas espaciais e/ou temporais diferentes (Wilby *et al.*, 1998). Trata-se de uma abordagem que oferece soluções em curto espaço de tempo, e que não necessita de grande poder computacional, constituindo uma solução viável para os pesquisadores da área. A principal limitação desta abordagem é que o modelo de desagregação fica restrito a uma região e às séries de dados utilizadas.

O segundo método de desagregação de saídas de um MCG para uma resolução espacial maior, consiste da utilização de aninhamento de modelos numéricos. São modelos climáticos de mesoescala ou regionais que, inicializados com os resultados de um MCG, descrevem melhor as condições climáticas de uma região. Esses modelos podem apresentar ainda a peculiaridade de possuir várias grades de espaçamentos diferentes, que podem ser aninhadas, proporcionando resultados de alta resolução (Pielke *et al.*, 1992). Essa metodologia, no entanto, necessita de tempo e grande capacidade computacional para a execução dos modelos, e ainda possui o agravante de que as simulações dos modelos de maior resolução podem ser contaminadas por erros nas condições limites simuladas pela baixa resolução do MCG (Frederick *et al.*, 1997).

Entre os métodos empíricos utilizados como ferramentas de desagregação, os mais largamente empregados até hoje são os de regressão múltipla (Palutikof *et al.*, 1997; Schubert, 1998). Nos últimos anos, no entanto, o método das redes neurais artificiais (RNAs) tem conquistado espaço entre pesquisadores da área, obtendo bons resultados como um instrumento de desagregação espacial (Cavazos, 1997; Trigo e Palutikof, 1999; Galvão e Trigo, 1999). Conforme Gardner e Dorling (1998), uma grande vantagem desse método sobre as demais técnicas estatísticas é o fato dele não precisar de suposições iniciais sobre a distribuição estatística dos dados. A sua aplicação tem provado ser particularmente efetiva na desagregação de precipitação, onde há relações não-lineares que

técnicas mais tradicionais, como regressão múltipla, não capturam muito bem (Hewitson e Crane, 1996).

Uma consideração também importante a ser feita na desagregação espacial de dados de precipitação é a escala temporal da mesma, visto que os MCGs fornecem seus resultados em nível diário, mas, no entanto, não simulam bem a precipitação nesta escala temporal. Desejando reduzir o erro associado a essas previsões, antes de serem utilizados, os resultados são normalmente agrupados em escalas temporais maiores, geralmente em formato sazonal, e, por conseguinte, a desagregação também tem sido efetuada neste formato.

2.4. REDES NEURAIS ARTIFICIAIS

Segundo Valença e Ludermir (1999a), as Redes Neurais são sistemas constituídos por elementos de processamento interconectados, chamados de neurônios, os quais estão dispostos em camadas (uma camada de entrada, uma ou várias intermediárias e uma de saída) e são responsáveis pela não-linearidade da rede, através do processamento interno de certas funções matemáticas.

A rede mais simples existente é conhecida por *Perceptron*, e foi proposta por Rosenblatt (1958) *apud* Hines (1997). Ela é composta de duas camadas, onde os neurônios da primeira estão conectados ao único neurônio da camada de saída.

O tipo de rede neural que tem estado em maior evidência na comunidade científica é a rede *perceptron* multicamadas. Poderosa técnica de regressão (Sarle, 1994), ela é capaz de modelar funções altamente não-lineares, e após ser treinada, generalizar com grande precisão. Trata-se de uma rede com uma ou mais camadas intermediárias, em que todos os neurônios de uma camada se ligam a todos da camada seguinte (Figura 2.4). Entende-se por generalização a capacidade que uma RNA tem de fornecer respostas adequadas para dados de entrada que não foram usados na calibração.

Numa rede multicamadas, o valor de saída de cada neurônio é multiplicado por um peso (w), cujo produto será a entrada para o neurônio da camada seguinte. Em cada neurônio são somados todos os produtos recebidos, sendo então adicionado ao somatório uma parcela conhecida por ruído (b). O resultado disto é, em seguida, transformado por

uma função de ativação que transmitirá o resultado da transformação aos neurônios da camada seguinte.

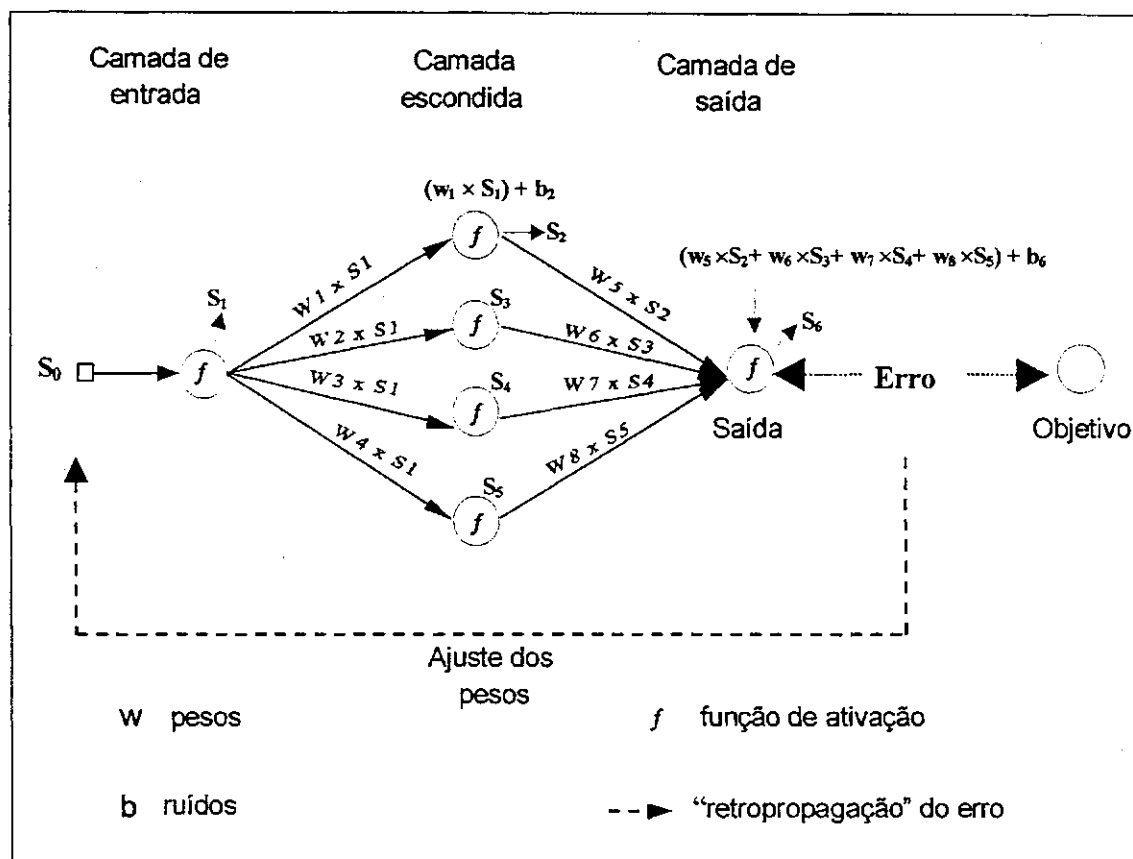


Figura 2.4 – Exemplo da estrutura de uma RNA multicamada com três camadas

O elemento responsável pela não-linearidade de uma rede neural é a função de ativação. Ela transforma as informações que chegam num dado neurônio, para serem utilizadas pelos neurônios da camada seguinte. Há vários tipos de função que podem ser empregadas numa rede. As mais comuns são a linear, a sigmóide, a gaussiana e a tangente hiperbólica. Segundo Azoff (1994), pode-se utilizar uma função de ativação diferente para cada neurônio existente numa rede.

O aprendizado de uma rede, processo que habilita a mesma à realização de novas tarefas, pode ser realizado de três formas diferentes (Valença e Ludermir, 1999a): através do estabelecimento de novas conexões na rede, da eliminação de conexões existentes, e da alteração dos pesos das conexões. Existem duas categorias básicas de aprendizado: o supervisionado e o não supervisionado. O primeiro consiste da necessidade de um conjunto de dados de entrada e suas respectivas respostas para que, de posse da diferença entre a

resposta desejada e a fornecida pela rede, se possa corrigir os pesos das conexões entre os neurônios. No aprendizado não supervisionado, a rede requer o conjunto de dados de entrada sem exigir o conjunto resposta. Ela aprende a detectar regularidades e correlações em suas entradas, moldando suas futuras respostas àquelas entradas. São as chamadas redes auto-organizáveis.

A rede multicamada necessita de aprendizado supervisionado, realizado através de um algoritmo de treinamento que busca o mínimo global através de incrementos nos valores dos pesos por uma quantidade proporcional à primeira derivada da diferença entre o valor de saída desejado e o fornecido pela rede, modificando os pesos da rede até reduzir tal diferença.

Atualmente, existem vários tipos de redes neurais. O que diferencia uma da outra são as funções de ativação utilizadas, se a rede possui camadas intermediárias ou não, como são efetuadas as conexões entre os neurônios, e o tipo de aprendizado empregado na rede.

As redes recorrentes são extensões das redes *Perceptron* multicamadas, com a particularidade de apresentar conexões de retorno entre camadas intermediárias ou de saída e a camada de entrada. As redes de Jordan, Elman (Figura 2.5) e Hopfield são exemplos de redes recorrentes (Beale e Jackson, 1990; Valença e Ludermir, 1999a).

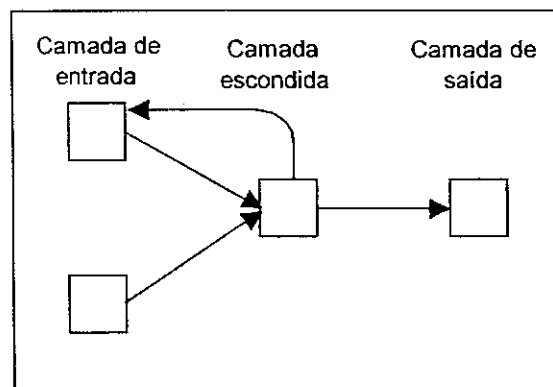


Figura 2.5 – Rede de Elman

As redes com função de base (RBF), conforme Valença e Ludermir (1999a), também podem ser consideradas como redes multicamadas que contém apenas uma camada intermediária, onde os neurônios dessa usam funções de base ortogonal, e os da camada de saída empregam a função linear. Essas redes apresentam rápida aprendizagem e

facilidade de projeto, porém, em alguns casos, podem necessitar de um número muito grande de neurônios na camada intermediária, tornando-se preferível a utilização de outras redes.

As redes auto-organizáveis são aquelas que estão submetidas ao aprendizado não supervisionado. Dentre estas, merecem destaque a de Kohonen (Caudill, 1997) e as redes compostas por blocos de sigmóides auto-organizáveis (SSBN) (Valença e Ludermir, 1999a).

Nos últimos anos tem havido um aumento na adoção das RNAs pela comunidade meteorológica. Elas têm sido utilizadas como técnicas de desagregação espacial, estabelecendo funções de transferência entre a circulação de grande escala e a precipitação num dado local (Hewitson e Crane, 1996; Cavazos, 1997). Isto se deve, principalmente, à capacidade que uma RNA tem de generalizar relações, após ser “treinada” (calibrada) usando casos observados.

Embora sejam análogas à regressão múltipla, o maior atrativo das RNAs é sua capacidade de representar qualquer função não-linear. Robinson (1991) promoveu em seu estudo uma comparação entre regressão múltipla e redes neurais, e apesar desta última ter se mostrado significativamente melhor do que a primeira, ele chama atenção para alguns pontos importantes para se obter bons resultados: o primeiro diz respeito à necessidade de uma quantidade satisfatória de dados (quantidade essa que depende de cada caso específico), para que a rede seja capaz de identificar a relação existente no problema. O segundo, refere-se ao tamanho mais adequado para a rede, que deverá conter mais ou menos neurônios e camadas de acordo com a complexidade do problema. A terceira e última observação feita é relacionada à utilização de valores padronizados para os parâmetros de aprendizagem da rede, que podem não ser os mais apropriados para algumas situações, tornando-se a escolha destes uma questão de experiência e/ou tentativas.

3. METODOLOGIA

3.1. DELIMITAÇÃO DA REGIÃO DE ESTUDO

A região escolhida para a aplicação do método de desagregação foi o norte do Nordeste do Brasil (norte-NEB), que compreende parte dos Estados do Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas e norte da Bahia. Essa escolha foi baseada nas seguintes considerações (Galvão, 1999):

- a região apresenta, no NEB, a maior variabilidade interanual de precipitação e maior escassez hídrica;
- tem sua economia com base no grande número de reservatórios construídos e na pequena agricultura familiar dependente de chuva;
- tem sido mais estudada climática e hidrológicamente;
- sua precipitação sazonal tem apresentado maior previsibilidade que a das outras sub-regiões do Nordeste.

Para representar a região escolhida neste estudo, selecionou-se a bacia hidrográfica do Rio Piancó, limitada pelo posto fluviométrico de mesmo nome. Ela possui uma área de 4.550 km² e está localizada na porção sudoeste do Estado da Paraíba, entre as latitudes 7°10'S e 7°53'S e longitudes 37°50'W e 38°44'W. Três fatos justificam essa escolha: a bacia do Rio Piancó apresenta características hidrometeorológicas típicas do norte-NEB; ela possui uma boa densidade espacial de postos pluviométricos, que fornecem séries históricas de longa duração e boa qualidade; e ela pode ser considerada de porte médio, representando bem as bacias contribuintes de reservatórios do Nordeste do Brasil, que possuem áreas de 200 km² a 15.000 km².

Neste trabalho, considerou-se a malha de resolução do modelo do ECMWF como referência (Figura 3.1), que gera células em regiões tropicais – segundo a abordagem espacial sugerida por Skelly e Henderson-Sellers (1996), e aqui adotada – de 275 km × 275 km. Entre estas, apenas uma célula contém toda a Bacia do Rio Piancó, se estendendo de

6,25°S a 8,75°S e de 36,25°W a 38,75°W, e onde estão incluídas regiões dos Estados do Ceará, Rio Grande do Norte, Paraíba e Pernambuco (Figura 3.2).

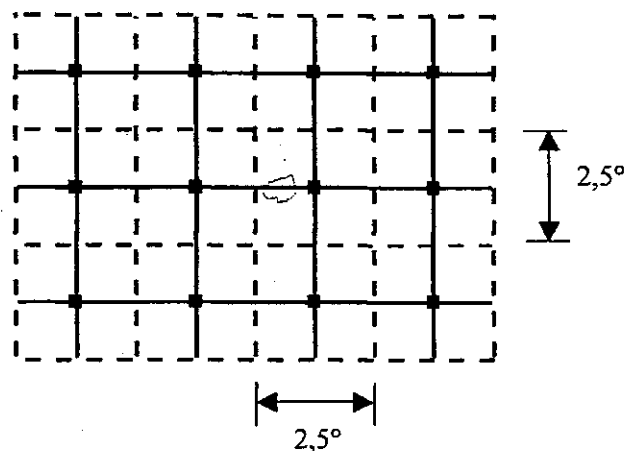
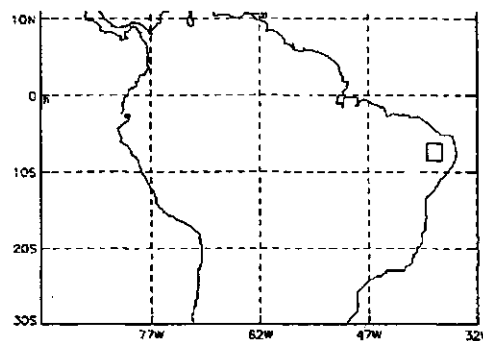


Figura 3.1 - Malha do modelo do ECMWF



localização da célula do MCG

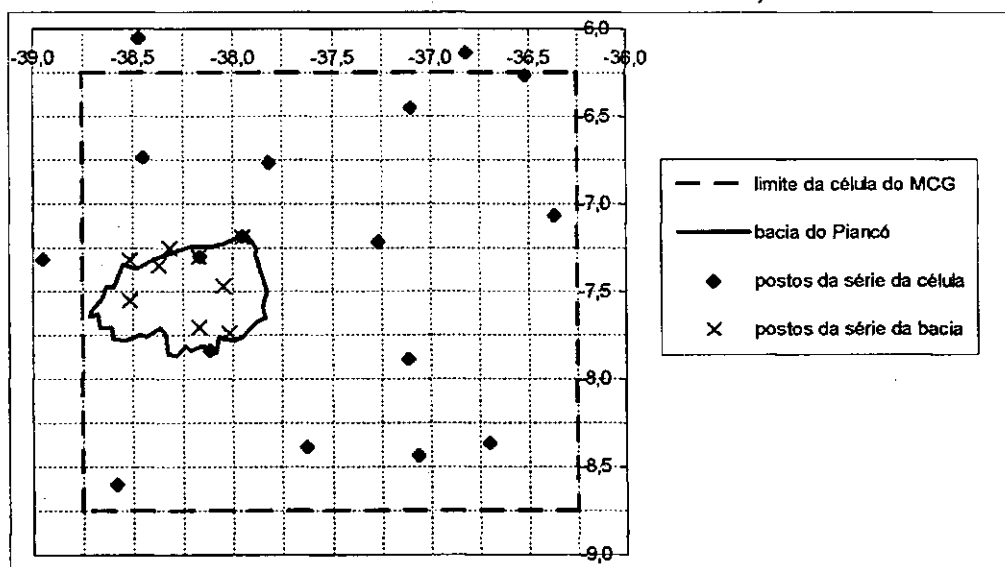


Figura 3.2 - Mapas esquemáticos (fora de escala) mostrando a célula do MCG sobre a Bacia do Rio Piancó, e os respectivos postos pluviométricos

A opção pelo modelo do ECMWF como referência se deve ao fato de que Galvão (1999), buscando verificar a previsibilidade do MCG, comparou valores previstos pelo modelo – executado nove vezes para produzir o conjunto de previsões para cada estação do ano, originando nove membros que diferem entre si apenas pelo dia em que o modelo foi inicializado – com dados observados de precipitação obtidos em postos pluviométricos da região da célula. Nesta comparação (Figura 3.3), ele observou que as previsões seguiram bem a tendência de anos secos e chuvosos e, apesar do modelo quase sempre subestimar a precipitação, obteve altos coeficientes de correlação entre a precipitação observada e a prevista por cada um dos membros ou a sua média (0,90), o que comprovou a boa previsibilidade do MCG do ECMWF no nível sazonal nesta região.

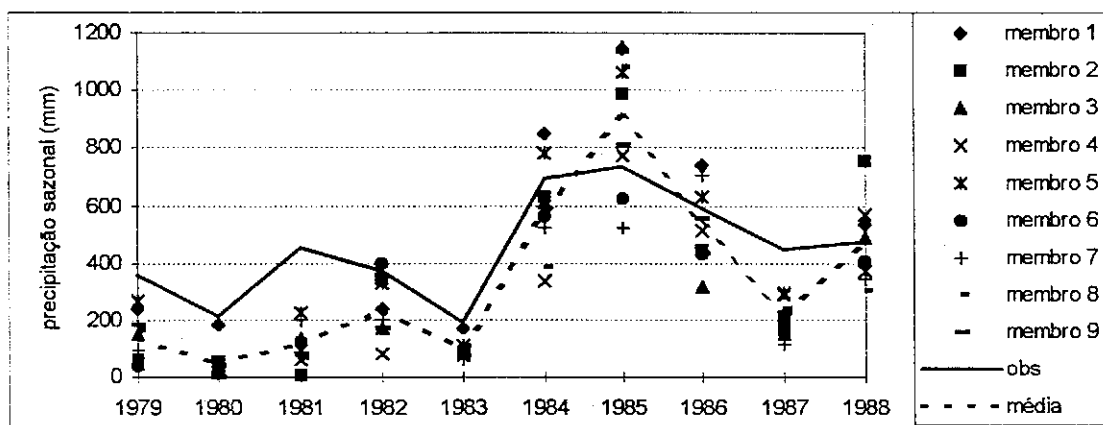


Figura 3.3 – Comparação entre o total precipitado de março-junho observado e previsto pelo MCG do ECMWF para a célula sobre a bacia do Piancó. A linha ‘média’ corresponde à média dos nove membros e a linha ‘obs’ corresponde à média de precipitação de 17 postos localizados na célula (Galvão, 1999)

Como já mencionado no Capítulo 2, os MCGs apresentam malhas horizontais com baixa resolução e, por isso, para se utilizar suas previsões em escalas menores, faz-se necessário o uso de técnicas que proporcionem a desagregação espacial. A Figura 3.4 ilustra dois níveis de redução de escala possíveis para a região escolhida neste estudo. Os valores obtidos no nó da malha do MCG poderiam ser desagregados para a escala de bacia, construindo relações de desagregação com valor oriundo da média de dados de postos pluviométricos localizados na mesma, ou ao nível de posto pluviométrico, estabelecendo relações diretamente com cada posto. Neste estudo, abordou-se apenas a desagregação para a escala de bacia, visto que o tempo imposto para o desenvolvimento da pesquisa não possibilitaria o estudo para a outra escala e a adequada análise de suas implicações.

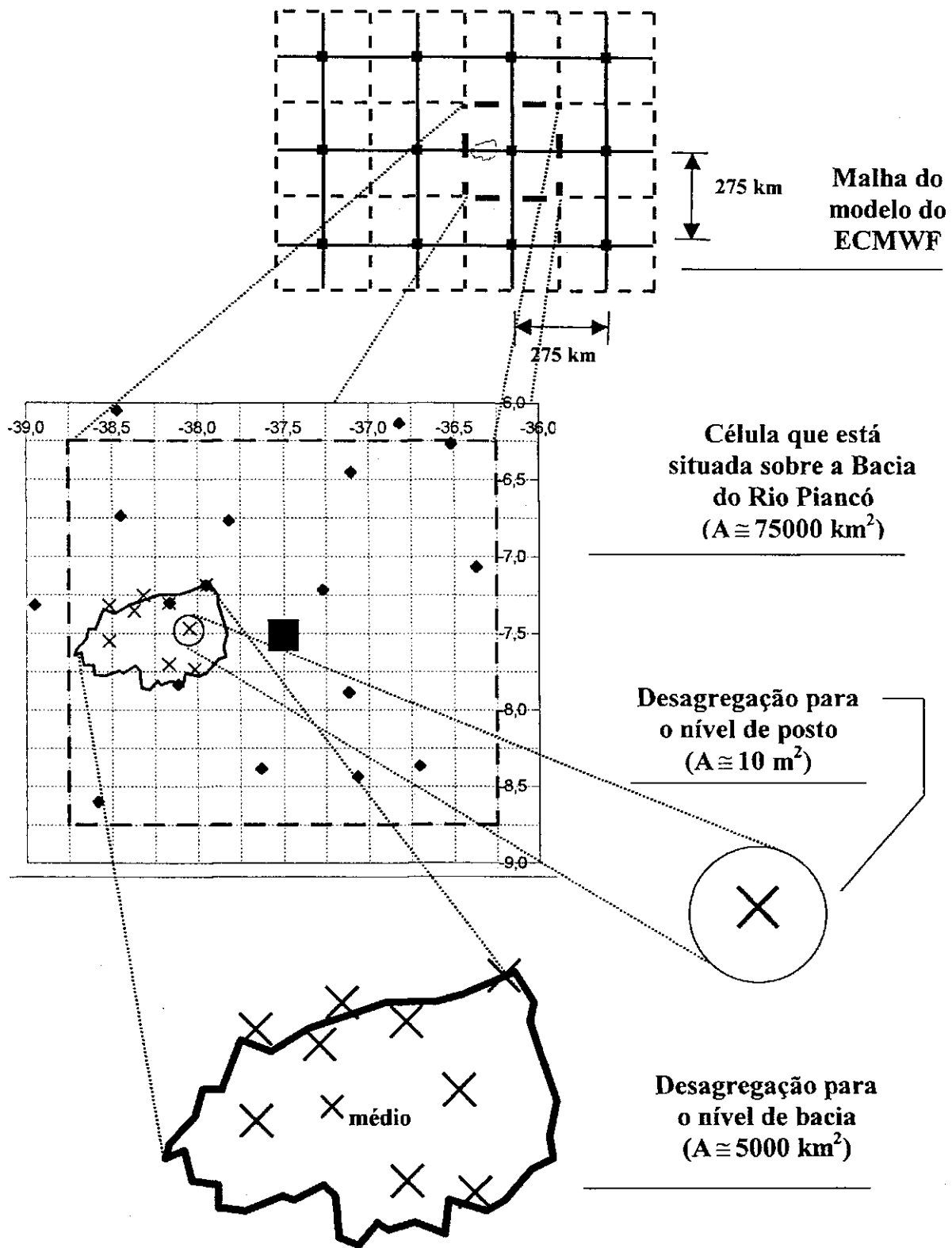


Figura 3.4 – Possibilidades de redução de escala para a região de estudo

É de suma importância a compreensão de que, neste estudo, não se utilizou o MCG ou qualquer resultado dele. Fez-se uso apenas de dados observados, com o propósito de estabelecer relações que possam desagregar, posteriormente, as previsões do MCG para a escala da bacia hidrográfica. Considerou-se aqui, apenas como referência para a desagregação, a malha do modelo do ECMWF.

3.2. COLETA DE DADOS PLUVIOMÉTRICOS

Os dados de precipitação que foram utilizados para a calibração e validação do método de desagregação foram obtidos de Galvão (1999), e originalmente extraídos dos bancos de dados da Superintendência do Desenvolvimento do Nordeste (SUDENE), do Departamento Nacional de Águas e Energia Elétrica (DNAEE), da Universidade Federal da Paraíba (UFPB), e de núcleos estaduais de meteorologia e recursos hídricos da região.

Os postos utilizados para representar a precipitação na bacia (Tabela 3.1 e Figura 3.2) foram selecionados com base na sua distribuição espacial, na extensão da série e na qualidade dos dados, conforme avaliação realizada por Molinier *et al.* (1994). Dos nove postos selecionados, quatro foram implantados em 1911 e cinco em 1933. Como poucos destes apresentavam dados disponíveis em anos posteriores a 1988, provavelmente devido a uma crise institucional vivida pela SUDENE, principal responsável na época pela coleta dos dados, optou-se por se considerar apenas o período de 1911 a 1988 nesta dissertação.

Tabela 3.1 – Postos pluviométricos da bacia do Rio Piancó

CÓDIGO DNAEE	POSTO	COORDENADAS / ALTITUDE	INÍCIO DA SÉRIE
00738013	Princesa Isabel	7°44'S - 38°01'W / 660 m	1911
00737006	Piancó	7°11'S - 37°57'W / 250 m	1911
00738017	Itaporanga	7°18'S - 38°10'W / 230 m	1911
00738020	Conceição	7°33'S - 38°31'W / 370 m	1911
00738023	Bom Jesus	7°21'S - 38°22'W / 470 m	1933
00738022	Bonito de Santa Fé	7°19'S - 38°31'W / 575 m	1933
00738015	Manáira	7°42'S - 38°10'W / 605 m	1933
00738010	Serra Grande	7°15'S - 38°19'W / 585 m	1933
00738014	Nova Olinda	7°28'S - 38°03'W / 315 m	1933

Galvão (1999) utilizou o método de Thiessen para calcular as precipitações médias diárias sobre a bacia do Rio Piancó. As falhas não foram preenchidas diretamente, mas levadas em conta à medida que não foram consideradas na determinação dos polígonos os postos que as apresentavam.

Os dados de precipitação obtidos para representar o valor gerado pelo MCG na célula que contém a bacia do Rio Piancó foram retirados da base de dados de Brito *et al.* (1991). Considerou-se dezessete postos com séries de longa duração, agrupadas mensalmente, dois dos quais coincidentes com os da bacia (Tabela 3.2 e Figura 3.2).

Tabela 3.2 – Postos pluviométricos da célula do MCG sobre a bacia do Rio Piancó

CÓDIGO DAAEE	POSTO	COORDENADAS / ALTITUDE	PERÍODO DA SÉRIE	LOCALIZAÇÃO
00636005	Florânia (Flores)	6°08'S - 36°49'W / 210 m	1911-1989	RN
00636025	Currais Novos	6°16'S - 36°31'W / 350 m	1911-1988	RN
00637004	Caicó	6°27'S - 37°06'W / 143 m	1913-1989	RN
00637032	Pombal	6°46'S - 37°49'W / 178 m	1911-1988	PB
00638011	Pereiro	6°03'S - 38°28'W / 220 m	1911-1989	CE
00638033	Antenor Navarro	6°44'S - 38°27'W / 240 m	1911-1990	PB
00736008	Soledade(Ibiapinópolis)	7°04'S - 36°22'W / 560 m	1913-1988	PB
00737002	Teixeira	7°13'S - 37°16'W / 770 m	1911-1989	PB
00737006	Piancó	7°11'S - 37°57'W / 250 m	1911-1985	PB
00737014	Monteiro	7°53'S - 37°07'W / 590 m	1912-1986	PB
00738008	Milagres	7°19'S - 38°57'W / 371 m	1912-1989	CE
00738017	Itaporanga(Misericórdia)	7°18'S - 38°10'W / 230 m	1911-1989	PB
00738032	Triunfo	7°50'S - 38°07'W / 1010 m	1912-1989	PE
00836032	Pesqueira	8°22'S - 36°42'W / 650 m	1920-1989	PE
00837003	Arcoverde (Rio Branco)	8°26'S - 37°04'W / 663 m	1914-1989	PE
00837025	Ibimirim (Jeritacó)	8°23'S - 37°38'W / 445 m	1934-1989	PE
00838020	Floresta	8°36'S - 38°35'W / 317 m	1911-1988	PE

Como Galvão (1999) mostrou que no nível sazonal, para a região da bacia do Piancó, a desagregação é viável, definiu-se esta escala temporal para o desenvolvimento do presente estudo. Acreditando-se que os MCGs também simulem razoavelmente bem a precipitação mensal, e que isto será em breve confirmado em outros estudos, optou-se por produzir relações de desagregação também na escala mensal. Portanto, objetivando-se construir essas relações, as médias diárias referentes à bacia foram acumuladas mensalmente, e depois agrupadas a nível sazonal (fevereiro a maio), procedimento também adotado para as médias mensais que representam a precipitação da célula.

3.3. DIAGNÓSTICO DAS RELAÇÕES DA PRECIPITAÇÃO ENTRE AS DUAS ESCALAS

Buscando-se a visualização das relações existentes entre a precipitação da célula e da bacia nas escalas sazonal e mensal, comparou-se as séries de dados das duas escalas espaciais, resultantes das médias entre dados dos postos pluviométricos selecionados, construindo-se gráficos que mostram o comportamento da precipitação em cada escala temporal.

Para a verificação da estabilidade de relações resultantes entre as séries sazonais da célula e da bacia, calculou-se coeficientes de correlação entre as duas séries para uma “janela” de 30 anos, móvel ano a ano sobre todo o período. O mesmo procedimento foi adotado com relação às séries mensais, calculando-se “janelas” móveis de 30 anos aplicadas para os meses de fevereiro a maio.

Galvão (1999) mostrou, através do traçado de isoietas na célula sobre a bacia do Piancó, que o padrão de distribuição espacial da precipitação sazonal, seja para anos secos, normais ou chuvosos, sempre segue aproximadamente a topografia da região, caracterizando uma distribuição espacial relativamente bem definida, o que favorece a desagregação espacial da precipitação na região da bacia do Piancó.

3.4. SELEÇÃO DO MÉTODO DE DESAGREGAÇÃO

Como já mencionado na seção 2.3, nos últimos anos, o método das Redes Neurais Artificiais (RNAs) tem obtido bons resultados como um instrumento de desagregação espacial, e tem provado ser particularmente efetivo na desagregação de precipitação, onde há uma relação não-linear que técnicas convencionais não capturam muito bem. Em comparação aos modelos dinâmicos com grades aninhadas, apresenta ainda a grande vantagem de não exigir alto custo computacional para a sua execução. Por tais motivos, o método das redes neurais artificiais foi julgado apropriado para o desenvolvimento desta dissertação. Um fato que possibilita a utilização deste método empírico é a existência de séries de dados de longa duração disponíveis para a modelagem.

Com base no estudo de Galvão (1999), no qual verificou que as séries sazonais da célula e da bacia do Piancó apresentam uma relação aproximadamente linear, recorreu-se ao método da Regressão Linear Simples para também se realizar a desagregação espacial da precipitação. O objetivo disto é de comparar e avaliar, posteriormente, a eficiência do método das Redes Neurais Artificiais.

3.5. CALIBRAÇÃO E VALIDAÇÃO DOS MODELOS

3.5.1 REGRESSÃO LINEAR SIMPLES

A regressão linear simples é um método estatístico que busca estabelecer uma equação matemática linear que descreva o relacionamento entre duas variáveis. Essa equação é então utilizada para estimar, ou prever valores futuros de uma variável quando se conhece os valores da outra.

Denominada de equação de regressão, ela possui a forma $y = ax + b$, onde x é a variável independente, a e b são os parâmetros a serem estimados, e y é a variável dependente. Essa equação é estimada através do método dos mínimos quadrados, técnica que define que a reta a ser estabelecida entre pontos que representam a relação entre as duas variáveis (Figura 3.5) deverá ser aquela que torna mínima a soma dos quadrados das distâncias da reta aos pontos, medidas na direção da variável dependente (y).

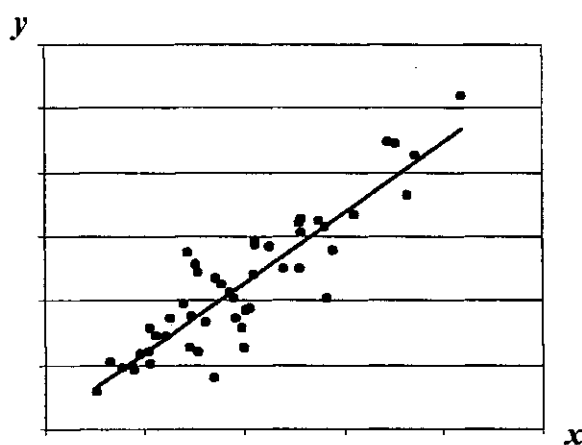


Figura 3.5 – Reta de ajuste da Regressão Linear obtida entre duas variáveis

Neste estudo, buscou-se encontrar equações de regressão que possibilitassem a desagregação espacial da precipitação, onde X representaria a precipitação prevista pelo MCG na escala da célula, e Y a precipitação correspondente na bacia do Piancó. Para a modelagem dessa desagregação, adotou-se o procedimento de se dividir cada série de dados em duas outras: uma para calibrar o modelo e a outra para validá-lo. Na fase de calibração estimou-se os parâmetros a e b através do já citado método dos mínimos quadrados, originando as equações de regressão. Na fase de validação, essas equações foram verificadas com a utilização do segundo grupo de dados, onde se compara os valores calculados pela equação de regressão com os dados observados da bacia do Piancó.

Para se verificar o grau de relacionamento entre as duas variáveis de uma equação de regressão, são feitas análises dos coeficientes de correlação calculados em cada fase da modelagem. Esse foi o procedimento adotado neste estudo.

Para a modelagem do método de regressão linear foram utilizadas planilhas eletrônicas do EXCEL (Microsoft, 1997).

3.5.2 REDES NEURAS ARTIFICIAIS

Para a implementação de uma rede neural artificial (RNA), algumas decisões devem ser tomadas, tais como a seleção dos dados – se de “treinamento” (expressão comumente utilizada para designar a calibração de uma rede neural) ou de validação, a arquitetura da rede, o algoritmo de treinamento e o número de iterações necessárias (Hines, 1997).

O fluxograma apresentado na Figura 3.6 indica os passos que devem ser tomados para se implementar uma rede neural com aprendizado supervisionado. O primeiro é a coleta dos dados. Uma vez coletados, eles são divididos em dois grupos: um para treinar e o outro para validar a rede. Essa seleção deve ser realizada de modo que o primeiro grupo contenha exemplos dos possíveis casos que possam ocorrer, para que a rede seja capaz de capturar a relação entrada-saída, fornecendo respostas adequadas para qualquer situação.

Após selecionado o conjunto de dados para o treinamento, deve-se definir a arquitetura da rede. Esta definição consiste na descrição de quantas camadas essa rede deve

ter, o número de neurônios existente em cada uma, a função de ativação escolhida para cada camada, e como elas são conectadas umas as outras e com as entradas da rede. A melhor arquitetura para a rede depende do tipo de problema a ser representado. Em geral, ela deve possuir a estrutura mais simples possível que consiga estabelecer a relação entrada-saída desejada.

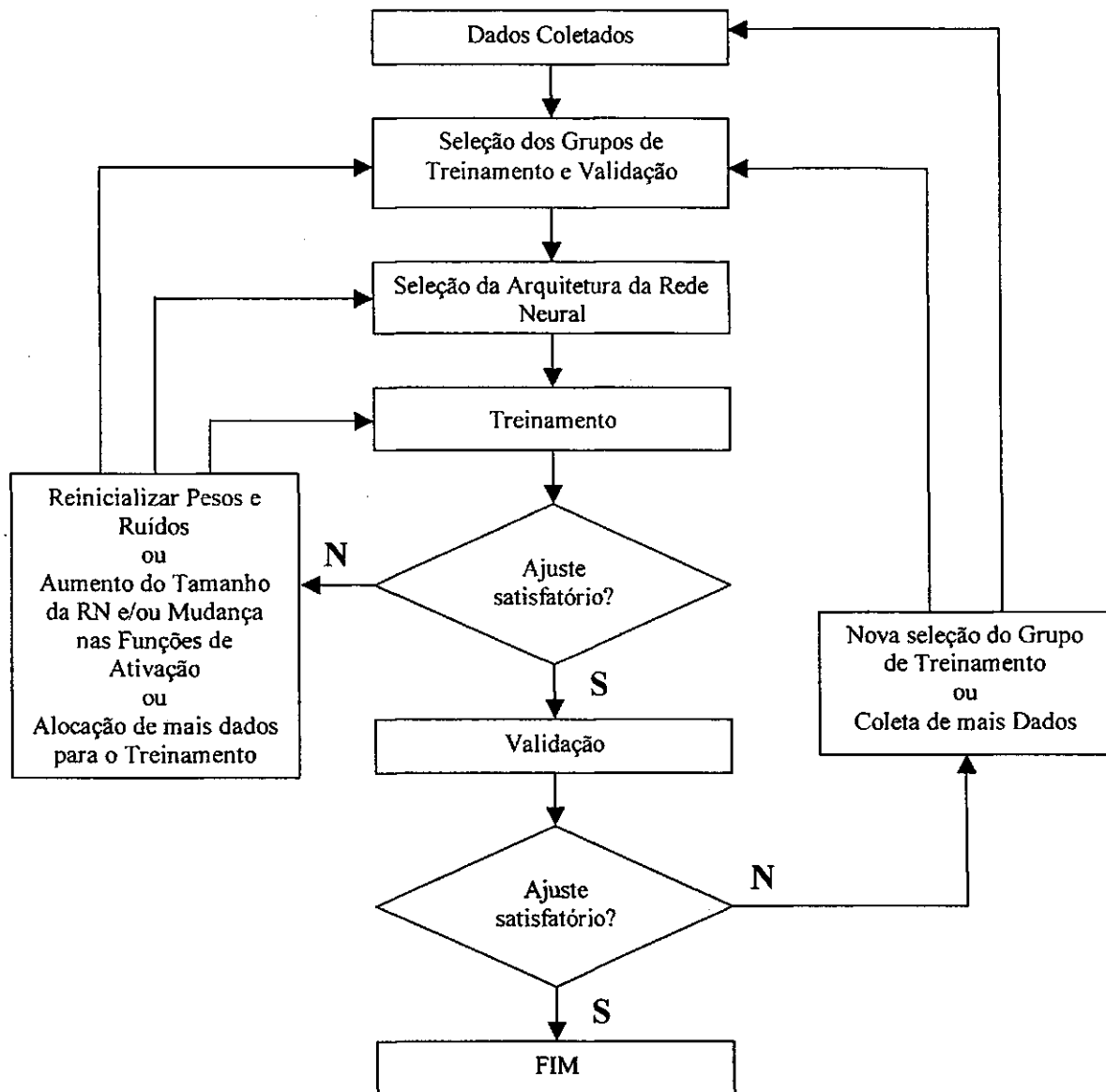


Figura 3.6 – Fluxograma do processo de treinamento e validação de uma RNA com aprendizado supervisionado (Hines, 1997)

Definida a arquitetura da rede (também chamada de “estrutura”), dá-se início à fase de treinamento da mesma, buscando-se um ajuste satisfatório entre dados observados e calculados pela rede.

Segundo Hines (1997), a rede pode não alcançar o resultado desejado devido às seguintes razões:

- o algoritmo de treinamento converge para um mínimo local;
- a rede não tem o número de graus de liberdade necessário para achar a relação entrada-saída;
- não há dados suficientes para executar o traçado desejado.

No primeiro caso, a solução seria reinicializar pesos e ruídos, e recomeçar o treinamento. No segundo, seria adicionar neurônios e/ou camadas, recomeçando o treinamento em seguida. O último caso necessitaria da alocação de mais dados para o conjunto de treinamento. Uma das situações que devem ser evitadas durante o treinamento de uma rede é a ocorrência de um superajustamento (*overfitting*). Isto acontece quando ela contém um número elevado de neurônios que faz com que a rede apenas “memorize” os dados, perdendo sua boa capacidade de generalização.

Após se alcançar um ajuste satisfatório no treinamento, o próximo passo seria a validação da rede, onde se emprega o conjunto de dados até então não utilizado. Esse conjunto deve ser uma amostra representativa do problema e, de acordo com Hines (1997), se um resultado satisfatório não for alcançado será devido ou a uma fraca generalização do conjunto de treinamento, que deverá ser resolvida com uma nova seleção do grupo, ou a uma má representatividade do conjunto de validação, indicando que este necessita de outros dados.

Atingindo-se resultados satisfatórios também na validação, a RNA estará pronta para ser utilizada, fornecendo respostas adequadas a dados de entrada que ainda não tenham sido usados, demonstrando a sua capacidade de generalização.

O tipo de rede neural adotado para a desagregação dos dados de precipitação sobre a bacia do Rio Piancó foi a rede *perceptron* multicamadas. Essa escolha teve por base a capacidade que este tipo de rede tem, já mencionada na seção 2.4, de poder generalizar com precisão.

O aprendizado é do tipo supervisionado, onde foi fornecido como entrada à rede, um conjunto de dados observados de precipitação em grande escala (representando a célula

do MCG), e como a respectiva resposta esperada, um conjunto de dados de precipitação observados na bacia hidrográfica do Rio Piancó. A diferença entre a resposta desejada e a fornecida pela rede é reduzida através da utilização de um algoritmo de treinamento, técnica de otimização projetada para minimizar uma função objetivo, modificando os pesos da rede até atingir tal fim. Com esse algoritmo, cada exemplo treinado gera um erro de saída que é usado para ajustar os pesos e ruídos, os quais são iniciados com valores aleatórios.

O algoritmo de treinamento utilizado na modelagem das redes neurais foi o algoritmo de Levenberg–Marquardt, um algoritmo de retropropagação (*backpropagation*), que ajusta pesos e ruídos modificando-os da camada de saída à camada de entrada. Segundo Hagan e Menhaj (1994), ele é muito mais eficiente que outras técnicas quando treina redes que têm até umas poucas centenas de parâmetros (pesos e ruídos), provocando uma convergência mais rápida e resultados com maior precisão. Em Demuth e Beale (1998), a comparação do desempenho de vários algoritmos de treinamento em um determinado problema mostrou que o algoritmo de Levenberg–Marquardt é amplamente superior aos demais, pois atinge a convergência em menor tempo, com um menor número de iterações, e com menor número de operações com ponto flutuante. O responsável pela velocidade e pela precisão com que uma rede converge, atingindo rapidamente um valor próximo ao mínimo global da função objetivo erro, é um parâmetro que o algoritmo de Levenberg–Marquardt apresenta. Esse parâmetro é variável no tempo, diminuindo após cada passo do algoritmo que reduza a função objetivo, e sofrendo acréscimos quando o algoritmo se afasta do mínimo da função, resultando na redução do erro a cada iteração do algoritmo de treinamento. A grande desvantagem atribuída a esse algoritmo é a capacidade computacional exigida por ele, que é bem superior à capacidade necessária à utilização dos demais algoritmos existentes. Todavia, como o nível de exigência computacional para a sua utilização não é demasiadamente alto, e previa-se, neste estudo, redes pequenas com apenas algumas dezenas de neurônios na camada escondida, adotou-se o mesmo como o algoritmo de treinamento responsável pela redução da diferença entre a resposta desejada e a fornecida por cada rede projetada, em vista de sua eficiência já comprovada.

Os coeficientes de correlação resultantes entre os dados observados e os fornecidos pela rede foram adotados como índice indicativo do desempenho do treinamento de cada rede. O treinamento seria encerrado quando o ganho nesses coeficientes fosse insignificante. Constatou-se que, em vista da rápida velocidade com que o algoritmo

escolhido provoca a convergência de uma rede, todas as redes projetadas convergiram em algumas dezenas de iterações.

Apesar da possibilidade de se utilizar uma função de ativação diferente para cada neurônio da rede, geralmente atribui-se uma única função para os neurônios de uma mesma camada escondida (Hines, 1997). Uma rede multicamada com uma única camada escondida de neurônios sigmoidais, que recebem entradas diretamente, e então transmitem suas saídas para uma camada de neurônios lineares, tem provado ser capaz de aproximar qualquer função contínua com satisfatória exatidão, sendo considerada como aproximador universal de funções (Hines, 1997; Demuth e Beale, 1998; Valença e Ludermir, 1999a). Com base nesta indicação, foram projetadas redes com apenas três camadas neste estudo. A camada de entrada de cada rede não faz processamento, ou seja, não se atribui a ela funções de ativação. Ela simplesmente envia os dados de entrada, modificados por um peso, para cada um dos neurônios da próxima camada. A segunda camada, única escondida em cada rede neural, é composta de neurônios com função de ativação sigmóide; e a camada de saída é formada por neurônios com função linear.

Para determinar o número de neurônios de uma camada, alguns projetistas escolhem iniciar com uma rede razoavelmente grande, encontrando com facilidade o ajuste desejado; então eles extraem neurônios até obter a menor rede que ainda atinja um resultado satisfatório. Outros projetistas preferem começar com uma rede pequena, fazendo-a “crescer” até alcançar o ajuste desejado, o que é recomendável uma vez que o tempo de treinamento aumenta bastante com o tamanho da rede (Valença e Ludermir, 1999b). Este foi o procedimento adotado para a determinação do número de neurônios da camada escondida de cada rede neural projetada para a desagregação. Os números de neurônios na camada de entrada e de saída foram determinados pela dimensão dos campos de entrada e de saída, respectivamente, ou seja, como se deseja relacionar um dado de precipitação média da célula a um único valor médio da bacia, as camadas de entrada e de saída foram formadas por apenas um neurônio cada uma. A configuração de rede da Figura 3.7 ilustra o processo de desagregação.

Um dos procedimentos clássicos de validação, conforme Valença e Ludermir (1999b), é escolher algumas arquiteturas para a rede, realizar o treinamento e fazer a validação de tal forma que, através de um índice indicativo (erro médio quadrático, coeficiente de correlação,...), se possa escolher a de melhor ajuste. Este foi o procedimento adotado neste trabalho, apesar de existirem divergências sobre tal procedimento, uma vez

que o grupo de dados destinado para a validação acaba por participar do processo de seleção da rede, o que vai de encontro à idéia original de validação, que exige dados independentes e nunca utilizados. Para contornar isso, o que se pode fazer é dividir a série disponível em três conjuntos de dados. O primeiro seria empregado no treinamento da rede, o segundo, utilizado para a verificação de acordo com o procedimento aqui descrito, e o terceiro grupo seria utilizado na validação propriamente dita.

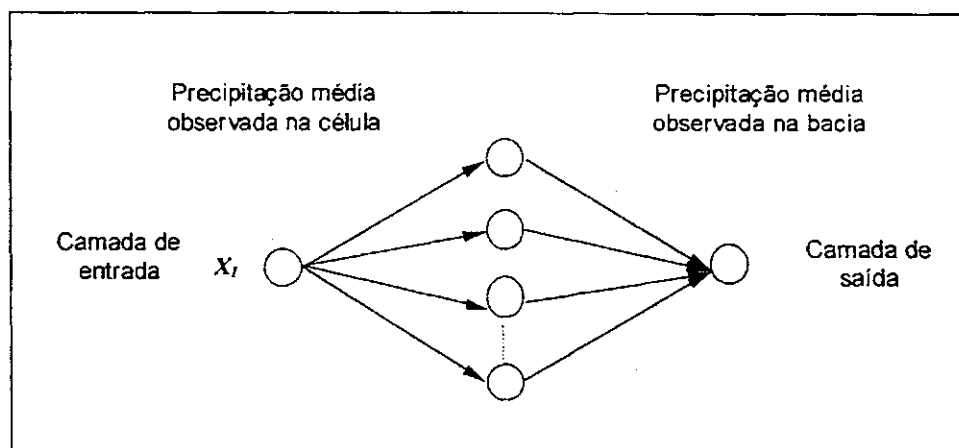


Figura 3.7 – Configuração da rede

O fluxograma da Figura 3.8 indica os passos que foram tomados para a escolha da melhor rede. Dividiu-se a série histórica em dois grupos, realizando-se o treinamento de várias redes de arquiteturas semelhantes – que diferem apenas pelo número de neurônios na camada escondida – com o primeiro grupo. Em seguida, efetuou-se a validação de todas elas com o segundo grupo de dados e, com base nos coeficientes de correlação gerados nas fases de treinamento e validação, extraiu-se destas um conjunto formado por aquelas que obtiveram, aparentemente, os melhores desempenhos. Por existir uma sensibilidade dos algoritmos de treinamento em relação aos pesos e ruídos iniciais, fez-se 10 calibrações para cada estrutura de rede selecionada, com o objetivo de se verificar a “estabilidade” de cada estrutura de rede. Entende-se por uma estrutura estável quando, nas redes resultantes das 10 calibrações, o algoritmo de treinamento consegue convergir para uma solução aceitável, fornecendo coeficientes de correlação semelhantes para todas elas. Aquela rede que fornecesse o melhor coeficiente de correlação na validação, e que ainda possuísse uma estrutura estável, seria a mais indicada para a desagregação espacial. No entanto, além destes critérios de seleção, considerou-se ainda o princípio da parcimônia

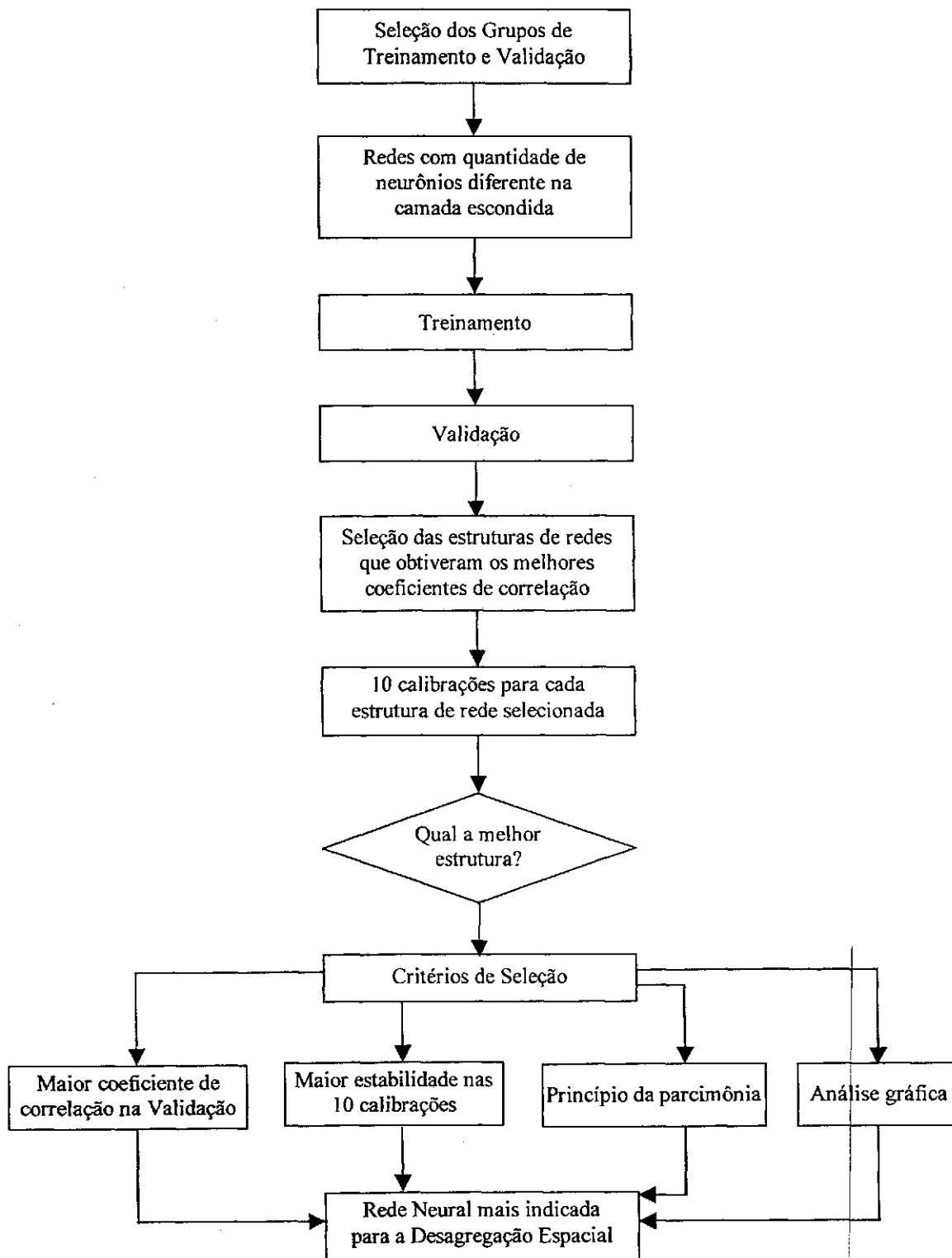


Figura 3.8 – Fluxograma do processo de seleção da rede neural para a desagregação espacial

para a efetiva escolha da melhor rede. Este princípio consiste na representação adequada do comportamento de um processo ou um sistema por um modelo com o menor número possível de parâmetros (Tucci, 1998). Isto é, a melhor rede será aquela que contenha o menor número de neurônios na camada escondida e apresente um desempenho semelhante às redes com mais neurônios.

A ferramenta utilizada para a implementação da rede neural foi o MATLAB, versão 5.3. Sistema interativo e linguagem de programação para computação técnica e científica em geral, o MATLAB permite a construção de “bibliotecas” de funções, denominadas de *toolboxes*, que o tornam apto a resolver classes particulares de problemas (Hanselman e Littlefield, 1999). Como essas *toolboxes* foram escritas usando a linguagem do MATLAB, isto permite não só o exame de comandos internos de algoritmos que a compõem, como também a criação de algoritmos novos.

Utilizou-se a *toolbox* de Redes Neurais do MATLAB neste estudo. Ela possibilita o uso dos mais variados tipos de redes neurais, permite a modificação de redes já existentes e a criação de outras desejáveis, e apresenta a enorme vantagem de possibilitar rapidez e eficácia na implementação e na simulação de uma rede neural.

As funções de ativação TANSIG e PURELIN, implementadas na *toolbox* de Redes Neurais, foram utilizadas nas camadas escondida e de saída, respectivamente, neste trabalho. A função de ativação sigmoideal tangente hiperbólica, denominada TANSIG, é frequentemente empregada em redes multicamadas. Ela gera saídas no intervalo de -1 a $+1$ (Figura 3.9), calculando-as de acordo com a equação abaixo, onde n é o somatório dos produtos das saídas da camada anterior pelos pesos de ligação, e comumente também acrescido de um ruído.

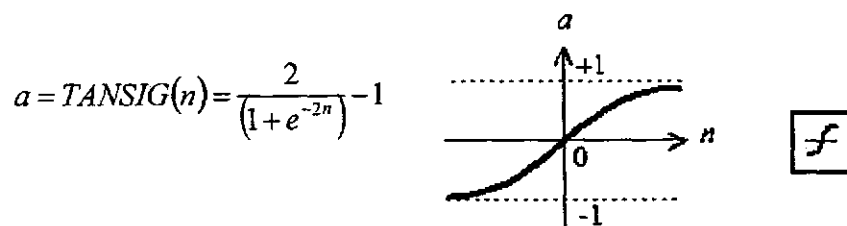


Figura 3.9 – Algoritmo, gráfico e símbolo utilizados pela função TANSIG (Demuth e Beale, 1998)

A função PURELIN é uma função de ativação linear também utilizada em redes multicamadas. É freqüentemente empregada em camadas de saída, pois não restringe os resultados da rede a intervalos, permitindo que os mesmos assumam qualquer valor (Figura 3.10).

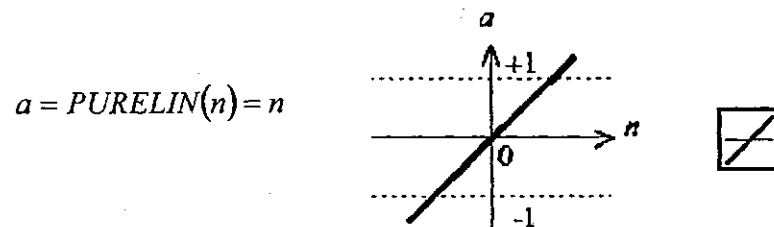


Figura 3.10 – Algoritmo, gráfico e símbolo utilizados pela função PURELIN (Demuth e Beale, 1998)

Uma rede do tipo *perceptron* multicamadas formada apenas por neurônios lineares comporta-se como uma equação de regressão. Para demonstrar isso, projetou-se uma rede com três camadas, sendo as duas últimas compostas de neurônios com a função de ativação PURELIN.

4. RESULTADOS

4.1. DIAGNÓSTICO DAS RELAÇÕES DA PRECIPITAÇÃO ENTRE AS DUAS ESCALAS

4.1.1. NÍVEL SAZONAL

As séries temporais da célula e da bacia da média dos dados dos postos pluviométricos selecionados estão mostradas na Figura 4.1. Verificou-se que entre estas existe uma relação aproximadamente linear (Figura 4.2), com um coeficiente de correlação de 0,85 para o período de 1911 a 1988.

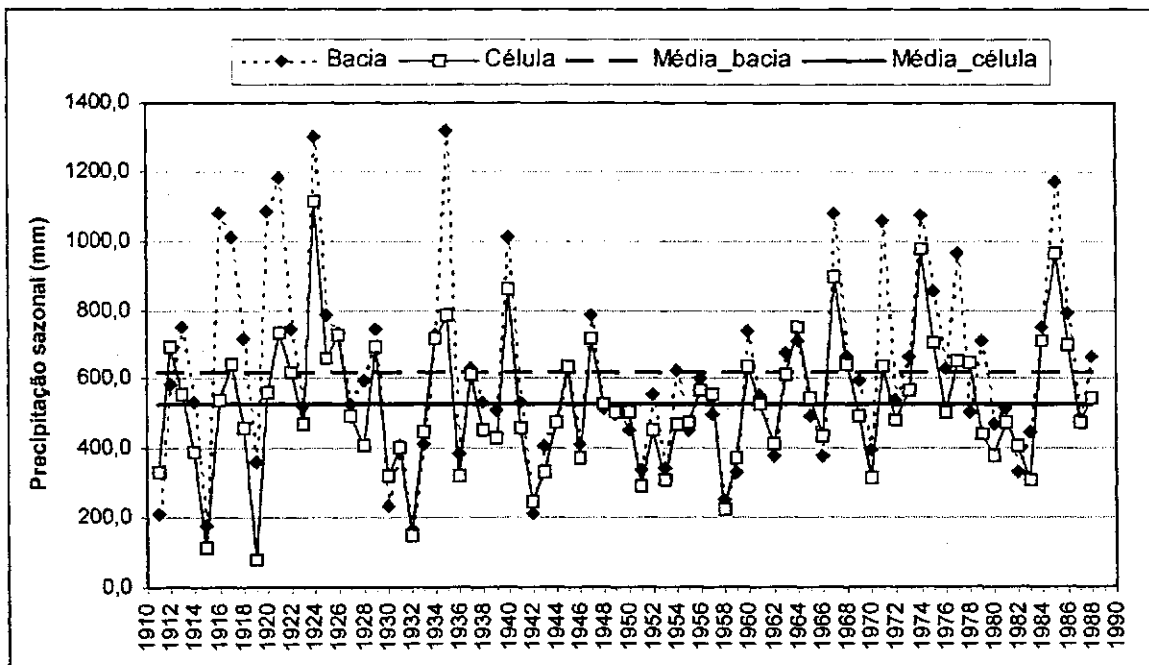


Figura 4.1 – Séries temporais de precipitação sazonal (fevereiro-maio) da célula e da bacia do Rio Piancó (1911-88)

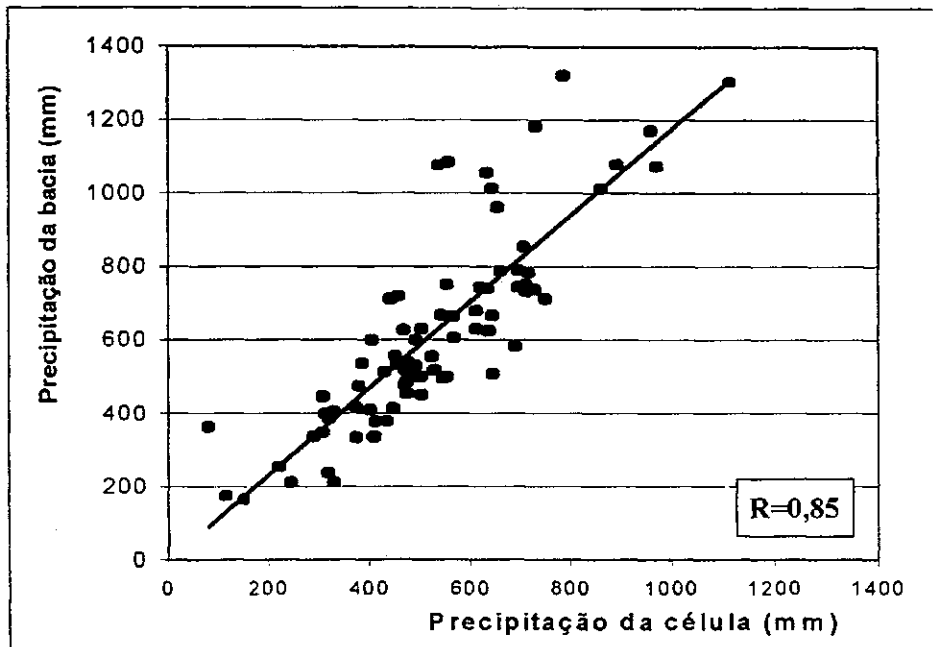


Figura 4.2 – Relação entre dados sazonais (fevereiro-maio) de precipitação da célula e da bacia, para o período de 1911 a 1988

Apesar da constatação, através da Figura 4.2, de que a relação entre as séries da célula e da bacia é aproximadamente linear, é nítida a existência de dados bastante dispersos. Ao se analisar a Figura 4.1, constatou-se que essa dispersão está relacionada, principalmente, às três primeiras décadas do período analisado. Atribui-se a isto o fato de cinco dos nove postos selecionados para representar a bacia terem sido instalados em 1933, atingindo efetiva operação em 1938. Outras possíveis causas seriam os meios e instrumentos de medição utilizados na época, ou ainda, a ocorrência de mudanças climáticas. Notou-se também que, além das três primeiras décadas, alguns anos também mostraram o comportamento da precipitação bem diferente dos demais. Estes anos foram: 1971, 1977, 1978 e 1979. Duas possibilidades são consideradas: ou se trata de anos realmente anômalos, relacionados a alguma anomalia nos sistemas atmosféricos; ou é o resultado de falhas em coleta e transcrição de dados. É importante salientar que a anomalia aqui considerada diz respeito não ao fato de que os totais precipitados desses anos fugiram ao intervalo esperado, mas refere-se a anos em que a relação entre a precipitação na célula e na bacia, observada nos outros anos, não foi verificada.

Desejando-se averiguar se a relação entre a precipitação na célula e na bacia é estável, calculou-se coeficientes de correlação entre as duas séries para uma “janela” de 30 anos, móvel ano a ano sobre todo o período (Figura 4.3). Os problemas detectados também

foram verificados, visto que os coeficientes passaram a ser menores ou maiores à medida que alguns dos anos mencionados eram incluídos ou excluídos da janela. Isto é comprovado ao se observar os coeficientes das janelas que tiveram início nos anos 1936 a 1941 (Figura 4.4), períodos em que as três primeiras décadas e os anos citados não estão inclusos, constatando-se que, para estes períodos, a relação bacia \times célula é bastante estável.

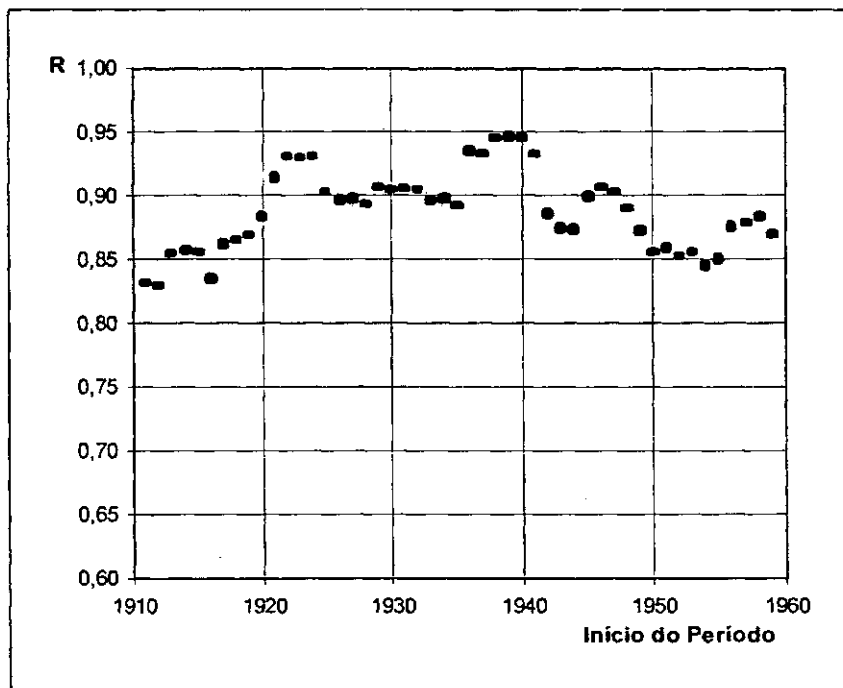


Figura 4.3 – Coeficientes de correlação entre as séries de precipitação sazonal da bacia e da célula, para uma janela móvel de 30 anos

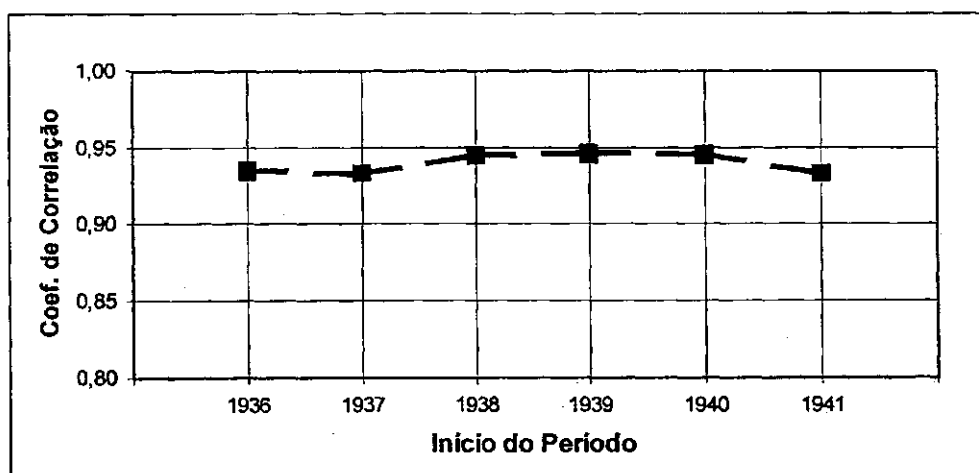


Figura 4.4 – Coeficientes de correlação entre as séries sazonais de precipitação da bacia e da célula, para uma janela móvel de 30 anos, iniciando entre 1936 e 1941

Visando aumentar a estabilidade da relação das precipitações entre as duas escalas, as três primeiras décadas das séries foram descartadas, utilizando-se apenas o período de 1939 a 1988 (Figura 4.5) para a obtenção da relação de desagregação. Após investigar-se os anos anômalos detectados, conferindo-se os dados originais e as operações realizadas entre estes, não se descobriu nada que os invalidasse e, portanto, optou-se por considerá-los corretos nesta pesquisa. O coeficiente de correlação linear encontrado para o período de 1939 a 1988 foi de 0,90.

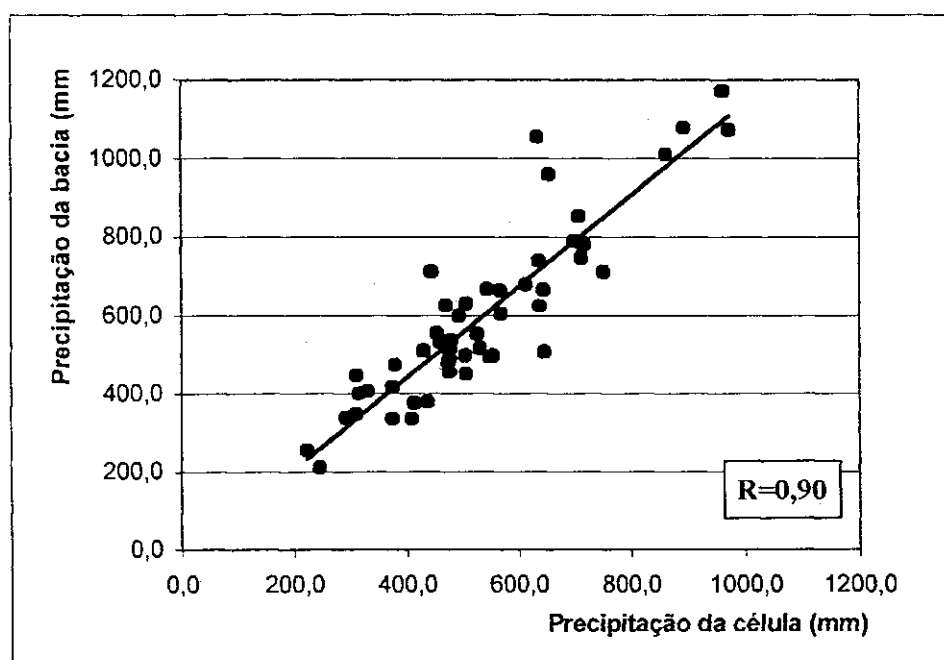
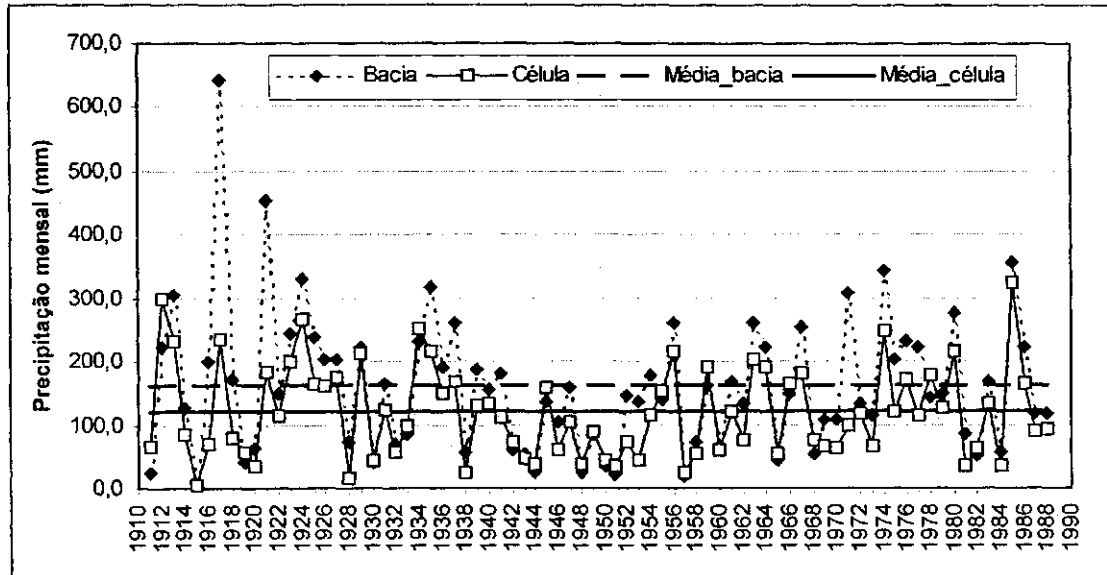


Figura 4.5 – Relação entre dados sazonais (fevereiro-maio) de precipitação da célula e da bacia, para o período de 1939 a 1988

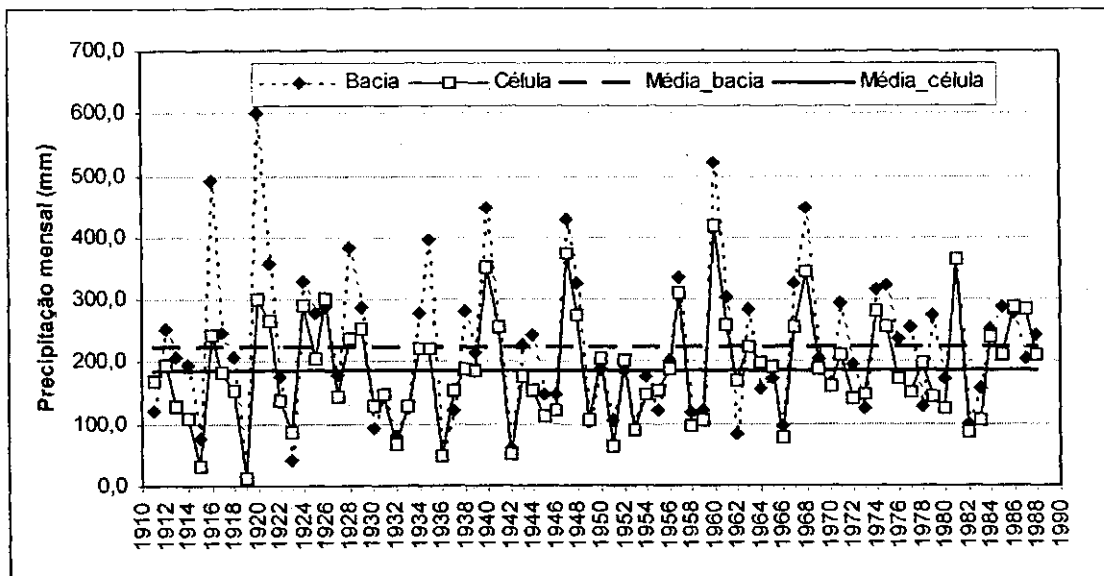
4.1.2. NÍVEL MENSAL

O mesmo procedimento foi adotado para a análise das séries de precipitação que se referem aos meses da quadra chuvosa da bacia do Piancó (Figura 4.6), séries essas também oriundas de médias entre os postos pluviométricos selecionados. Elas também apresentaram uma relação aproximadamente linear entre dados da célula e da bacia (Figura

4.7). Os coeficientes de correlação obtidos entre essas séries mensais, para o período de 1911 a 1988, estão expostos na Figura 4.7 e na Tabela 4.1.

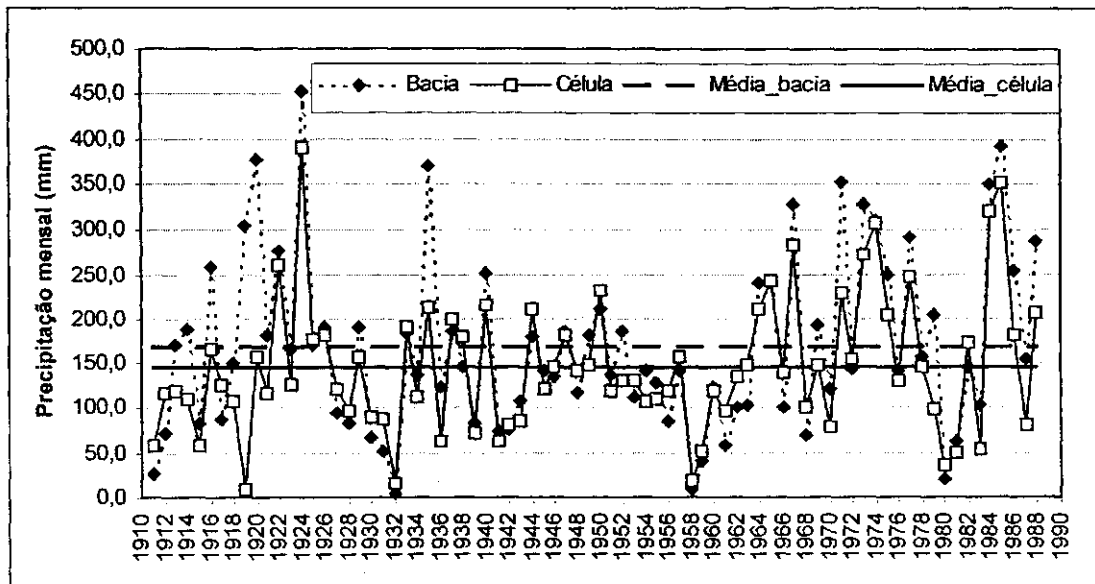


a) fevereiro

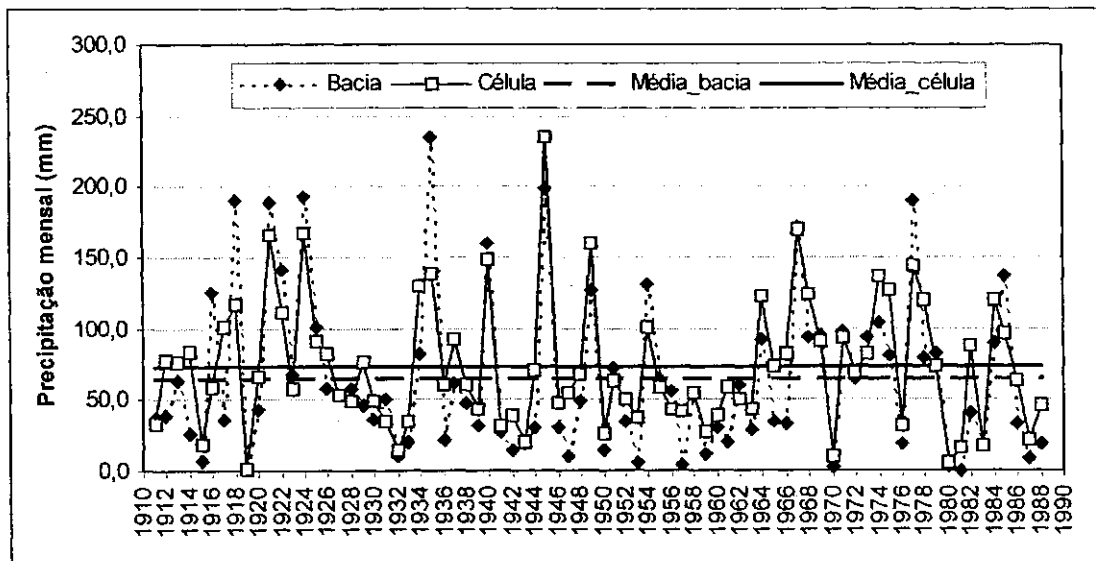


b) março

Figura 4.6 – Séries temporais de precipitação da célula e da bacia do Rio Piancó, para os meses de: a) fevereiro; b) março; c) abril; e d) maio



c) abril



d) maio

Figura 4.6 – Séries temporais de precipitação da célula e da bacia do Rio Piancó, para os meses de: a) fevereiro; b) março; c) abril; e d) maio (continuação)

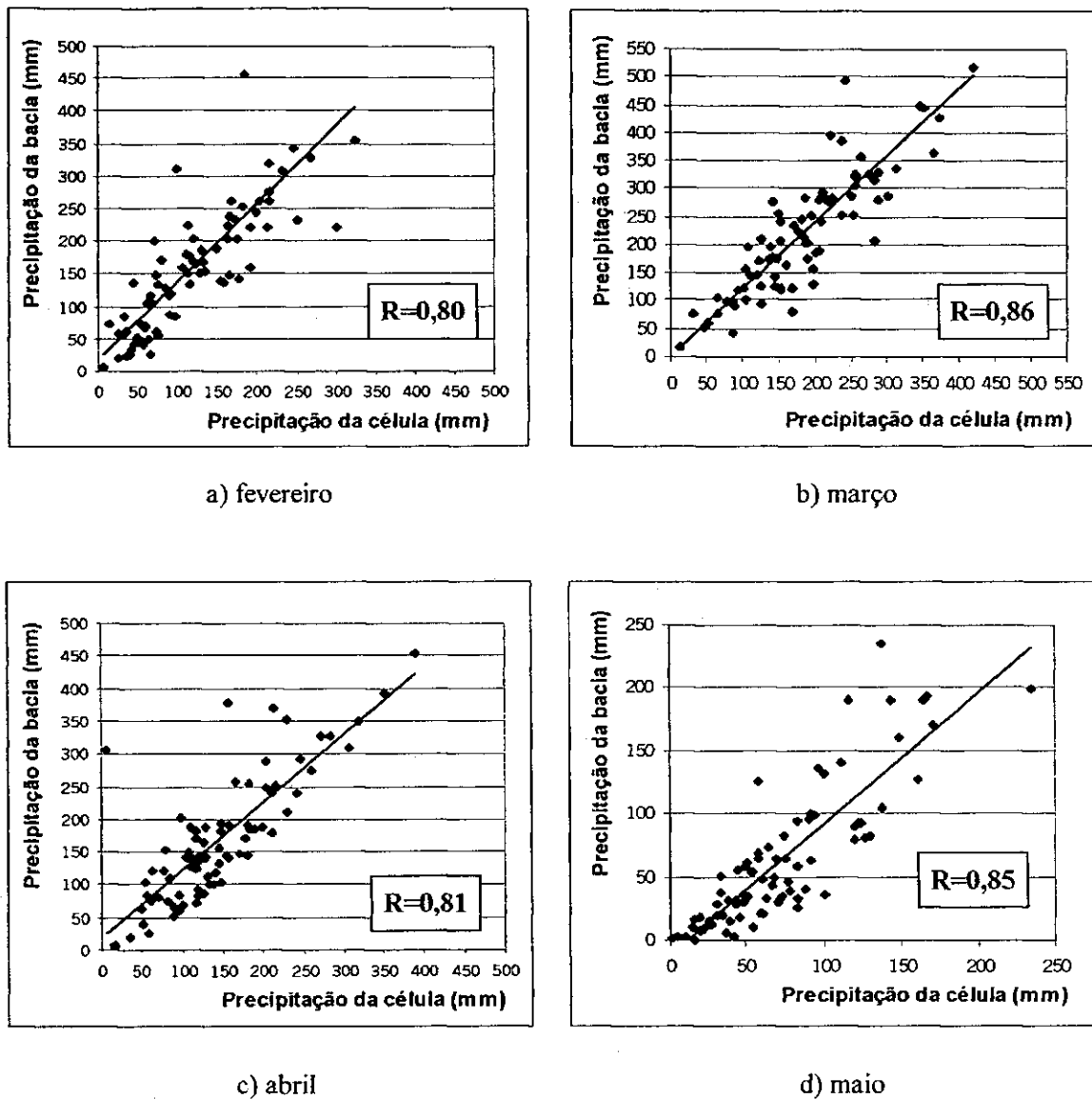


Figura 4.7 – Relação entre dados mensais de precipitação da célula e da bacia, para o período de 1911 a 1988, nos meses de: a) fevereiro; b) março; c) abril; e d) maio

Objetivando verificar se a relação existente entre os dados da célula e da bacia é estável para cada mês, calculou-se coeficientes de correlação para janelas móveis de 30 anos (Figura 4.8). Da mesma forma que para a escala sazonal, detectou-se problemas nas três primeiras décadas e nos anos de 1971, 1977, 1978 e 1979. Isto pode ser verificado também ao se observar o comportamento dos coeficientes das janelas que tiveram início

nos anos 1936 a 1941 (Figura 4.9), períodos em que as três primeiras décadas, e os meses que integram os anos citados, não estão inclusos.

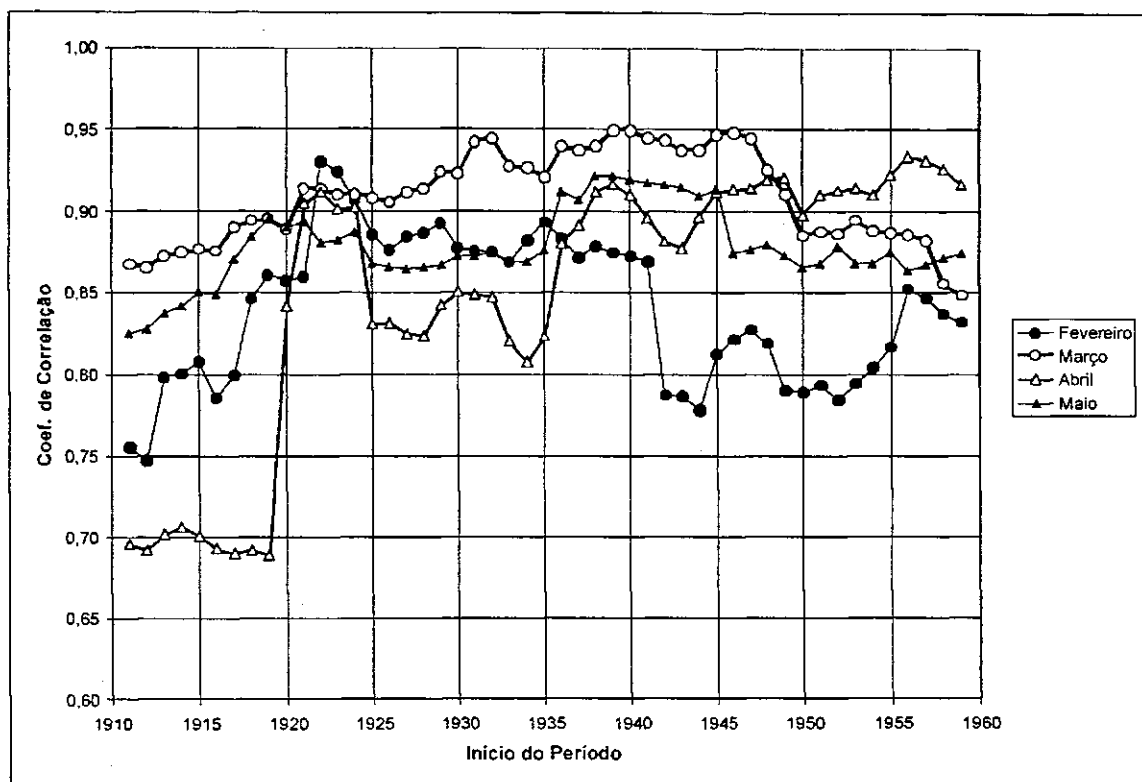


Figura 4.8 – Coeficientes de correlação entre as séries mensais de precipitação da bacia e da célula, para uma janela móvel de 30 anos

Para o mês de fevereiro, verificou-se que grandes distorções entre os dados da bacia e os da célula ocorreram nos anos de 1971 e 1977. O mês de março apresentou a maior distorção no ano de 1979, e o mês de abril se destacou em 1971 e 1979. Considerou-se como “grandes distorções” quando a diferença entre dados mensais da célula e da bacia atingiu 60% da média dos dados da célula. Nesse contexto, o mês de maio não apresentou diferenças significativas entre os seus dados.

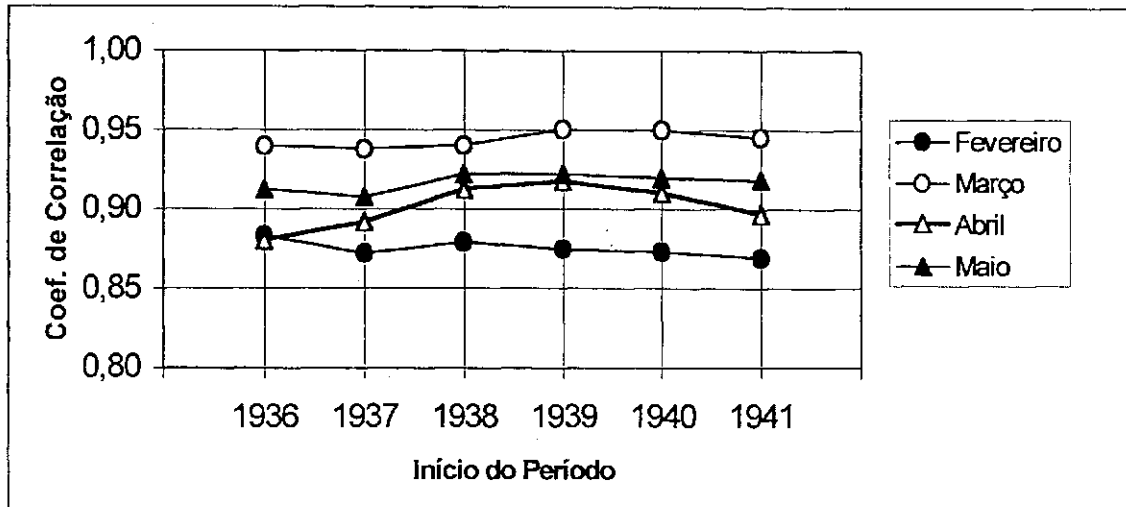


Figura 4.9 – Coeficientes de correlação entre as séries mensais de precipitação da bacia e da célula, para uma janela móvel de 30 anos, iniciando entre 1936 e 1941

Semelhante às conclusões resultantes da análise feita para o nível sazonal, optou-se também por descartar as três primeiras décadas das séries mensais, visando aumentar a estabilidade das relações entre as escalas. Os meses “anômalos” também foram considerados corretos, uma vez que foram investigados e não se achou nada que os invalidasse.

Os coeficientes de correlação encontrados entre as séries mensais do período de 1939 a 1988 (Figura 4.10) estão expostos na Tabela 4.1.

Tabela 4.1 - Coeficientes de correlação obtidos entre as séries mensais para os períodos de 1911-88 e 1939-88

Período da Série	Coeficiente de Correlação			
	Fevereiro	Março	Abril	Maio
1911-88	0,80	0,86	0,81	0,85
1939-88	0,85	0,90	0,92	0,90

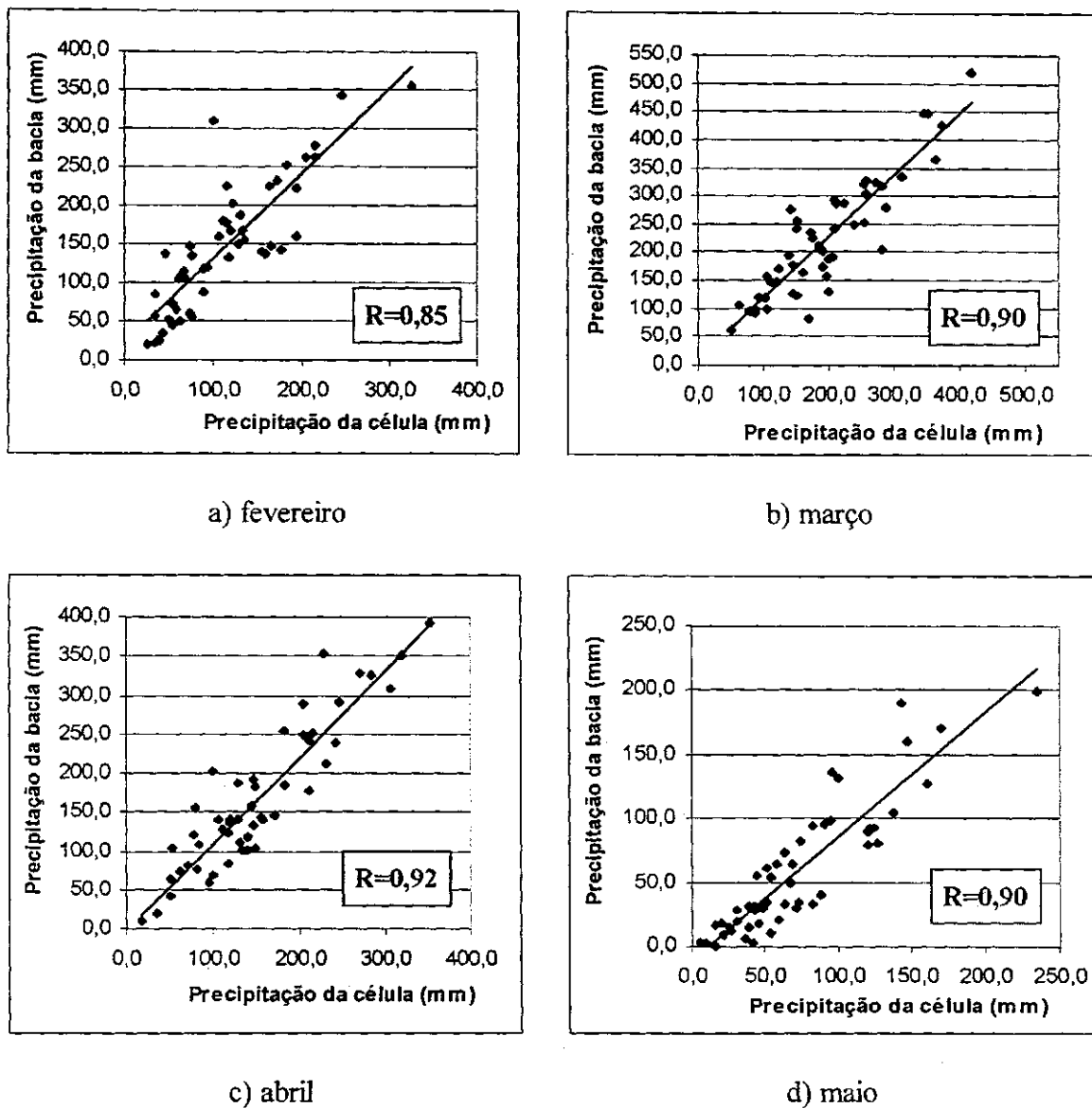


Figura 4.10 – Relação entre dados mensais de precipitação da célula e da bacia, para o período de 1939 a 1988, nos meses de: a) fevereiro; b) março; c) abril; d) maio

Ao fim da etapa de diagnóstico, observou-se que os dados de precipitação sazonal e mensal nas escalas espaciais do modelo de circulação global e da bacia do Piancó apresentam uma relação aproximadamente linear, com altas correlações, o que torna possível a realização da transferência da previsão entre as duas escalas, sem perdas significativas de previsibilidade, pelo menos na região estudada.

4.2. DESAGREGAÇÃO VIA REGRESSÃO LINEAR

Definiu-se o período de 1939 a 1968 para a calibração dos dois métodos, e o período de 1969 a 1988 para a validação dos mesmos. Esta definição foi arbitrária, considerando-se apenas um grupo maior de dados para a calibração.

4.2.1. NÍVEL SAZONAL

Na calibração do modelo para a escala temporal sazonal (Figura 4.11), obteve-se a seguinte equação de regressão

$$P_b = 1,134P_c - 38,686 \quad (4.1)$$

onde P_c é o total sazonal precipitado na célula e P_b a precipitação correspondente na bacia, em mm.

O coeficiente de correlação obtido no período de calibração foi de 0,95.

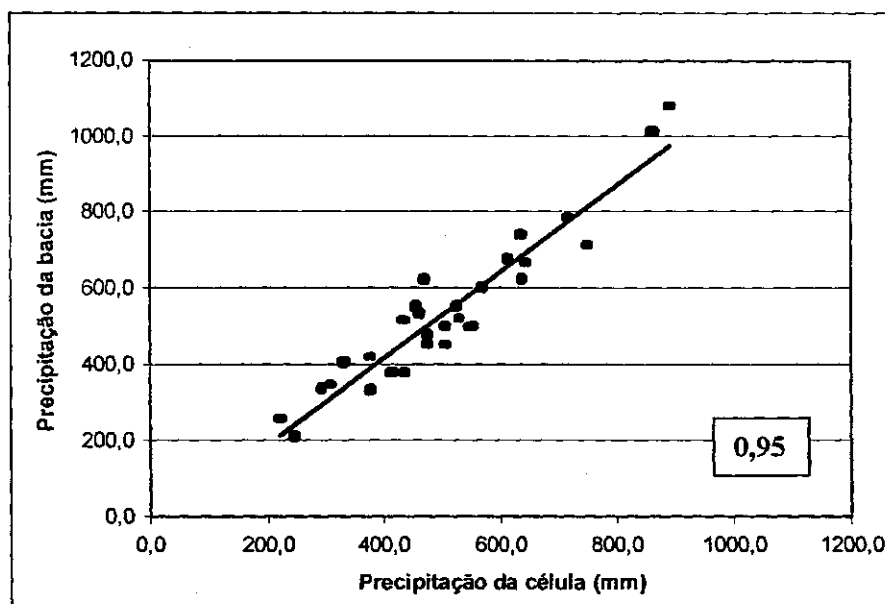


Figura 4.11 – Calibração do modelo de Regressão Linear entre dados de precipitação sazonal, para o período de 1939 a 1968

Para a validação do ajuste (Figura 4.12), o coeficiente de correlação encontrado entre os dados observados da bacia e os valores obtidos da equação de regressão foi de 0,86. Esta queda no coeficiente de correlação se deve, especialmente, aos anos identificados como anômalos (1971, 1977, 1978 e 1979). Isso pode ser também observado na Figura 4.13, que apresenta a relação entre os dados observados e os calculados pela equação para todo o período (1939-88).

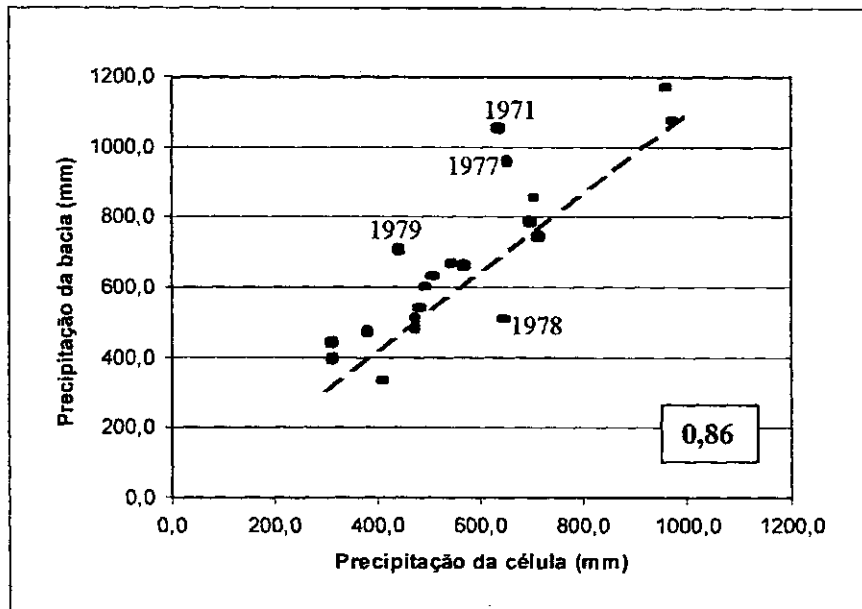


Figura 4.12 – Validação modelo de Regressão Linear entre dados de precipitação sazonal, para o período de 1969 a 1988

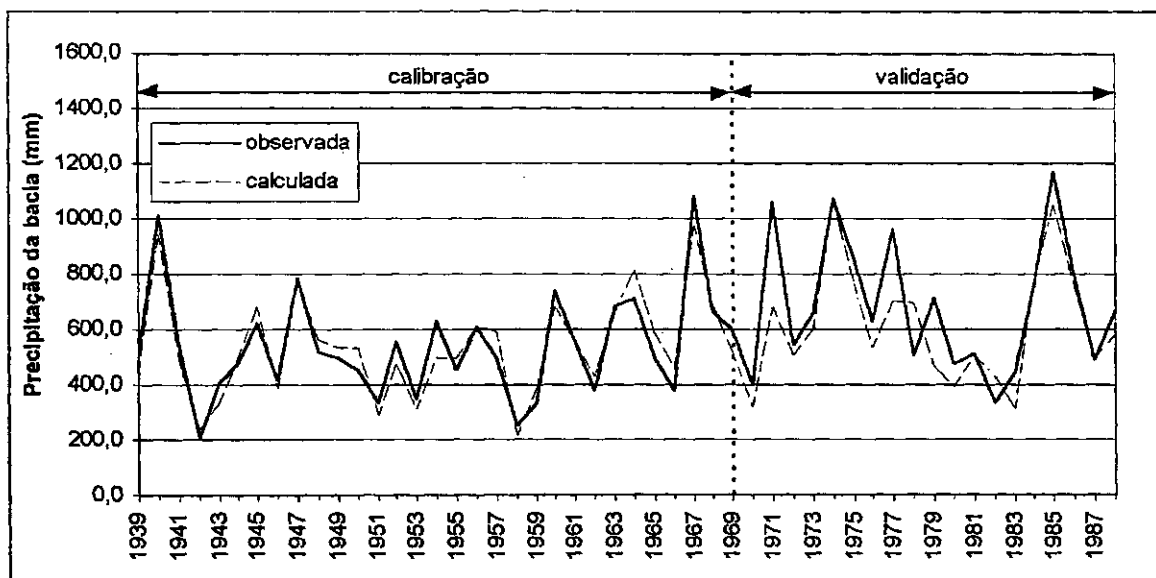


Figura 4.13 – Comparação entre o total precipitado sazonal observado e calculado pelo modelo de Regressão Linear para a bacia do Piancó

4.2.2. NÍVEL MENSAL

Na calibração dos modelos de regressão para os meses de Fevereiro a Maio (Figura 4.14), as equações obtidas para cada mês foram as seguintes:

$$\text{Fevereiro: } P_b = 1,111P_c + 9,443 \quad (4.2)$$

$$\text{Março: } P_b = 1,214P_c - 16,243 \quad (4.3)$$

$$\text{Abril: } P_b = 1,065P_c - 9,554 \quad (4.4)$$

$$\text{Maio: } P_b = 0,957P_c - 12,807 \quad (4.5)$$

onde P_c é o total mensal precipitado na célula e P_b a precipitação correspondente na bacia, em mm.

Efetuiu-se a validação das equações de ajuste, e novamente verificou-se a queda dos coeficientes de correlação devido aos dados dos denominados anos anômalos, como se pode ver na Figura 4.15. Excluiu-se das séries os anos que apresentaram anomalia (indicados na Figura 4.15), para a verificação da influência desses no comportamento dos coeficientes de correlação. Constatou-se que, para os meses de fevereiro, março e abril, os coeficientes apresentariam um aumento significativo. O mês de fevereiro, que apresentou um coeficiente de 0,83, passaria a ter um de 0,96, março que é de 0,73, subiria para 0,85, e o mês de abril passaria de 0,93 a 0,96.

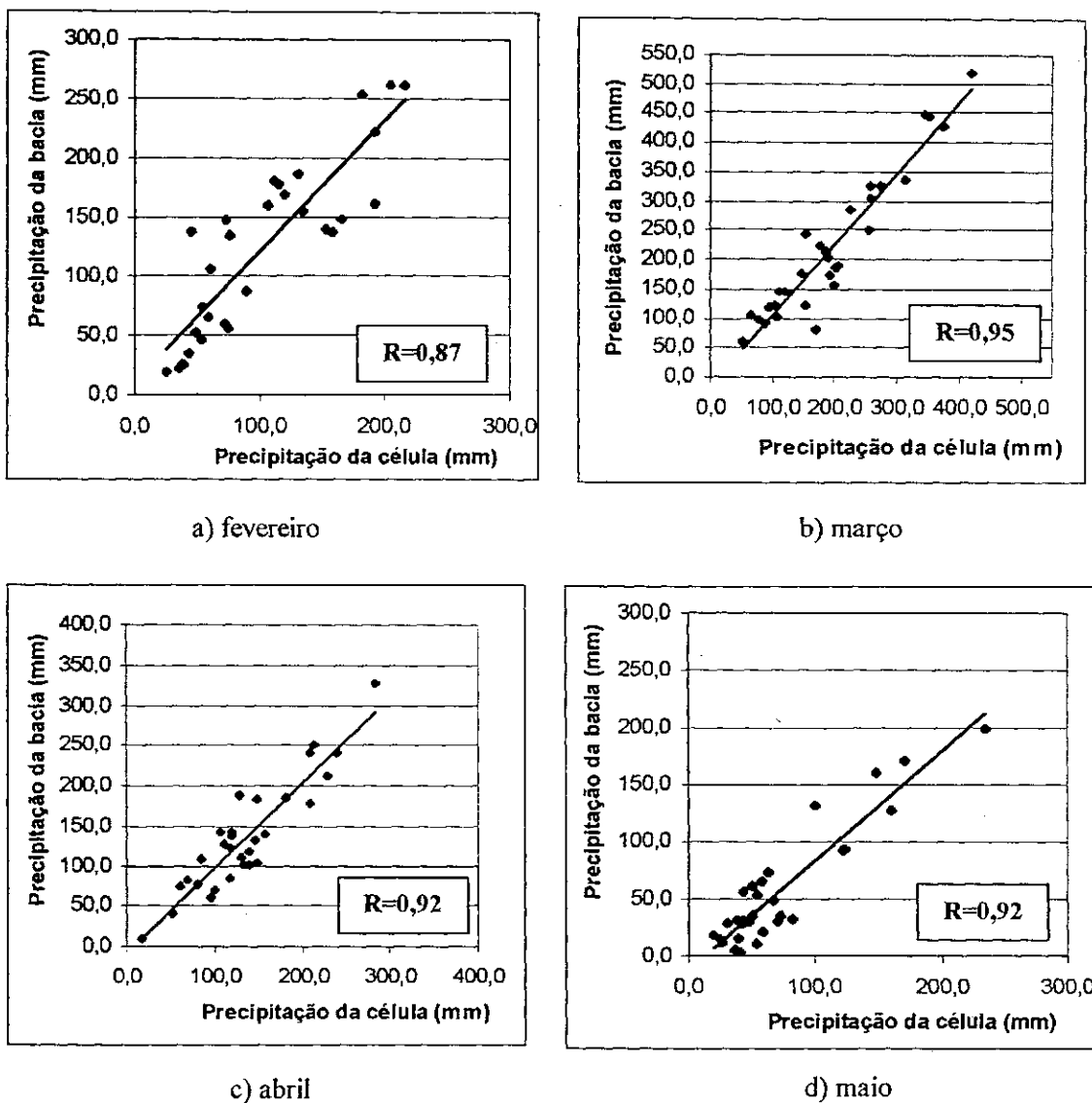


Figura 4.14 – Calibração do modelo de Regressão Linear entre dados de precipitação mensal, para o período de 1939 a 1968, nos meses de: a) fevereiro; b) março; c) abril; d) maio

Os coeficientes encontrados entre as séries nos períodos de calibração e de validação estão expostos na Tabela 4.2. A Figura 4.16 apresenta a relação resultante entre os dados observados e calculados pelas equações de regressão, para cada mês, em todo o período (1939-88).

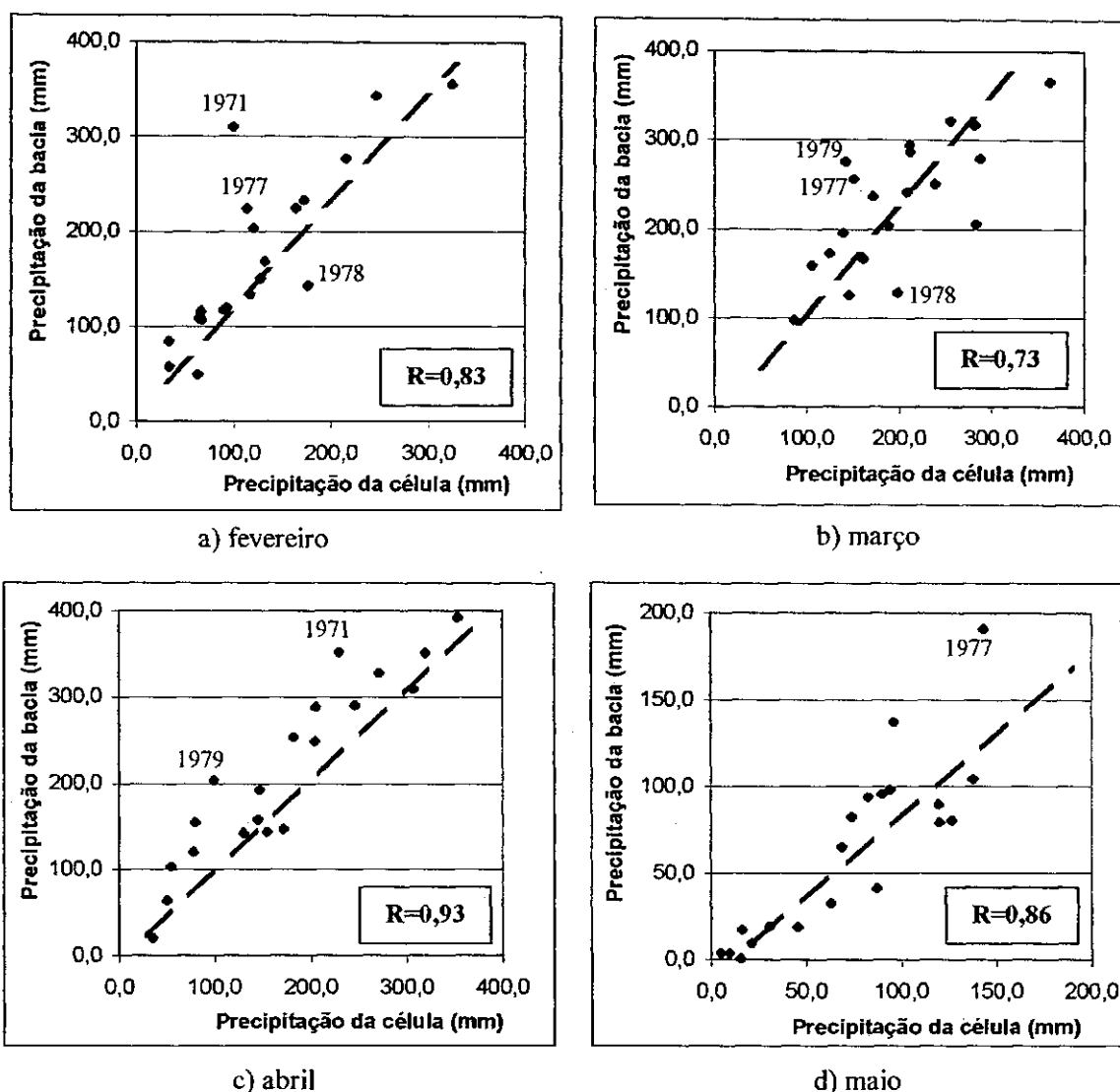
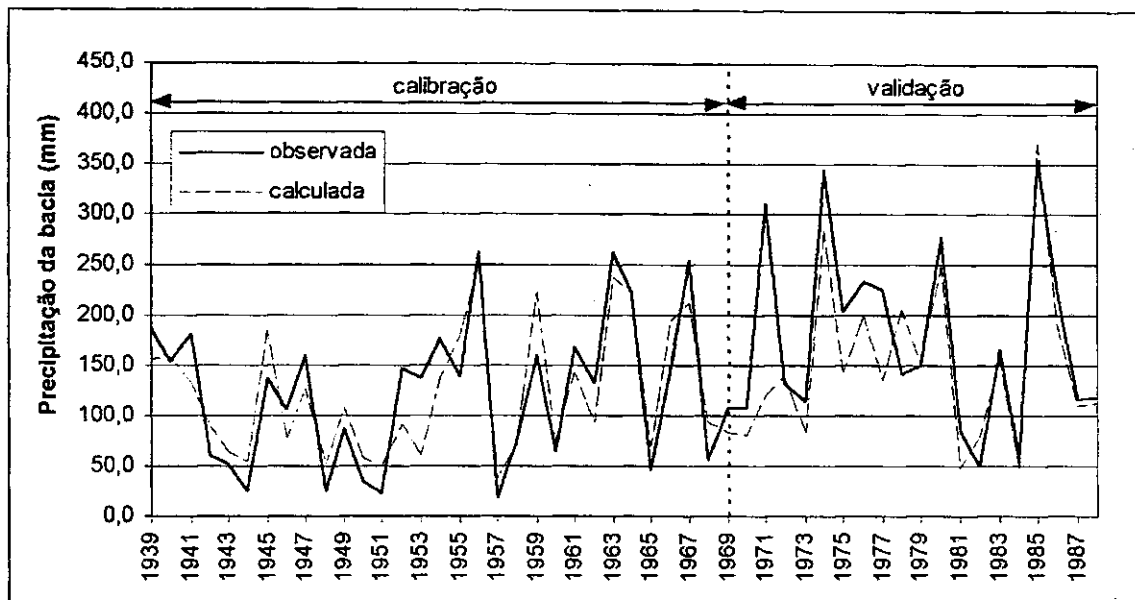


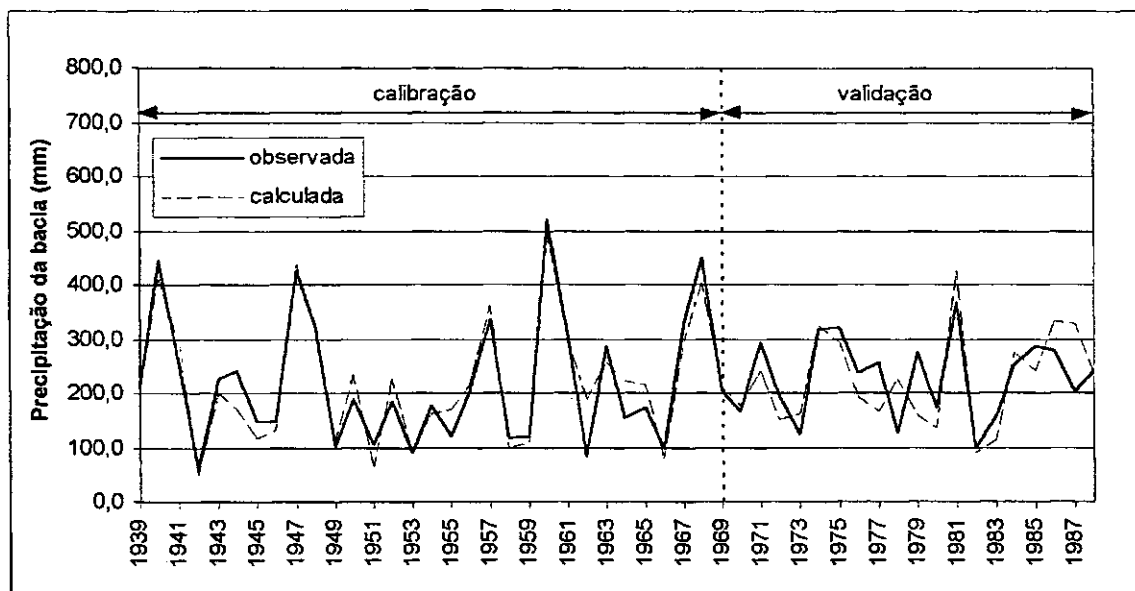
Figura 4.15 – Validação do modelo de Regressão Linear entre dados de precipitação mensal, para o período de 1969 a 1988, nos meses de: a) fevereiro; b) março; c) abril; d) maio

Tabela 4.2 – Coeficientes de correlação obtidos entre séries observadas e calculadas pela regressão linear

Escala temporal		Calibração	Validação
Sazonal		0,95	0,86
Mensal	Fevereiro	0,87	0,83
	Março	0,95	0,73
	Abril	0,92	0,93
	Maior	0,92	0,86

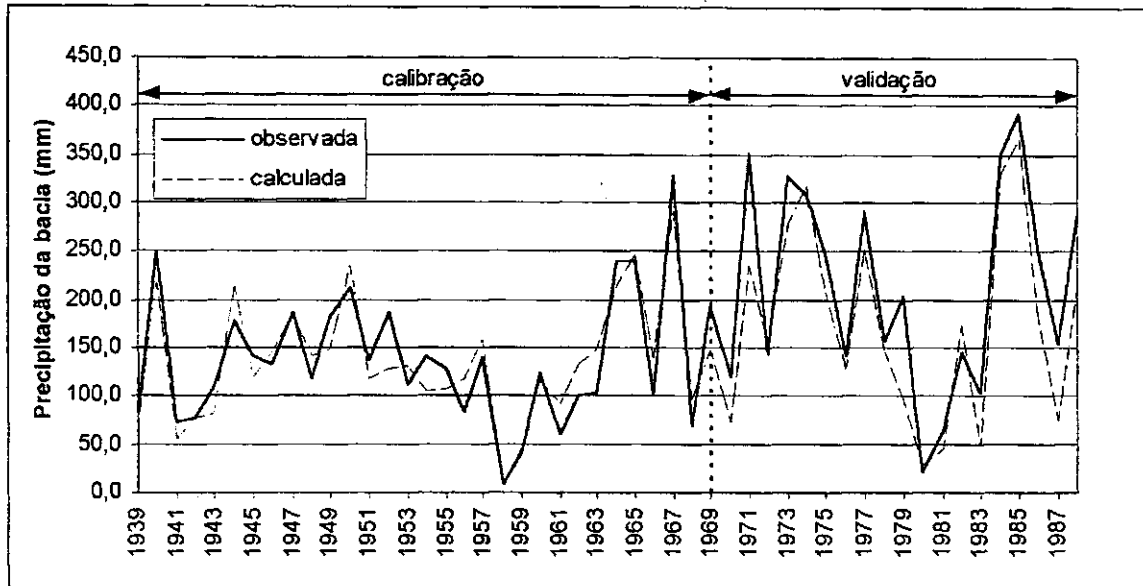


a) Fevereiro

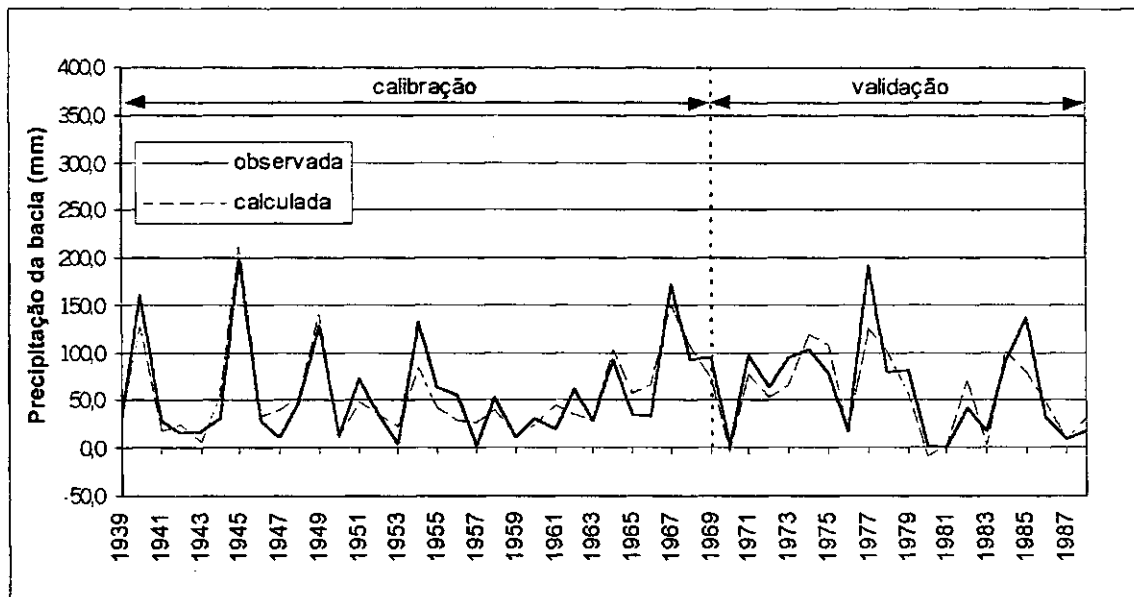


b) Março

Figura 4.16 – Comparação entre o total precipitado observado e calculado pelo modelo de Regressão Linear, para a bacia do Piancó, nos meses de: a) fevereiro; b) março; c) abril; d) maio



c) Abril



d) Maio

Figura 4.16 – Comparação entre o total precipitado observado e calculado pelo modelo de Regressão Linear, para a bacia do Piancó, nos meses de: a) fevereiro; b) março; c) abril; d) maio (continuação)

4.3. DESAGREGAÇÃO VIA REDES NEURAIS ARTIFICIAIS

4.3.1. REDES NEURAIS ARTIFICIAIS COMO REGRESSÃO LINEAR

Objetivando-se confirmar que uma rede neural formada apenas por neurônios lineares se comporta como uma equação de regressão, projetou-se uma rede com três camadas (Figura 4.17), onde a camada escondida é composta de dois neurônios PURELIN. Forneceu-se como dados de entrada e de saída, as séries sazonais de precipitação da célula e da bacia, respectivamente.

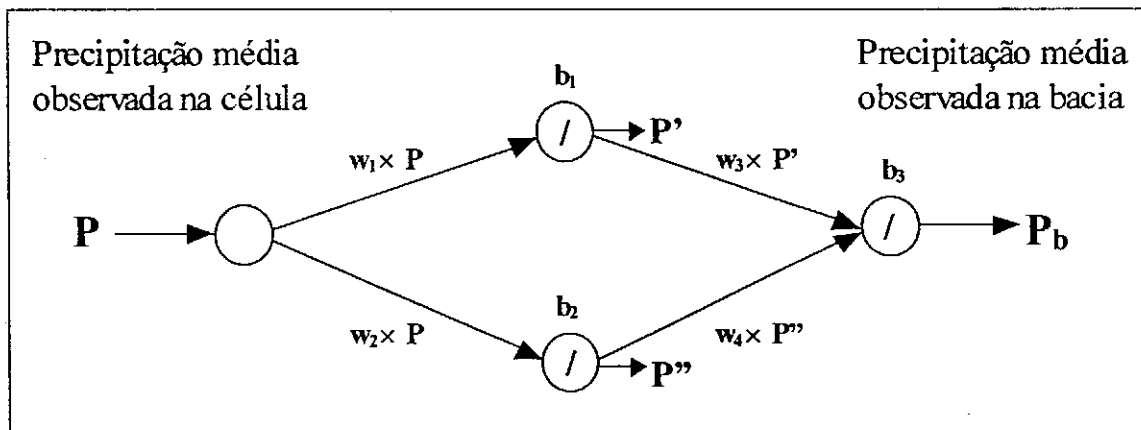


Figura 4.17 – Configuração da rede composta por neurônios lineares

As operações que devem ser efetuadas para se obter a equação que representa a rede neural são:

$$P' = w_1 \times P + b_1 \quad (4.6)$$

$$P'' = w_2 \times P + b_2 \quad (4.7)$$

$$P_b = w_3 \times P' + w_4 \times P'' + b_3 \quad (4.8)$$

onde:

P é a precipitação média observada na célula;

w_1, w_2, w_3 e w_4 são os pesos da rede;

$b_1, b_2,$ e b_3 são os ruídos inseridos nos três neurônios lineares;

P' corresponde à saída do primeiro neurônio linear da camada escondida;

P'' corresponde à saída do segundo neurônio linear da camada escondida;

P_b é a precipitação média a ser calculada para a bacia.

Substituindo as equações (4.6) e (4.7) na equação (4.8), obtém-se:

$$\begin{aligned} P_b &= w_3 \times (w_1 \times P + b_1) + w_4 \times (w_2 \times P + b_2) + b_3 \\ \Rightarrow P_b &= w_3 w_1 P + w_3 b_1 + w_4 w_2 P + w_4 b_2 + b_3 \\ \Rightarrow P_b &= (w_3 w_1 + w_4 w_2) P + (w_3 b_1 + w_4 b_2 + b_3) \end{aligned} \quad (4.9)$$

Como todos os elementos que estão inseridos entre os parênteses da equação (4.9) possuem valores constantes numa rede neural, pode-se igualar cada membro contido em parêntese a uma única constante.

$$w_3 w_1 + w_4 w_2 = a \quad (4.10)$$

$$w_3 b_1 + w_4 b_2 + b_3 = c \quad (4.11)$$

Substituindo as equações (4.10) e (4.11) na equação (4.9), obtém-se como a equação que representa a rede neural:

$$P_b = a \times P + c \quad (4.12)$$

mostrando que uma rede do tipo *perceptron* multicamadas formada por neurônios lineares comporta-se como uma equação de regressão.

Utilizando-se as séries de precipitação coletadas para esta pesquisa, a matriz dos pesos e ruídos resultante do treinamento (calibração) da rede é:

Pesos_e_ruídos =

$$\begin{pmatrix} -0,8404 \\ 0,7345 \\ -3,0606 \\ -5,1750 \\ 3,1113 \\ 5,1032 \\ -2,7586 \end{pmatrix} \begin{matrix} w_1 \\ w_2 \\ b_1 \\ b_2 \\ w_3 \\ w_4 \\ b_3 \end{matrix}$$

Aplicando estes valores na equação (4.9), obteve-se a seguinte equação:

$$P_b = 1,1336P - 38,6901 \quad (4.13)$$

Comparando a equação (4.13) com a equação de ajuste (4.1) resultante da regressão linear (seção 4.1), observa-se que elas são realmente semelhantes, mostrando que o algoritmo de treinamento da rede obteve valores dos parâmetros compatíveis com os obtidos pelo método dos mínimos quadrados na regressão linear.

4.3.2. PRÉ-PROCESSAMENTO DOS DADOS DE ENTRADA DE UMA REDE NEURAL

Antes de se iniciar a modelagem de um sistema qualquer através do método de redes neurais, vários estudos sugerem que se efetue o pré-processamento ou normalização dos dados de entrada da rede (Lawrence, 1991; Azoff, 1994; Gately, 1996; entre outros). Segundo Eklund (1994), trata-se de um processo em que se converte um conjunto de dados a um intervalo menor do que o original, tornando mais fácil o treinamento da rede, facilitando sua convergência. O autor afirma ainda que a normalização não só melhora o desempenho de uma rede neural, como reduz o seu tamanho, mostrando, inclusive, que

redes com um única camada que tiveram seus dados de entrada normalizados, podem “competir” com redes multicamadas cujos dados não tenham sido normalizados.

As simulações realizadas em redes projetadas com totais precipitados apresentaram alertas no MATLAB que diziam que os resultados obtidos poderiam não ser corretos. Verificou-se que essas redes geravam resultados bastante diferentes a cada simulação (Figura 4.18), apresentando-se muito instáveis, e expondo a necessidade de se fazer uma normalização dos dados antes da modelagem. Diante disto e com base na literatura, que recomenda que todos os dados de entrada de uma rede sejam normalizados – com exceção para dados que já estejam num “formato normalizado”, como os dados binários, ou para dados que se apresentem na mesma ordem de grandeza (Azoff, 1994) – efetuou-se a normalização dividindo-se cada valor utilizado da série pelo máximo valor (ou próximo dele) registrado na mesma, proporcionando o uso de dados que variam apenas entre 0 e +1, facilitando o treinamento e resultando em redes mais estáveis.

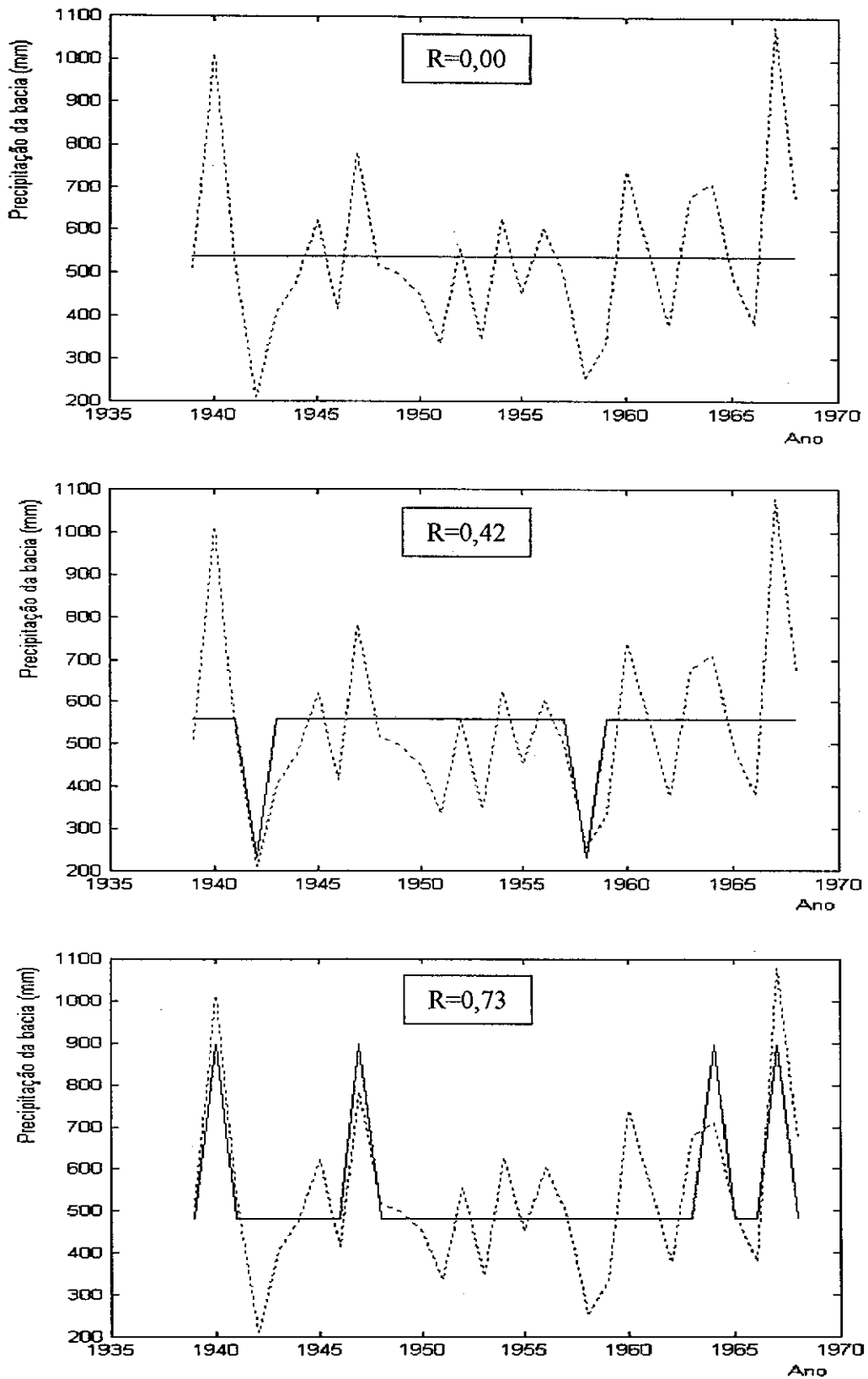


Figura 4.18 – Resultados obtidos em 10 simulações de uma rede neural com 5 neurônios sigmóides na camada escondida e 1 neurônio linear na camada de saída, em totais precipitados. A linha contínua é o resultado da simulação e a tracejada apresenta os dados observados

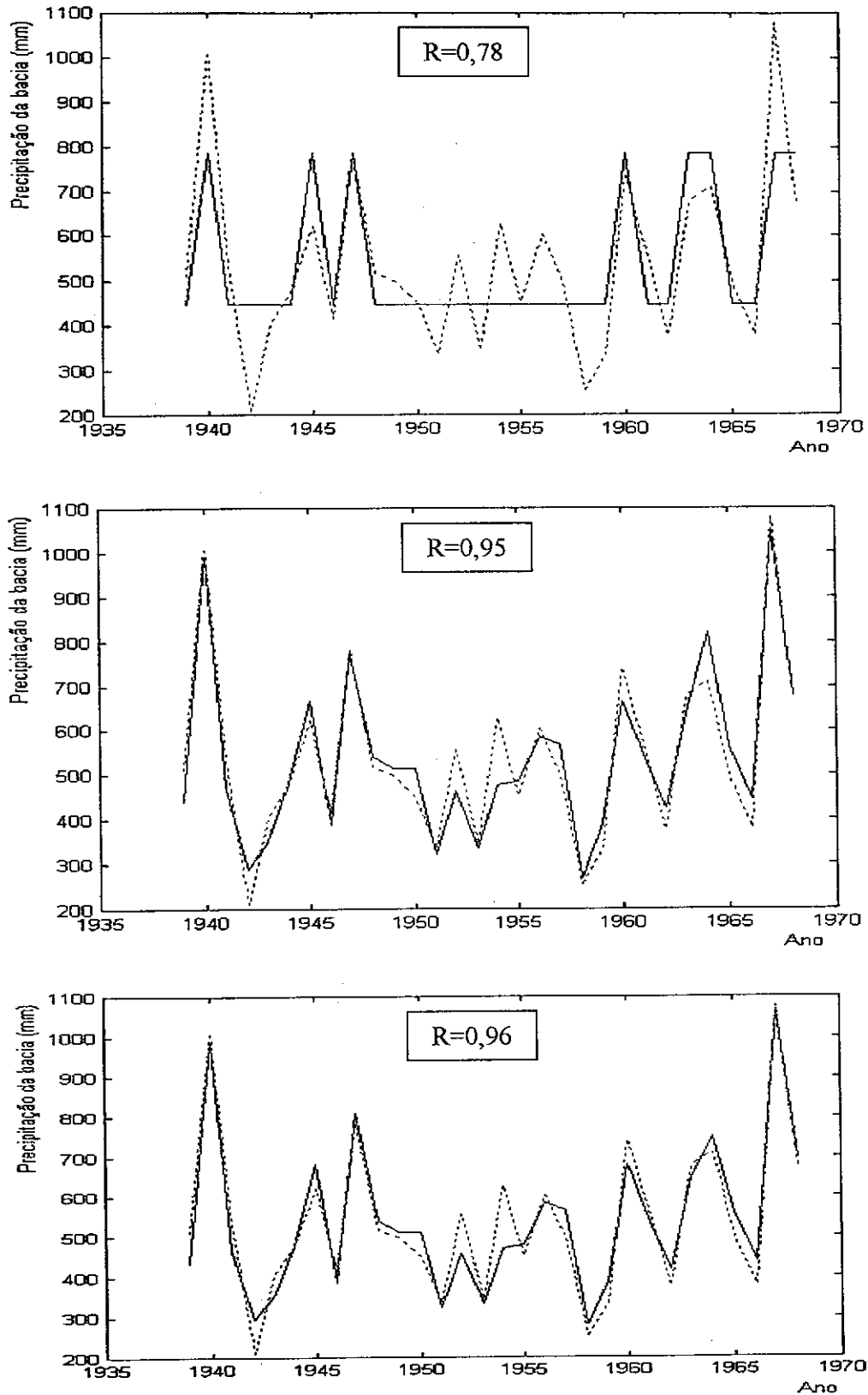


Figura 4.18 – Resultados obtidos em 10 simulações de uma rede neural com 5 neurônios sigmóides na camada escondida e 1 neurônio linear na camada de saída, em totais precipitados. A linha contínua é o resultado da simulação e a tracejada apresenta os dados observados (continuação)

4.3.3. MODELAGEM DAS REDES NEURAIS

Pelos motivos já descritos no capítulo anterior, todas as redes projetadas neste estudo são do tipo multicamadas, com apenas uma camada escondida, composta de neurônios sigmoidais, e com a camada de saída formada por apenas um neurônio linear.

Conforme o procedimento adotado para a determinação do número de neurônios adequado para compor a camada escondida (Figura 3.8), projetou-se redes semelhantes onde este número variou de 1 a 10 neurônios. A princípio, para a rede projetada para a desagregação da precipitação sazonal, esse número chegou a 15 neurônios, onde se constatou que tal magnitude não se fazia necessária, uma vez que redes com mais de 10 neurônios não apresentavam bons resultados na validação, caracterizando sempre a ocorrência de superajustamento.

As dez simulações realizadas com o objetivo de verificar a estabilidade de cada estrutura de rede selecionada pelo desempenho resultaram num conjunto de 10 redes, proporcionando o processo de seleção da melhor estrutura através dos critérios adotados, de estabilidade, de melhor coeficiente de correlação, análise gráfica e do princípio da parcimônia.

4.3.4. NÍVEL SAZONAL

As estruturas de rede que numa primeira seleção apresentaram os melhores resultados para a desagregação em escala sazonal, foram simuladas dez vezes, resultando em redes que estão apresentadas na Tabela 4.3.

A estrutura de rede neural escolhida para realizar a desagregação espacial da precipitação para a bacia do Rio Piancó, em escala sazonal, foi a de 4 neurônios na camada escondida.

Tabela 4.3 – Coeficientes de correlação obtidos nas simulações das estruturas de rede selecionadas para a verificação da estabilidade do algoritmo de treinamento, no nível sazonal

Número de neurônios na camada escondida											
3 neurônios			4 neurônios			5 neurônios			6 neurônios		
Rede	Treina- mento	Validação	Rede	Treina- mento	Validação	Rede	Treina- mento	Validação	Rede	Treina- mento	Validação
1	0,96	0,86	11	0,97	0,87	21	0,98	0,86	31	0,97	-0,34
2	0,96	0,84	12	0,96	0,83	22	0,97	0,85	32	0,98	0,86
3	0,96	0,86	13	0,97	0,87	23	0,97	0,86	33	0,98	0,84
4	0,96	0,86	14	0,97	0,87	24	0,97	0,87	34	0,96	-0,34
5	0,96	0,86	15	0,97	0,87	25	0,97	-0,42	35	0,97	-0,42
6	0,96	0,86	16	0,96	0,83	26	0,98	0,86	36	0,97	-0,39
7	0,96	0,84	17	0,96	0,83	27	0,97	0,85	37	0,98	0,87
8	0,96	0,86	18	0,97	0,87	28	0,97	0,86	38	0,97	-0,34
9	0,96	0,86	19	0,97	0,87	29	0,97	0,87	39	0,98	0,86
10	0,96	0,86	20	0,97	0,87	30	0,97	-0,42	40	0,97	-0,32

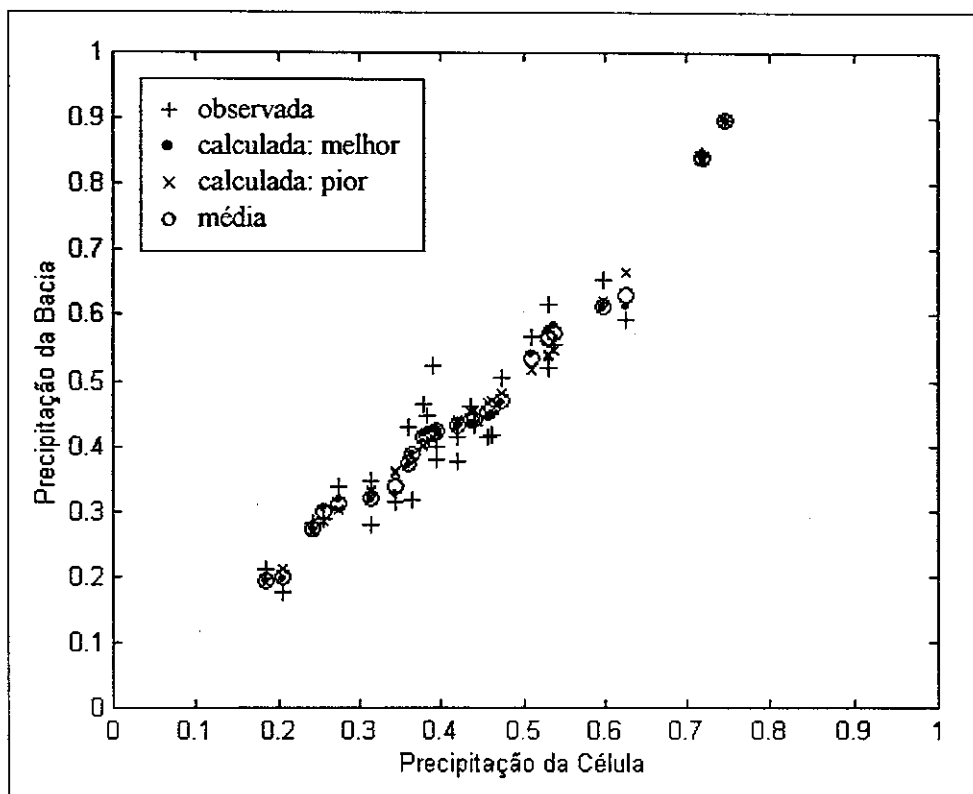
Os valores fornecidos por essas redes, e a média calculada entre esses valores, estão mostrados na Figura 4.19, juntamente com os dados observados da bacia. Como algumas redes mostraram o mesmo comportamento, mas diferiram de outras, manteve-se o mesmo símbolo no gráfico para as primeiras, e atribuiu-se símbolos diferentes às últimas.

A normalização dos dados foi realizada dividindo-se cada valor utilizado por 1200, que é aproximadamente o máximo valor registrado na série.

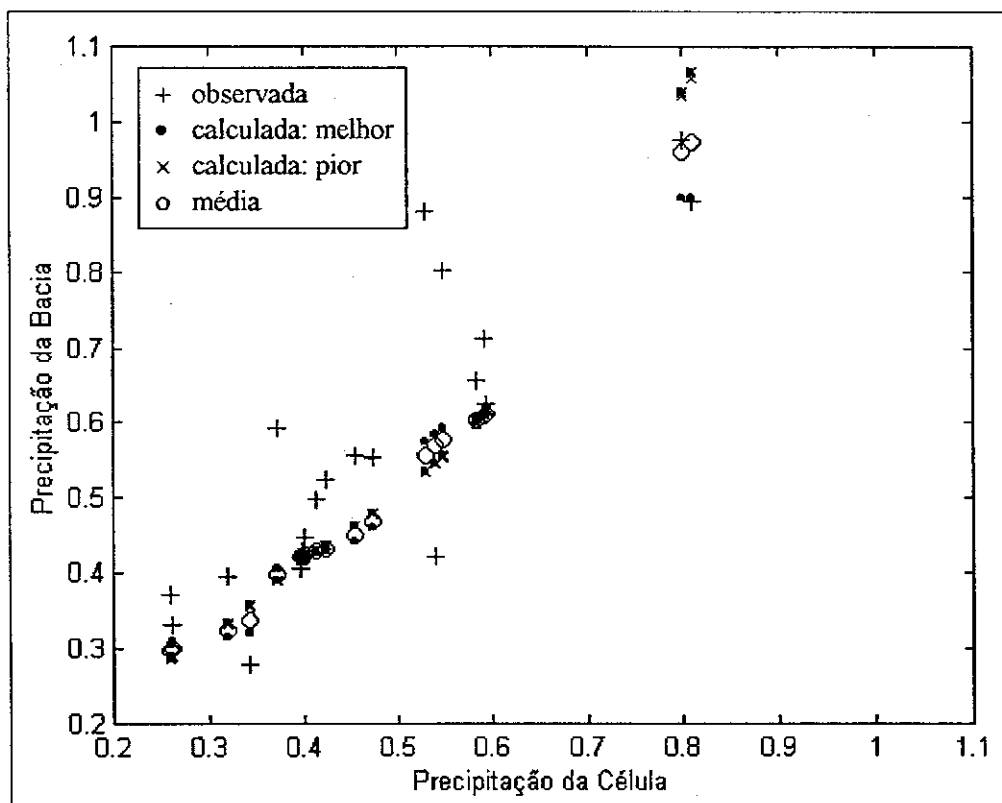
Como se pode verificar na Tabela 4.3, sete redes forneceram resultados semelhantes, obtendo os mesmos coeficientes no treinamento e na validação, e por isso, apresentando dados sobrepostos (Figura 4.19). Portanto, qualquer uma dessas sete redes poderia ser julgada adequada para realizar a desagregação a nível sazonal. A Figura 4.20 mostra o comportamento, no treinamento e na validação, da Rede 11 – escolhida ao acaso entre as sete redes – em comparação aos dados observados. Nesta figura, pode se observar que a rede selecionada consegue capturar bem o comportamento não linear da precipitação durante o treinamento, conseguindo uma alta correlação. No entanto, devido aos já citados anos anômalos (1971, 1977-79), essa correlação cai na fase de validação.

Além da opção de se eleger uma única rede para a desagregação espacial da precipitação, uma outra alternativa seria a utilização de todas as dez redes para se obter uma faixa de valores previstos, estabelecendo um intervalo onde seria maior a possibilidade de acerto nas previsões.

A Figura 4.21 possibilita a comparação entre os dados sazonais observados e os calculados pela Rede 11 ao longo de toda a série (1939-88).

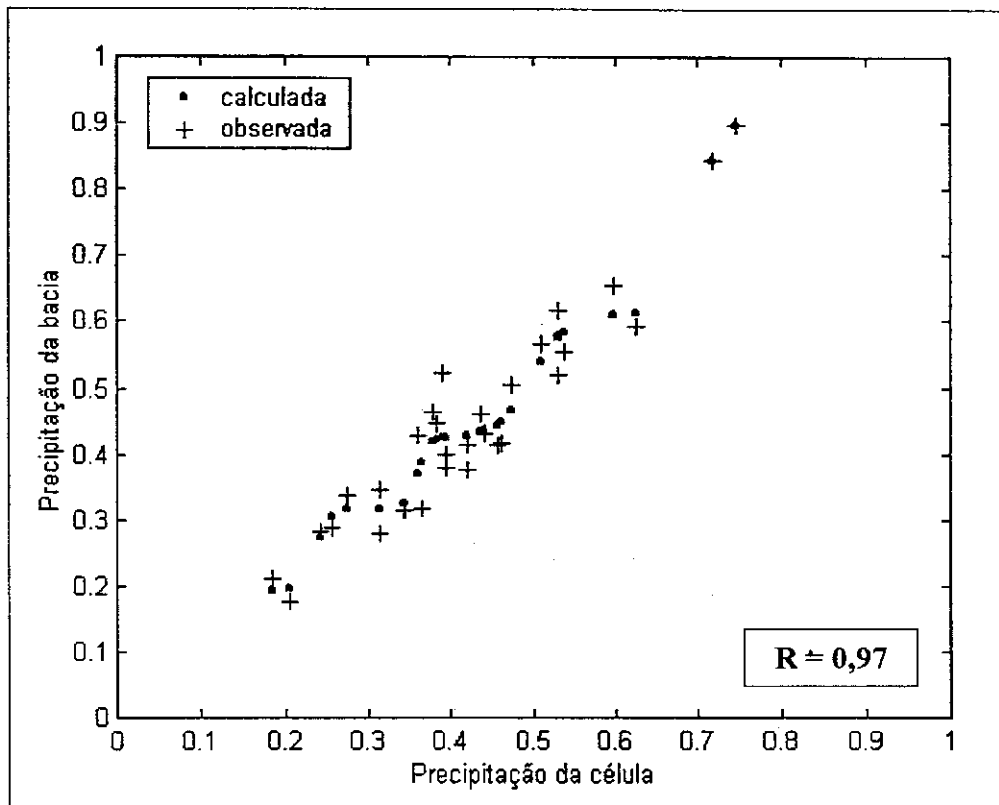


a) Treinamento

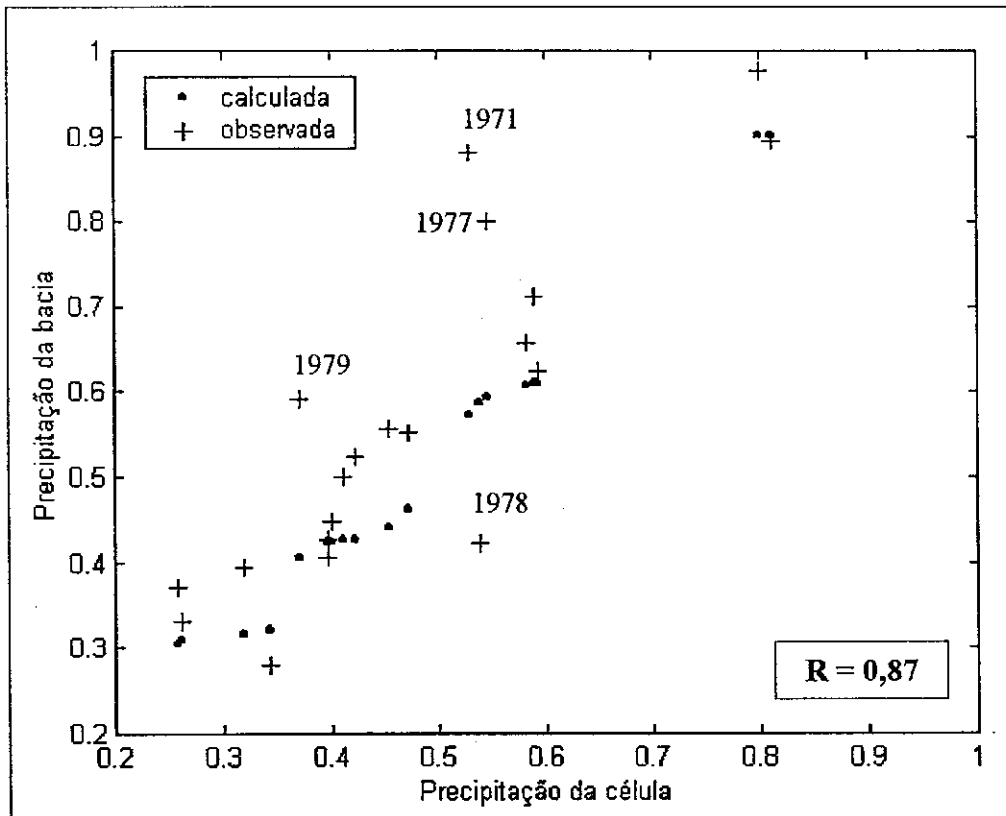


b) Validação

Figura 4.19 – Relações entre as séries normalizadas sazonais da célula e da bacia, para o conjunto de dez redes com 4 neurônios na camada escondida, no período de a) treinamento; e de b) validação

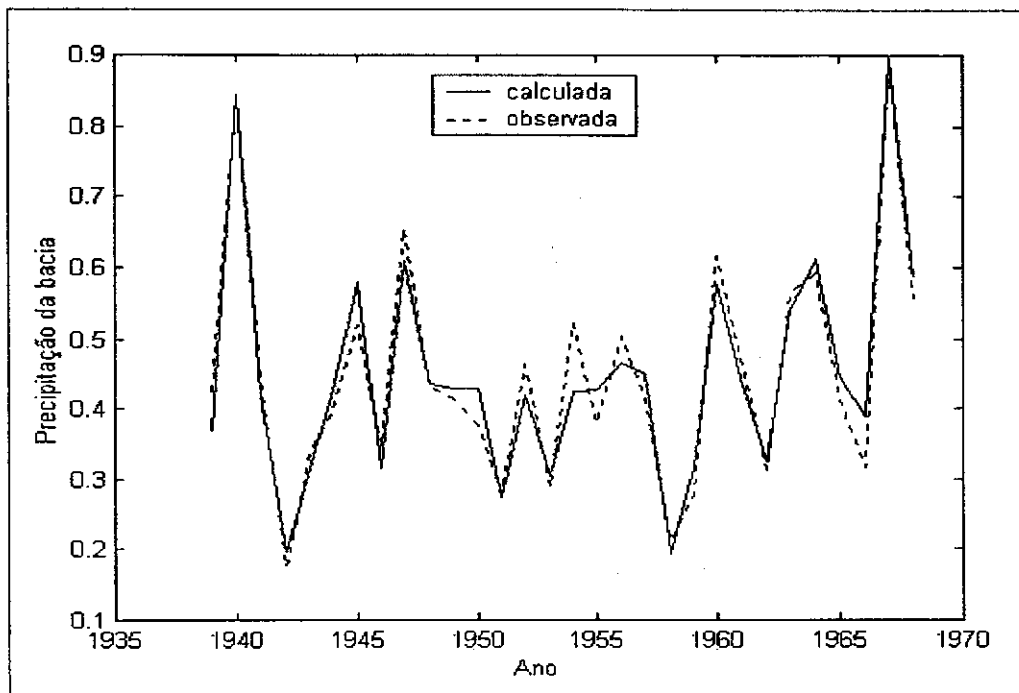


a) treinamento

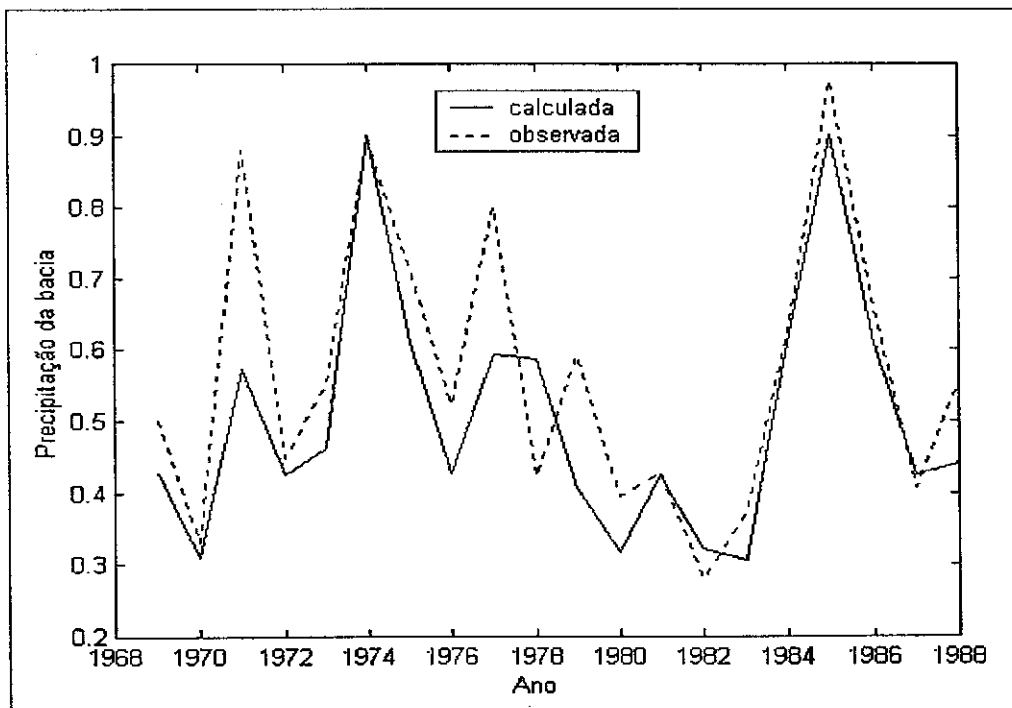


b) validação

Figura 4.20 – Relações resultantes da Rede 11 entre as séries normalizadas sazonais da célula e da bacia, no período de: a) treinamento; e de b) validação



a) treinamento



b) validação

Figura 4.21 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, nos períodos de: a) treinamento e b) validação

4.3.5. Nível Mensal

Entre as séries mensais, observou-se que o mês mais chuvoso na região da bacia do Piancó é o mês de Março, e o que apresenta os menores valores de precipitação é o mês de Maio. No procedimento adotado para a normalização, conforme descrito na seção 4.3.2, cada dado da série foi dividido pelo máximo valor (ou próximo dele) registrado na mesma. Diante disto, as séries foram normalizadas com a utilização de valores diferentes, que estão indicados na Tabela 4.4.

Tabela 4.4 – Valores pelos quais as séries mensais foram normalizadas

Fevereiro	Março	Abril	Maio
400	550	400	240

As estruturas de rede que numa primeira seleção apresentaram os melhores resultados para a desagregação espacial em escala mensal foram simuladas dez vezes, resultando em redes que estão apresentadas na Tabela 4.5. Destas, aquelas que tiveram suas estruturas consideradas como adequadas para realizar a desagregação espacial da precipitação nos meses de fevereiro a maio, para a bacia do Rio Piancó, estão expostas na Tabela 4.6, com os coeficientes de correlação obtidos das redes resultantes das dez simulações, nas fases de treinamento e de validação.

Tabela 4.5 – Coeficientes de correlação obtidos nas simulações das estruturas de rede selecionadas, para os meses de fevereiro a maio

FEVEREIRO								
1 neurônio			2 neurônios			3 neurônios		
Rede	Treinamento	Validação	Rede	Treinamento	Validação	Rede	Treinamento	Validação
1	0,88	0,83	11	0,92	0,81	21	0,90	-0,53
2	0,88	0,83	12	0,90	0,85	22	0,91	0,86
3	0,88	0,83	13	0,90	0,85	23	0,91	0,82
4	0,88	0,83	14	0,90	0,85	24	0,93	0,81
5	0,88	0,83	15	0,90	0,85	25	0,91	0,82
6	0,88	0,83	16	0,90	0,85	26	0,91	0,83
7	0,88	0,83	17	0,90	0,85	27	0,91	0,82
8	0,88	0,83	18	0,90	0,85	28	0,91	0,82
9	0,88	0,83	19	0,90	0,85	29	0,91	0,83
10	0,88	0,83	20	0,90	0,85	30	0,91	0,83

Tabela 4.5 – Coeficientes de correlação obtidos nas simulações das estruturas de rede selecionadas, para os meses de fevereiro a maio (continuação)

MARÇO											
1 neurônio			2 neurônios			3 neurônios			4 neurônios		
Rede	Treina-mento	Validação	Rede	Treina-mento	Validação	Rede	Treina-mento	Validação	Rede	Treina-mento	Validação
1	0,96	0,71	11	0,96	0,71	21	0,97	0,78	31	0,96	0,72
2	0,96	0,71	12	0,96	0,71	22	0,96	0,71	32	0,97	0,72
3	0,96	0,71	13	0,96	0,71	23	0,96	0,71	33	0,97	0,72
4	0,96	0,71	14	0,96	0,71	24	0,97	0,78	34	0,97	0,72
5	0,96	0,71	15	0,96	0,71	25	0,96	0,71	35	0,97	0,72
6	0,96	0,71	16	0,96	0,71	26	0,97	0,78	36	0,97	0,72
7	0,96	0,71	17	0,96	0,71	27	0,97	0,78	37	0,97	0,72
8	0,96	0,71	18	0,96	0,71	28	0,97	0,78	38	0,97	0,72
9	0,96	0,71	19	0,96	0,71	29	0,96	0,71	39	0,97	0,72
10	0,96	0,71	20	0,96	0,71	30	0,96	0,71	40	0,97	0,72
ABRIL											
3 neurônios			4 neurônios			5 neurônios			6 neurônios		
Rede	Treina-mento	Validação	Rede	Treina-mento	Validação	Rede	Treina-mento	Validação	Rede	Treina-mento	Validação
1	0,93	0,92	11	0,94	0,92	21	0,94	0,92	31	0,95	0,87
2	0,93	0,92	12	0,94	0,92	22	0,94	0,92	32	0,94	0,90
3	0,93	0,92	13	0,94	0,92	23	0,94	0,92	33	0,94	0,88
4	0,92	0,79	14	0,94	0,92	24	0,94	0,91	34	0,94	0,85
5	0,93	0,92	15	0,94	0,92	25	0,94	0,92	35	0,94	0,19
6	0,92	0,79	16	0,94	0,92	26	0,94	0,90	36	0,95	0,86
7	0,93	0,92	17	0,94	0,92	27	0,94	0,91	37	0,95	0,86
8	0,93	0,92	18	0,94	0,92	28	0,94	0,92	38	0,95	0,86
9	0,92	0,79	19	0,94	0,92	29	0,94	0,91	39	0,95	0,89
10	0,93	0,92	20	0,94	0,92	30	0,95	0,89	40	0,95	0,89
MAIO											
1 neurônio			2 neurônios			3 neurônios			4 neurônios		
Rede	Treina-mento	Validação	Rede	Treina-mento	Validação	Rede	Treina-mento	Validação	Rede	Treina-mento	Validação
1	0,93	0,85	11	0,93	0,86	21	0,95	0,84	31	0,95	0,87
2	0,93	0,85	12	0,93	0,86	22	0,93	-0,53	32	0,95	0,87
3	0,93	0,85	13	0,93	0,85	23	0,93	0,84	33	0,95	0,87
4	0,93	0,85	14	0,93	0,86	24	0,95	0,84	34	0,94	0,86
5	0,93	0,85	15	0,92	0,86	25	0,93	-0,51	35	0,95	0,87
6	0,93	0,85	16	0,93	0,86	26	0,95	0,79	36	0,96	0,71
7	0,93	0,85	17	0,93	0,86	27	0,93	-0,53	37	0,93	0,81
8	0,93	0,85	18	0,93	0,85	28	0,93	-0,54	38	0,95	0,88
9	0,93	0,85	19	0,93	0,85	29	0,93	-0,53	39	0,95	0,86
10	0,93	0,85	20	0,93	0,85	30	0,93	-0,55	40	0,95	0,87

Tabela 4.6 – Coeficientes de correlação obtidos das redes julgadas mais adequadas para a desagregação, nos meses de fevereiro a maio

FEVEREIRO			MARÇO			ABRIL			MAIO		
2 neurônios			3 neurônios			4 neurônios			4 neurônios		
Rede	Treina- mento	Validação	Rede	Treina- mento	Validação	Rede	Treina- mento	Validação	Rede	Treina- mento	Validação
11	0,92	0,81	21	0,97	0,78	11	0,94	0,92	31	0,95	0,87
12	0,90	0,85	22	0,96	0,71	12	0,94	0,92	32	0,95	0,87
13	0,90	0,85	23	0,96	0,71	13	0,94	0,92	33	0,95	0,87
14	0,90	0,85	24	0,97	0,78	14	0,94	0,92	34	0,94	0,86
15	0,90	0,85	25	0,96	0,71	15	0,94	0,92	35	0,95	0,87
16	0,90	0,85	26	0,97	0,78	16	0,94	0,92	36	0,96	0,71
17	0,90	0,85	27	0,97	0,78	17	0,94	0,92	37	0,93	0,81
18	0,90	0,85	28	0,97	0,78	18	0,94	0,92	38	0,95	0,88
19	0,90	0,85	29	0,96	0,71	19	0,94	0,92	39	0,95	0,86
20	0,90	0,85	30	0,96	0,71	20	0,94	0,92	40	0,95	0,87

Os valores fornecidos durante a fase de treinamento pelas redes referentes ao mês de fevereiro, março, abril e maio, e as médias calculadas entre esses valores, estão mostrados na Figura 4.22, juntamente com os dados observados da bacia. Os resultados da fase de validação estão expostos na Figura 4.23.

Para o mês de Fevereiro, apenas a Rede 11 apresentou um comportamento diferente das demais, atribuindo-se a ela, portanto, um símbolo diferente das outras nas Figuras 4.22 e 4.23. Nestas figuras e na Tabela 4.6, percebe-se que nove redes forneceram resultados semelhantes, obtendo os mesmos coeficientes no treinamento e na validação. Portanto, qualquer uma das nove poderia ser escolhida como a rede de melhor desempenho para a desagregação no mês de Fevereiro. A Figura 4.24 mostra o comportamento, no treinamento e na validação, da Rede 13 – escolhida ao acaso entre as nove redes – em comparação aos dados observados. Nesta figura (como em todas as outras redes projetadas a nível mensal e sazonal), pode se observar que a rede selecionada consegue capturar bem o comportamento não linear da precipitação durante o treinamento, conseguindo uma alta correlação. No entanto, devido aos já citados anos anômalos (1971, 1977, 1978 e 1979), essa correlação cai na fase de validação.

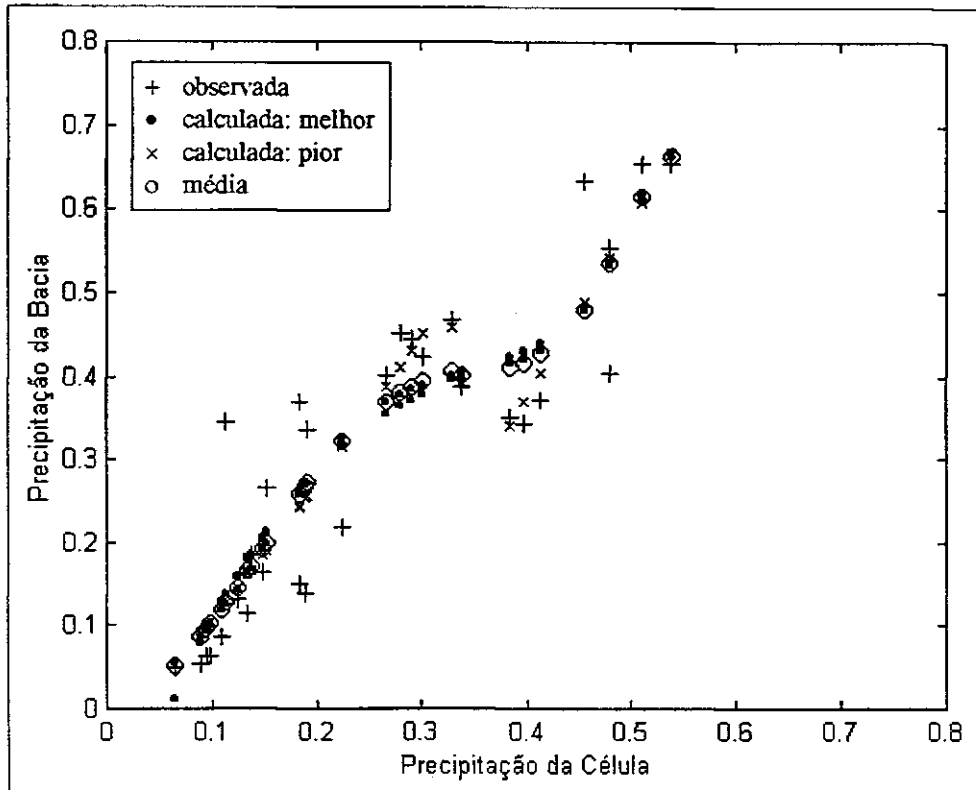
Para o mês de Março, atribuiu-se às redes destacadas na Tabela 4.6 um símbolo diferente das demais. Como elas têm o comportamento semelhante e forneceram os mesmos resultados, selecionou-se arbitrariamente a Rede 26 como a mais indicada para a

desagregação espacial da precipitação no mês de Março. A Figura 4.25 mostra o comportamento, no treinamento e na validação, desta rede em comparação aos dados observados. Mais uma vez registrou-se queda no coeficiente de correlação da validação em relação ao coeficiente obtido no treinamento devido aos anos anômalos.

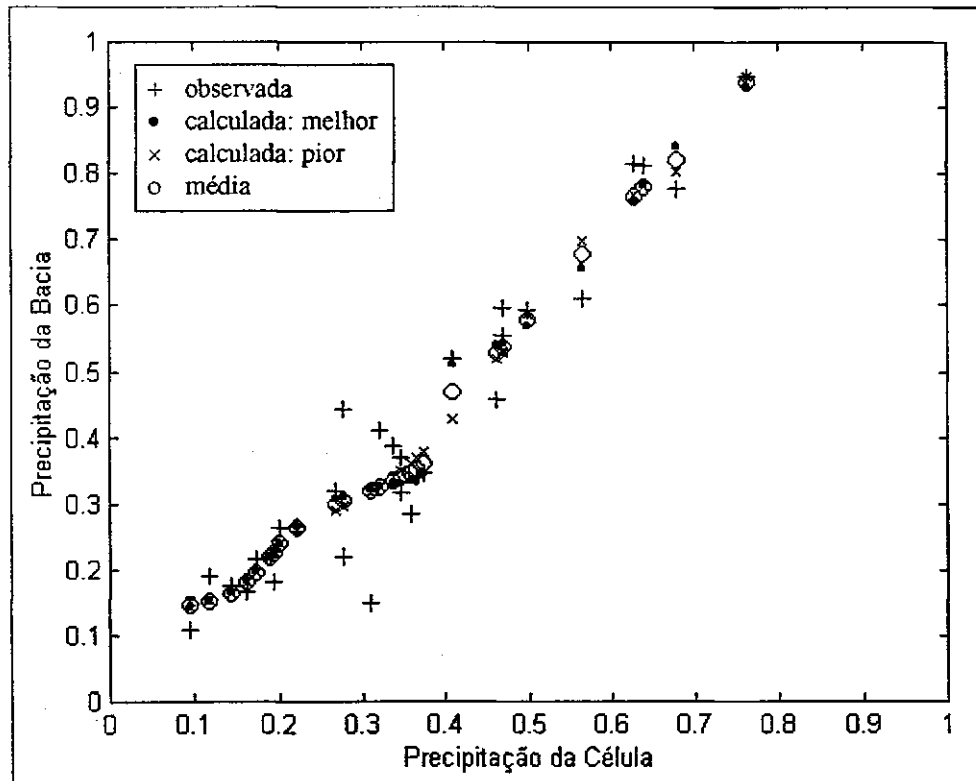
Com relação ao mês de Abril, através da Tabela 4.6 e das Figuras 4.22 e 4.23, observou-se que todas as redes se comportaram de maneira semelhante, indicando se tratar de uma estrutura de rede altamente estável. Apontou-se, então, a Rede 14 para representar o conjunto resultante, recebendo a denominação de a rede neural mais adequada para a desagregação espacial da precipitação no mês de Abril. A Figura 4.26 mostra o comportamento, no treinamento e na validação, desta rede em comparação aos dados observados.

Como das redes que compõem o conjunto referente ao mês de Maio – com exceção da Rede 37, que não foi considerada por motivos que serão expostos adiante – apenas a sexta apresentou um comportamento diferente das demais (Tabela 4.6, Figura 4.22 e Figura 4.23), atribuiu-se a ela, no gráfico, um símbolo diferente das outras. Apesar de apresentarem comportamentos semelhantes, a Rede 38 mostrou um desempenho um pouco melhor que as demais, e por isso foi selecionada para a desagregação espacial da precipitação no mês de Maio. A Figura 4.27 mostra o comportamento, no treinamento e na validação, desta rede em comparação aos dados observados.

As Figuras 4.28, 4.29, 4.30 e 4.31 possibilitam a comparação entre os dados observados e os calculados pelas redes escolhidas para a desagregação espacial da precipitação nos meses de fevereiro, março, abril e maio (respectivamente).

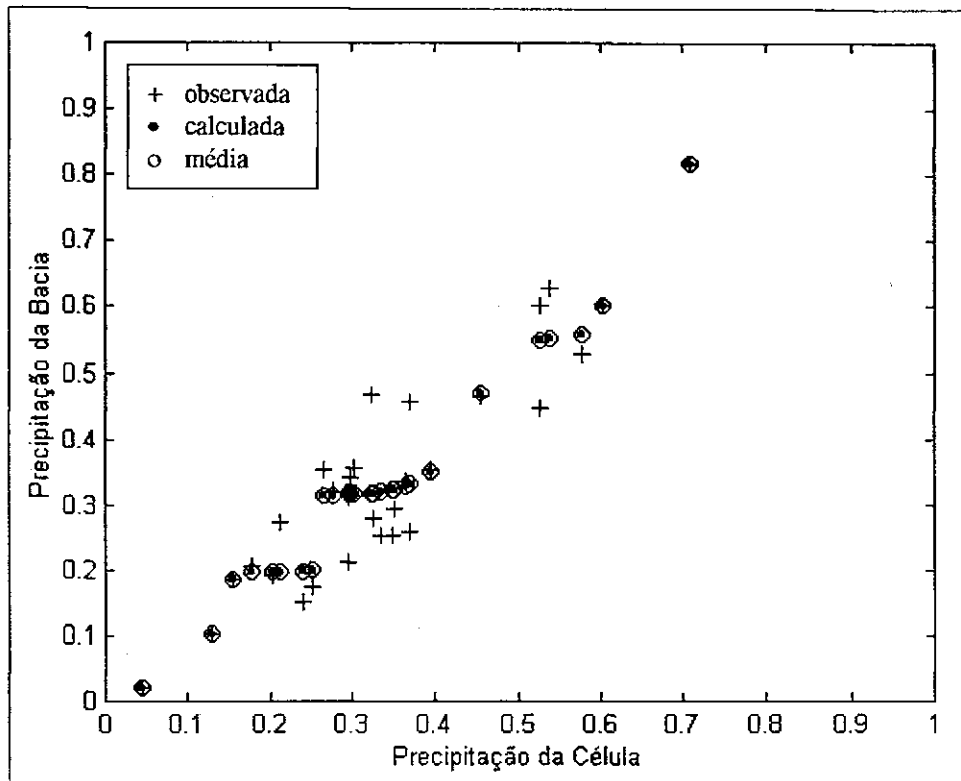


a) fevereiro

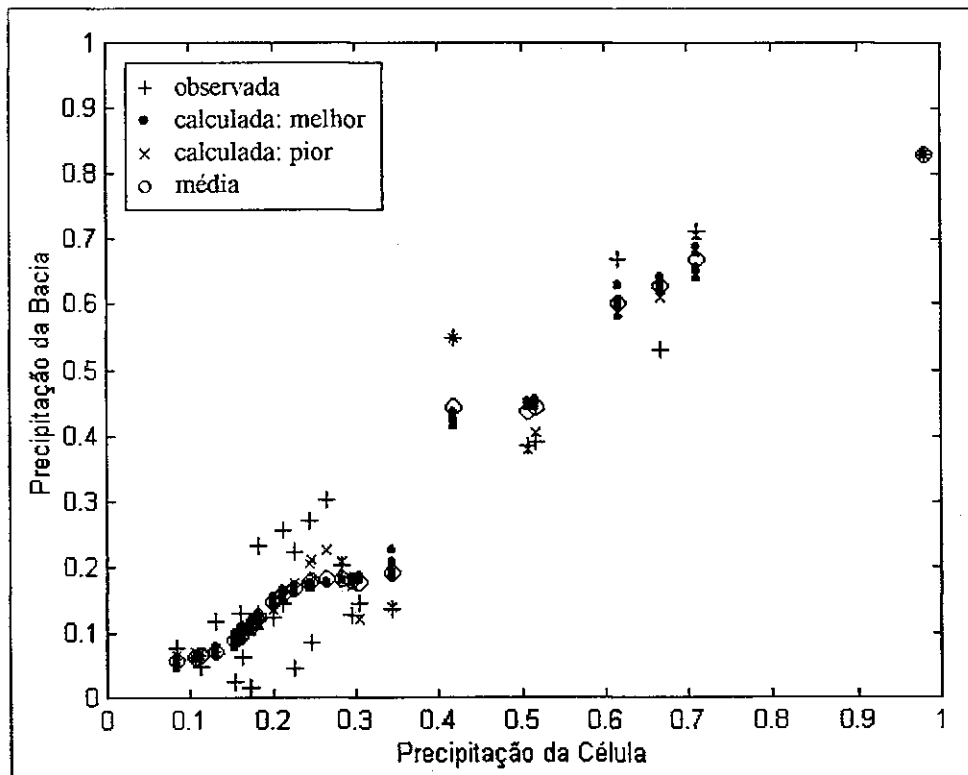


b) março

Figura 4.22 – Relações entre as séries normalizadas da célula e da bacia para o conjunto de dez redes, no período de treinamento (1939-68), referentes aos meses de a) fevereiro; b) março; c) abril; d) maio

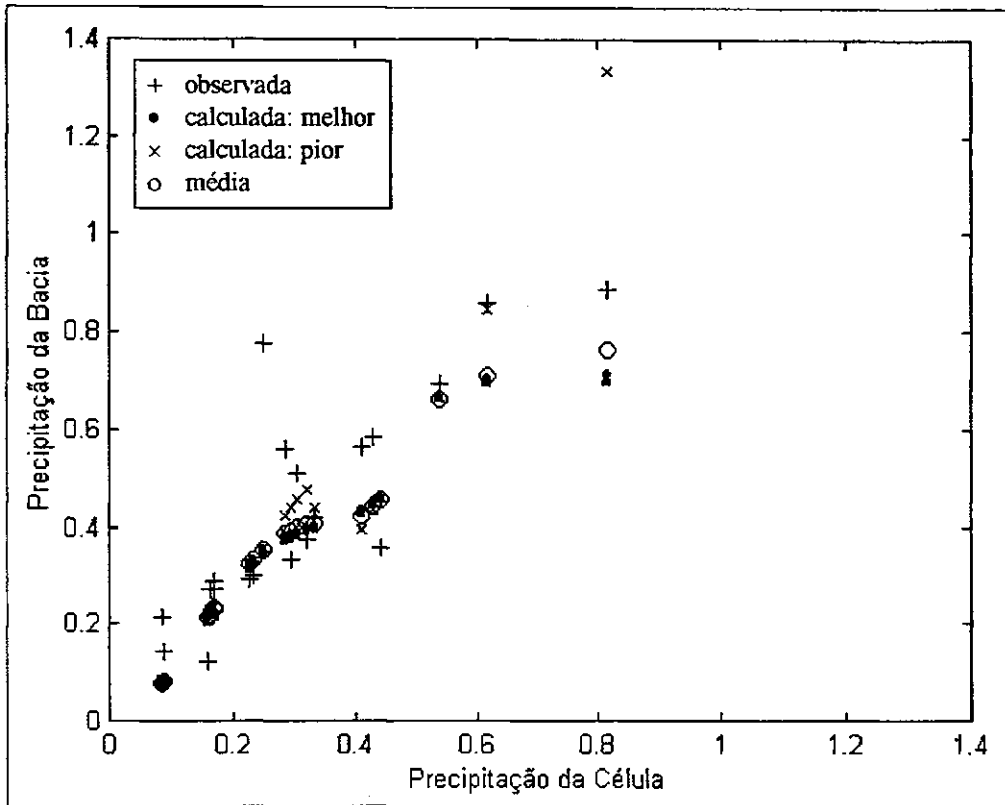


c) abril

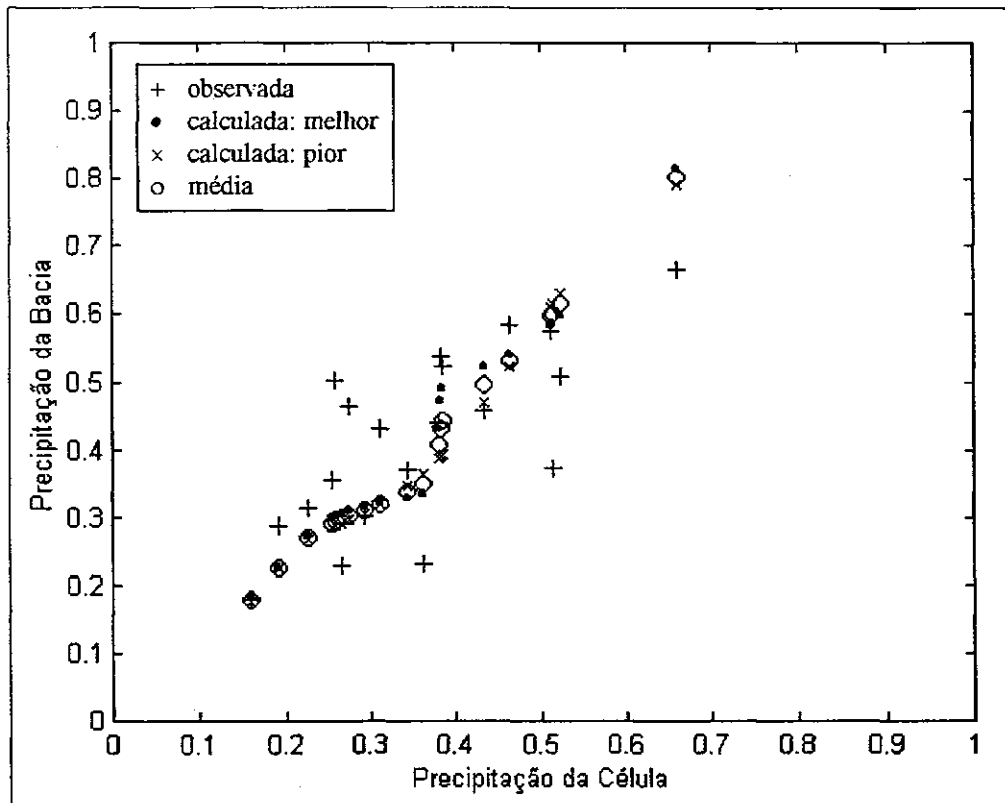


d) maio

Figura 4.22 – Relações entre as séries normalizadas da célula e da bacia para o conjunto de dez redes, no período de treinamento (1939-68), referentes aos meses de a) fevereiro; b) março; c) abril; d) maio (continuação)

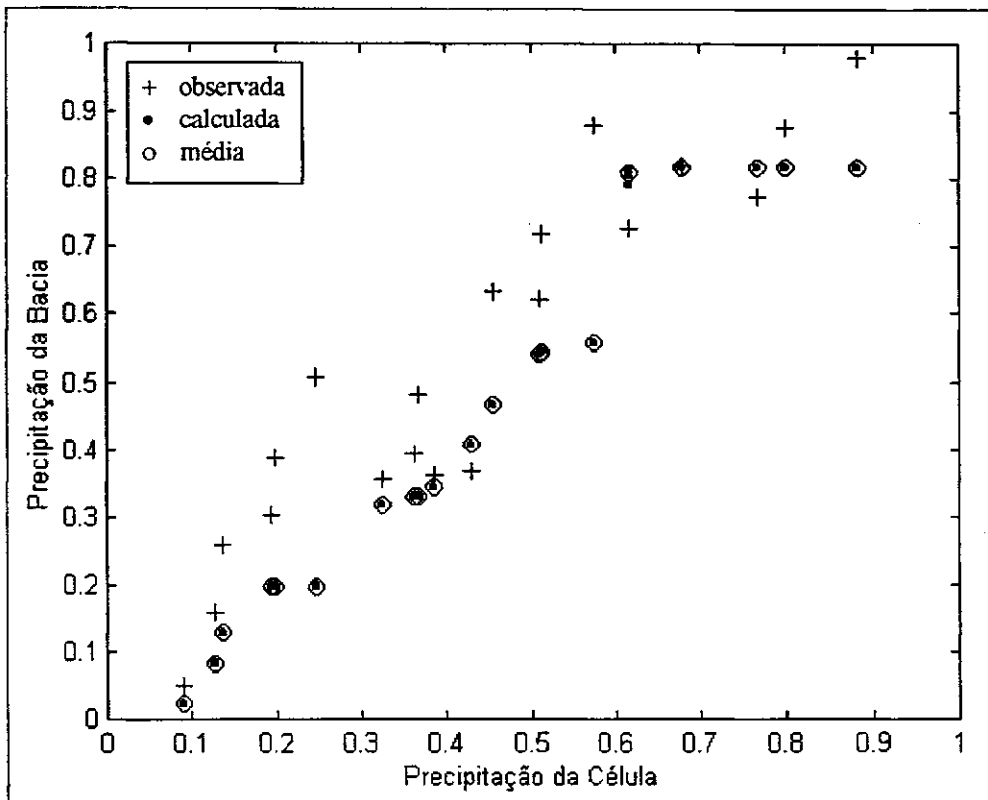


a) fevereiro

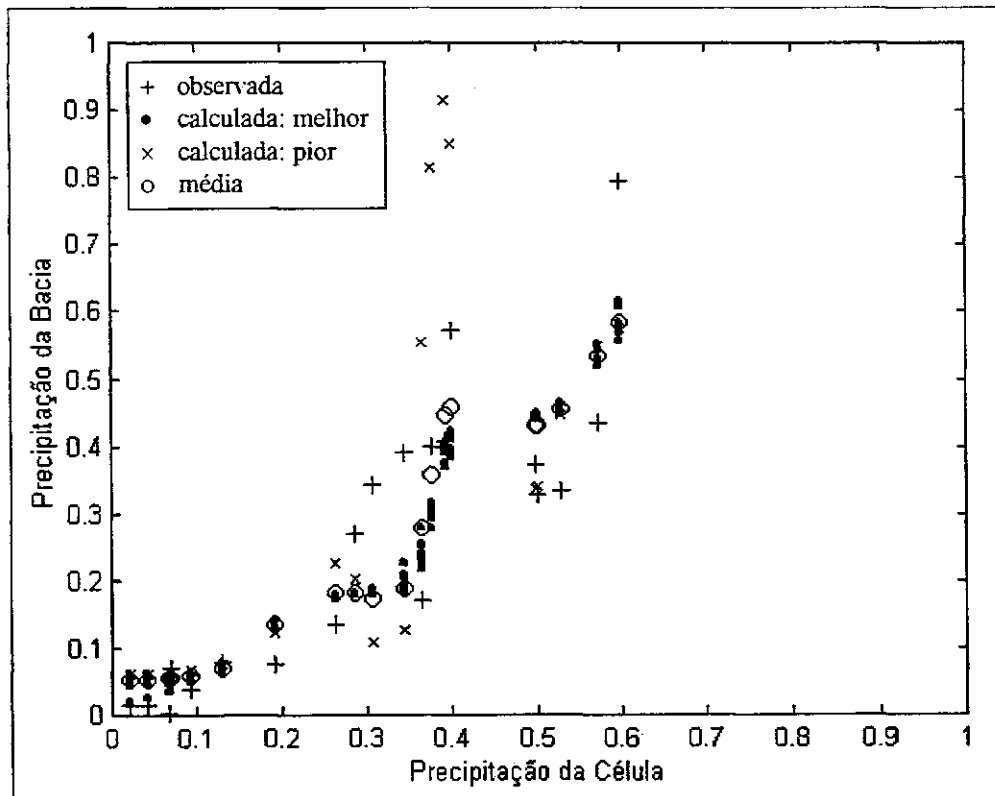


b) março

Figura 4.23 – Relações entre as séries normalizadas da célula e da bacia para o conjunto de dez redes, no período de validação (1969-88), referentes aos meses de a) fevereiro; b) março; c) abril; d) maio

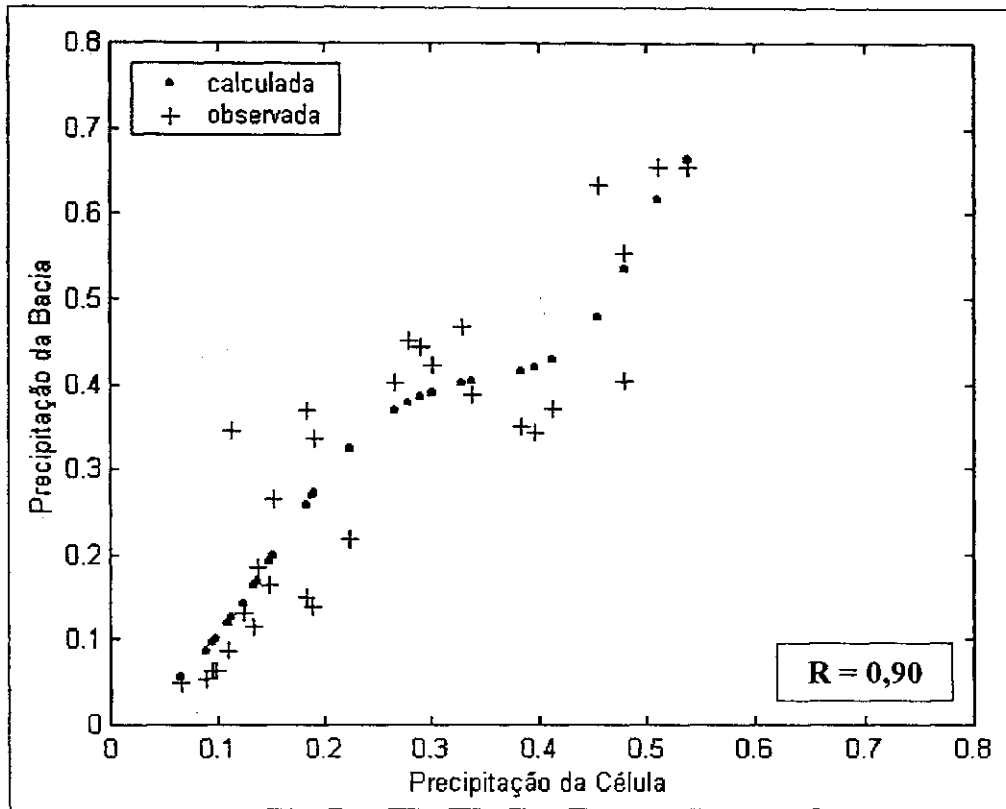


c) abril

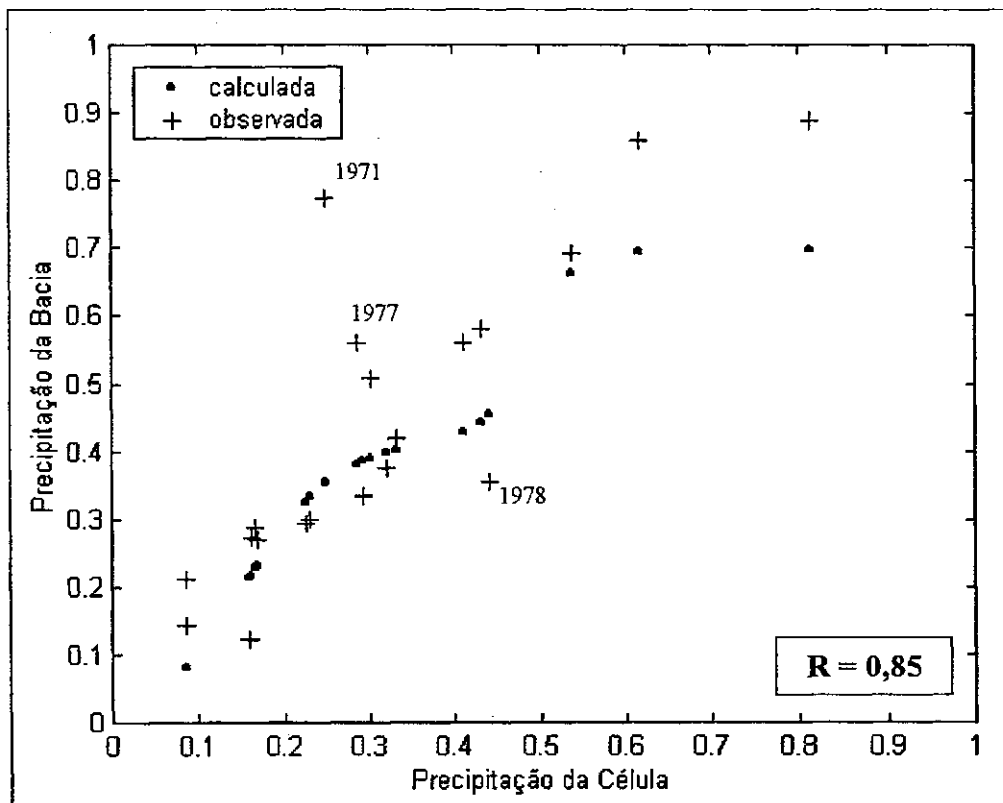


d) maio

Figura 4.23 – Relações entre as séries normalizadas da célula e da bacia para o conjunto de dez redes, no período de validação (1969-88), referentes aos meses de a) fevereiro; b) março; c) abril; d) maio (continuação)

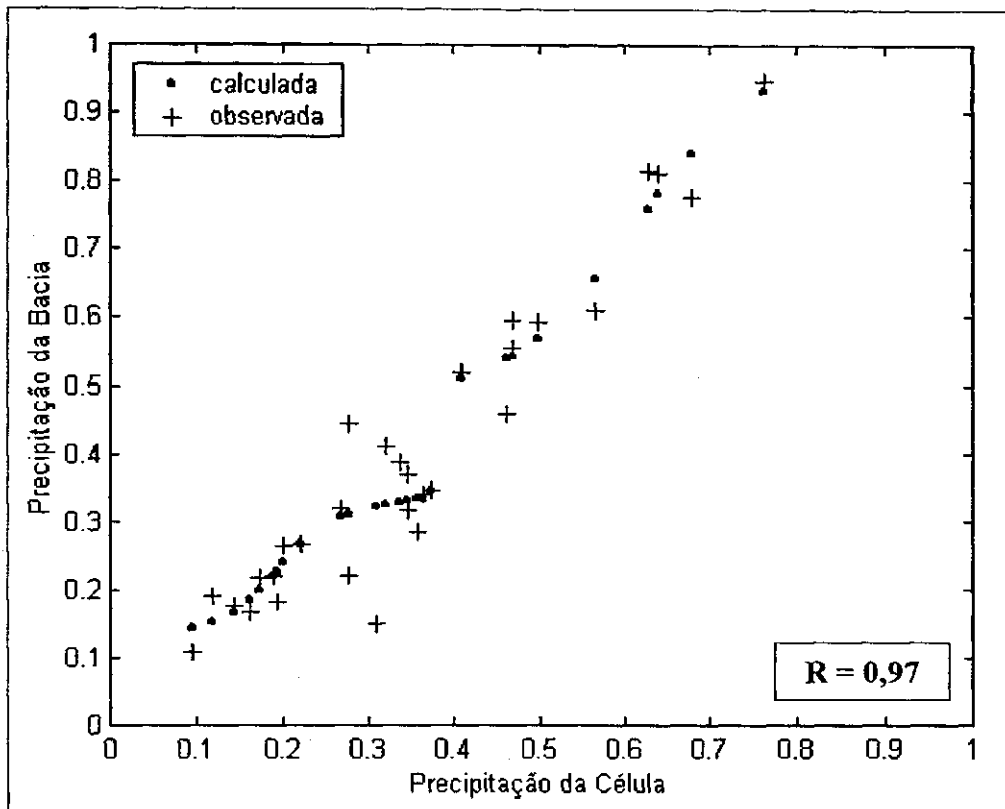


a) treinamento

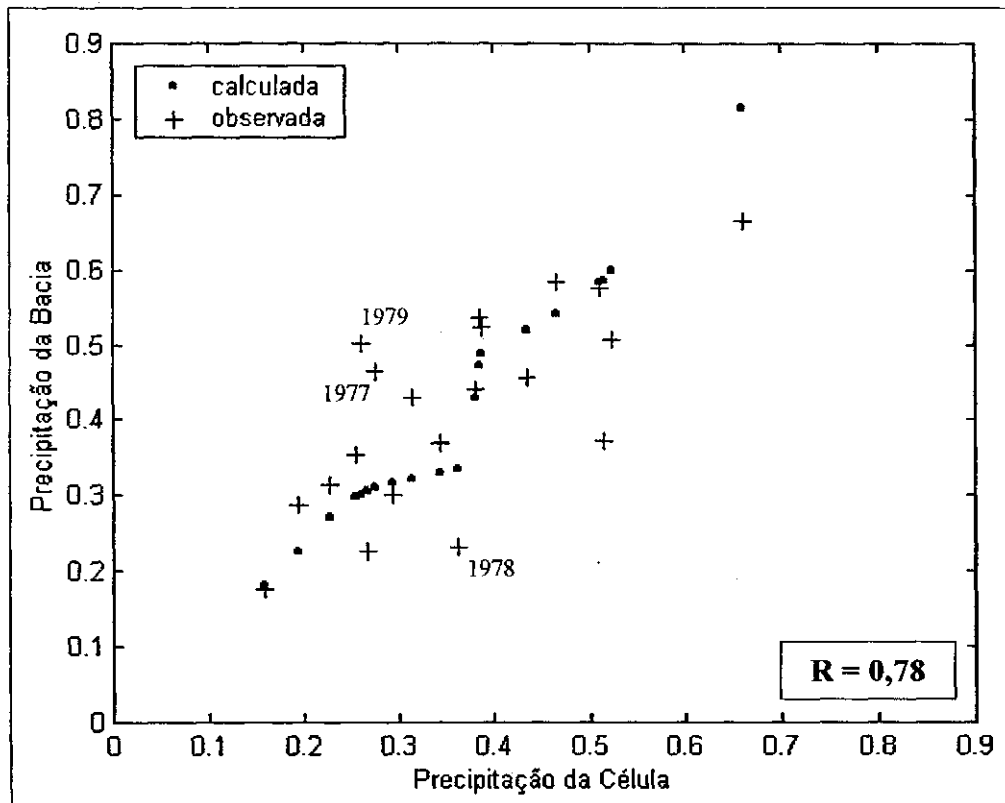


b) validação

Figura 4.24 – Relações resultantes da Rede 13 entre as séries normalizadas da célula e da bacia, referentes ao mês de Fevereiro, no período de a) treinamento; e de b) validação

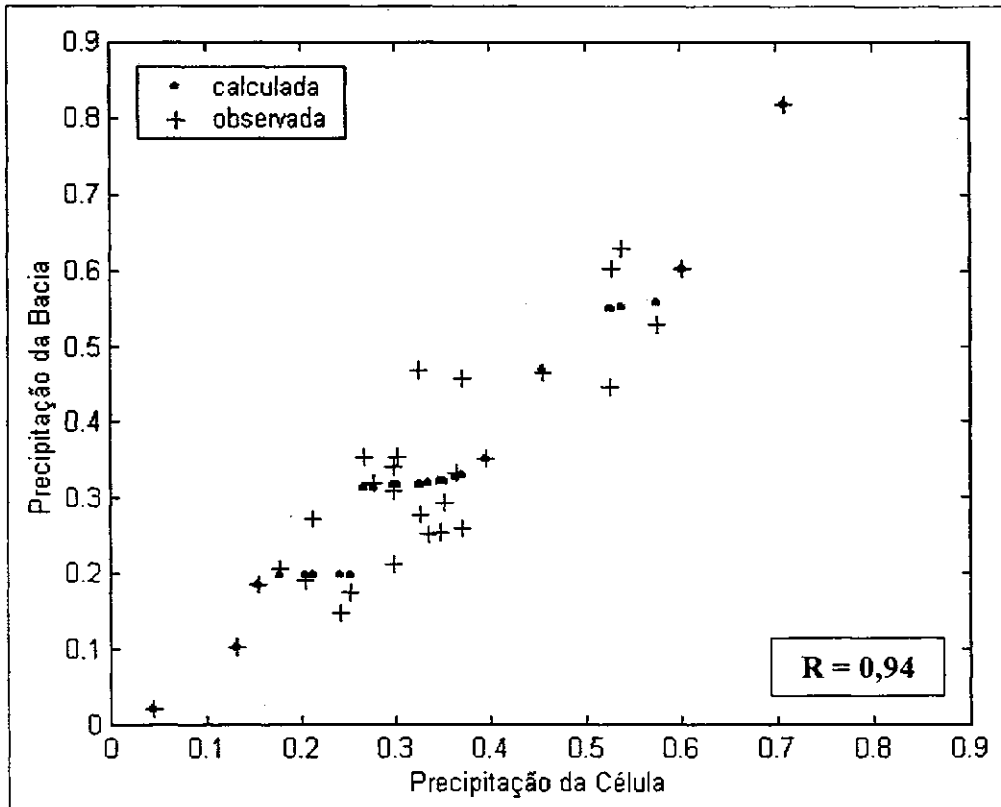


a) treinamento

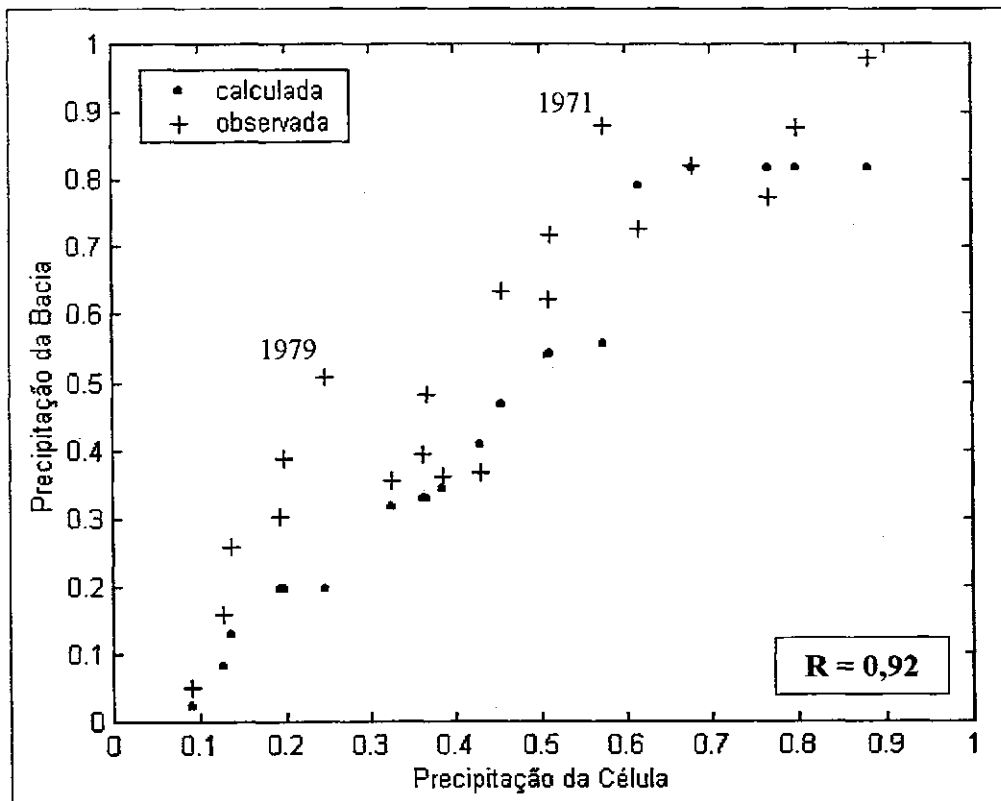


b) validação

Figura 4.25 – Relações resultantes da Rede 26 entre as séries normalizadas da célula e da bacia, referentes ao mês de Março, no período de a) treinamento; e de b) validação

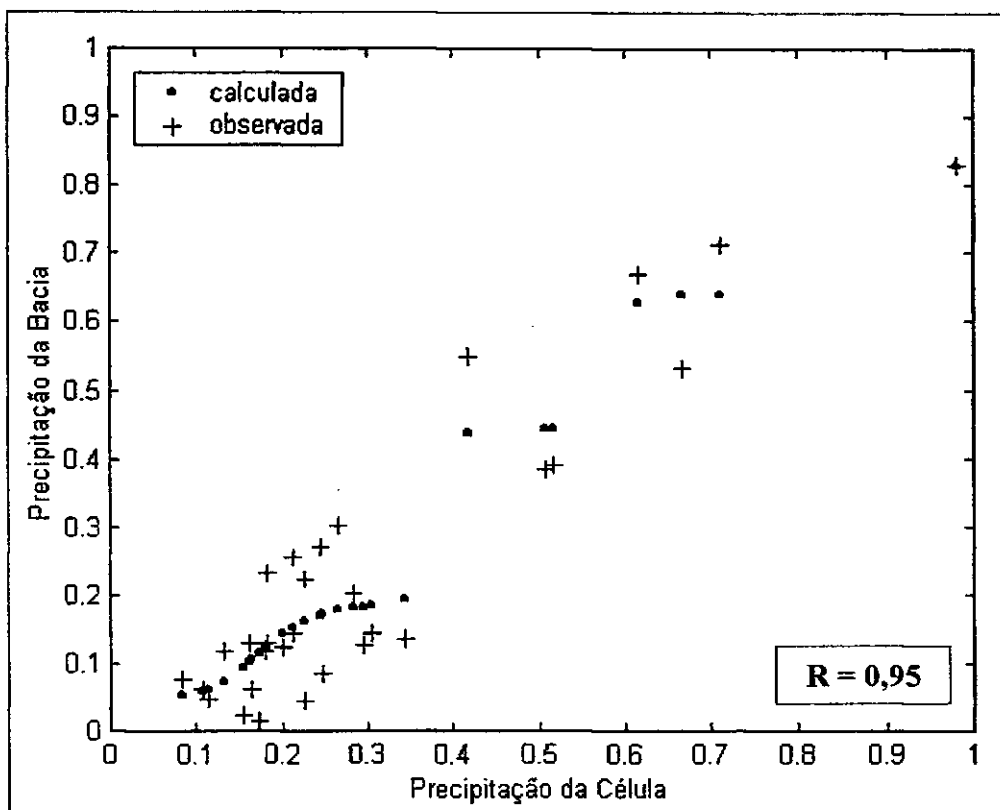


a) treinamento

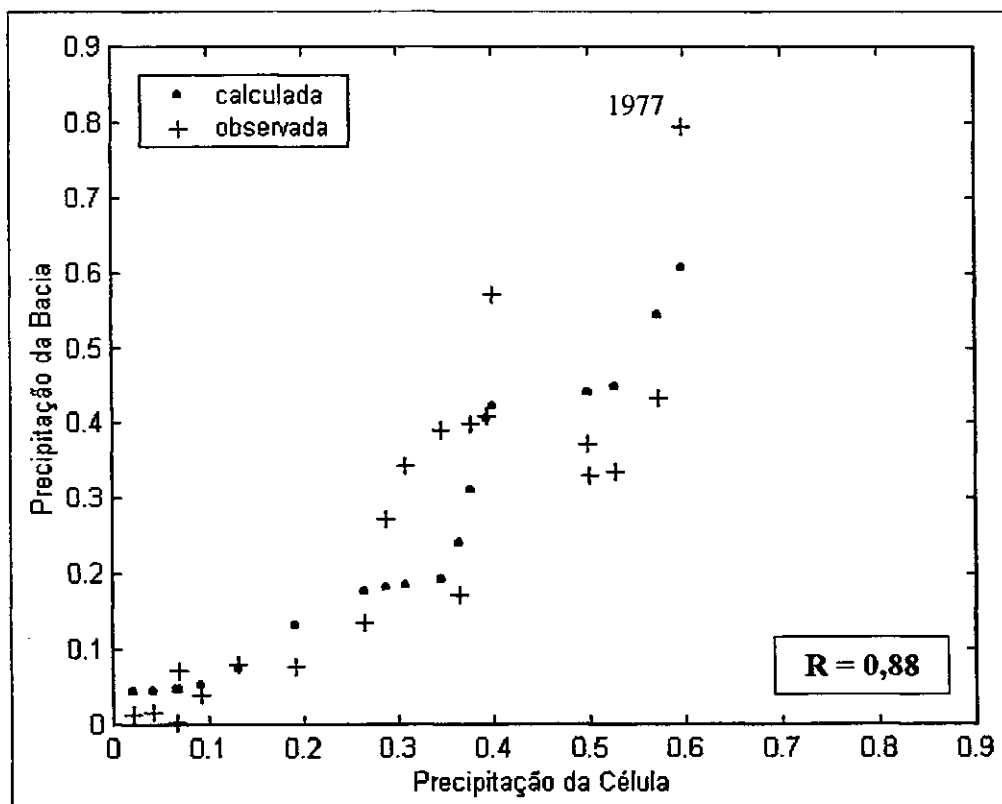


b) validação

Figura 4.26 – Relações resultantes da Rede 14 entre as séries normalizadas da célula e da bacia, referentes ao mês de Abril, no período de: a) treinamento; e b) validação

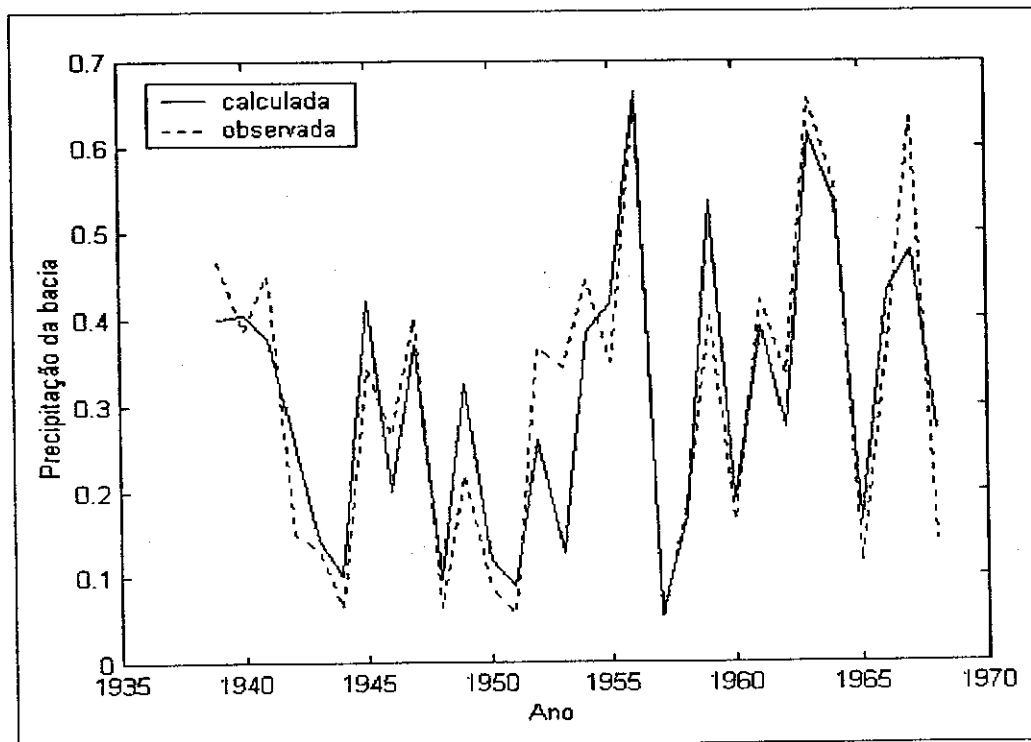


a) treinamento

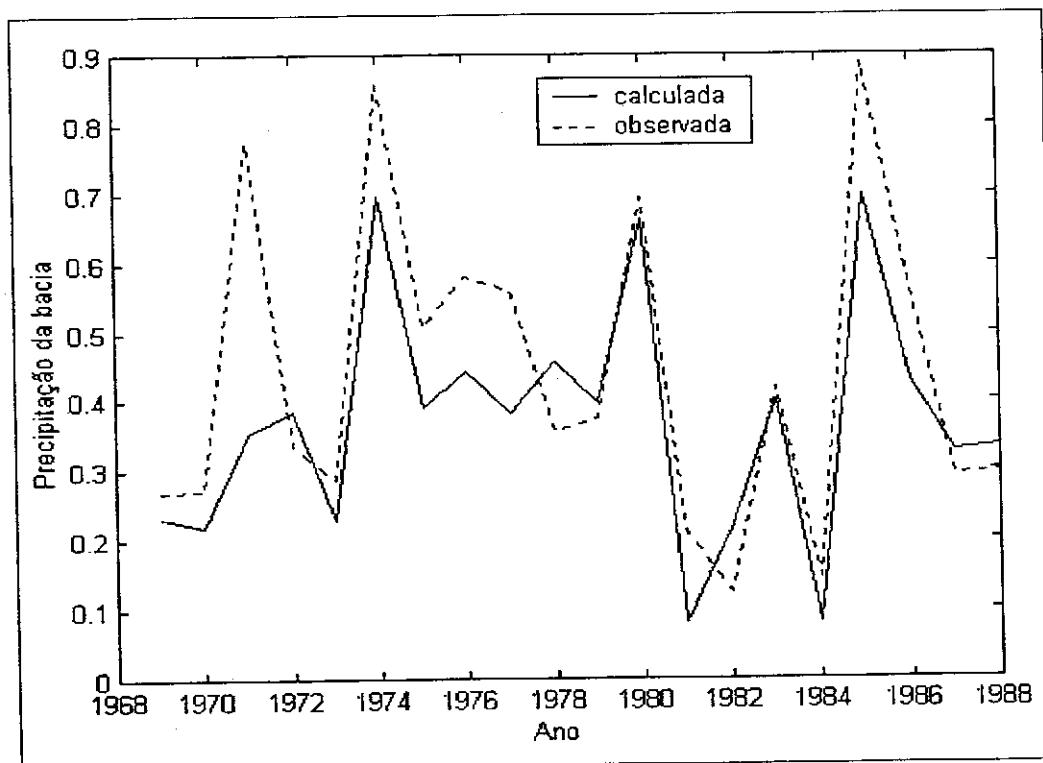


b) validação

Figura 4.27 – Relações resultantes da Rede 38 entre as séries normalizadas da célula e da bacia, referentes ao mês de Maio, no período de: a) treinamento; e b) validação

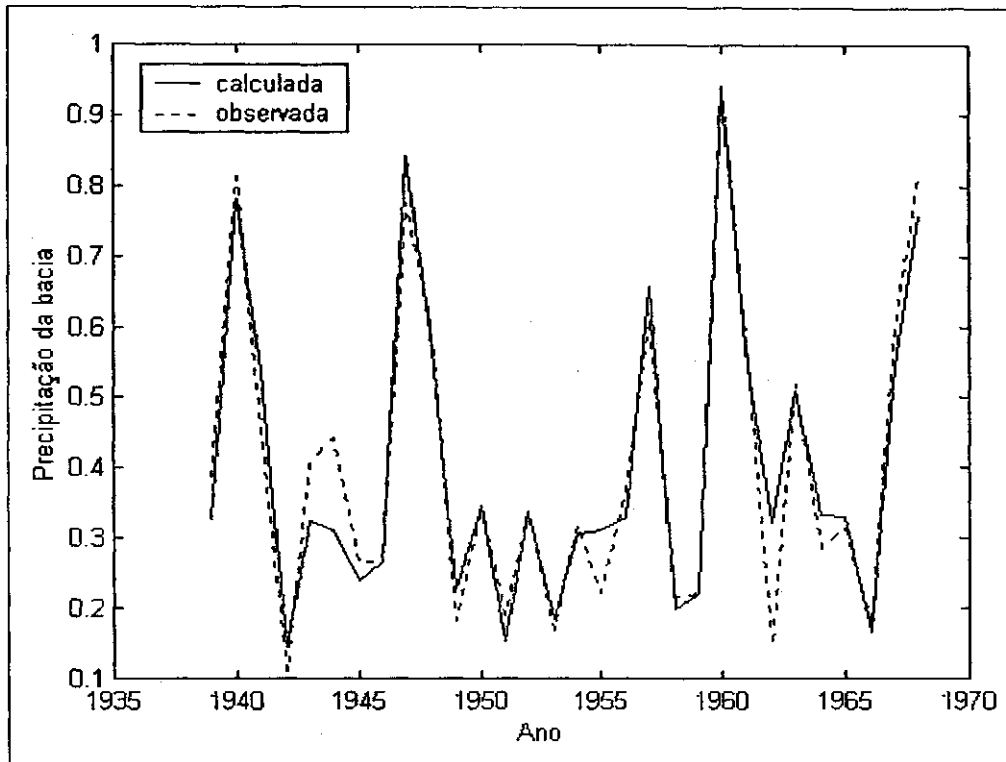


a) treinamento

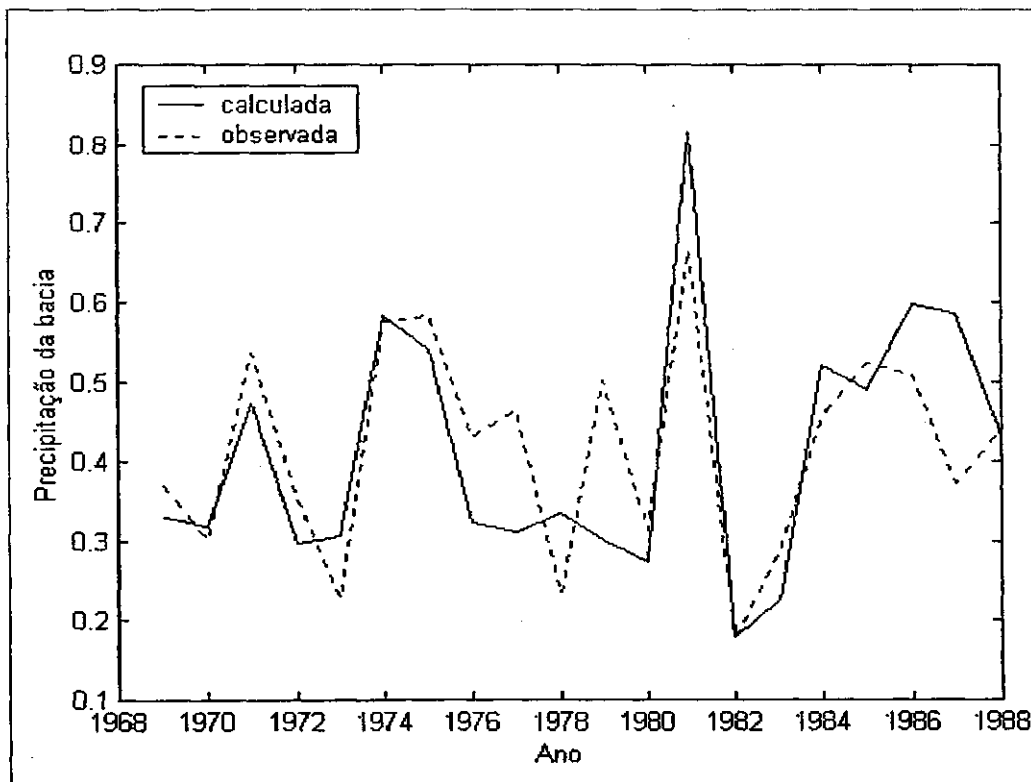


b) validação

Figura 4.28 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, no mês de Fevereiro, nos períodos de: a) treinamento e b) validação

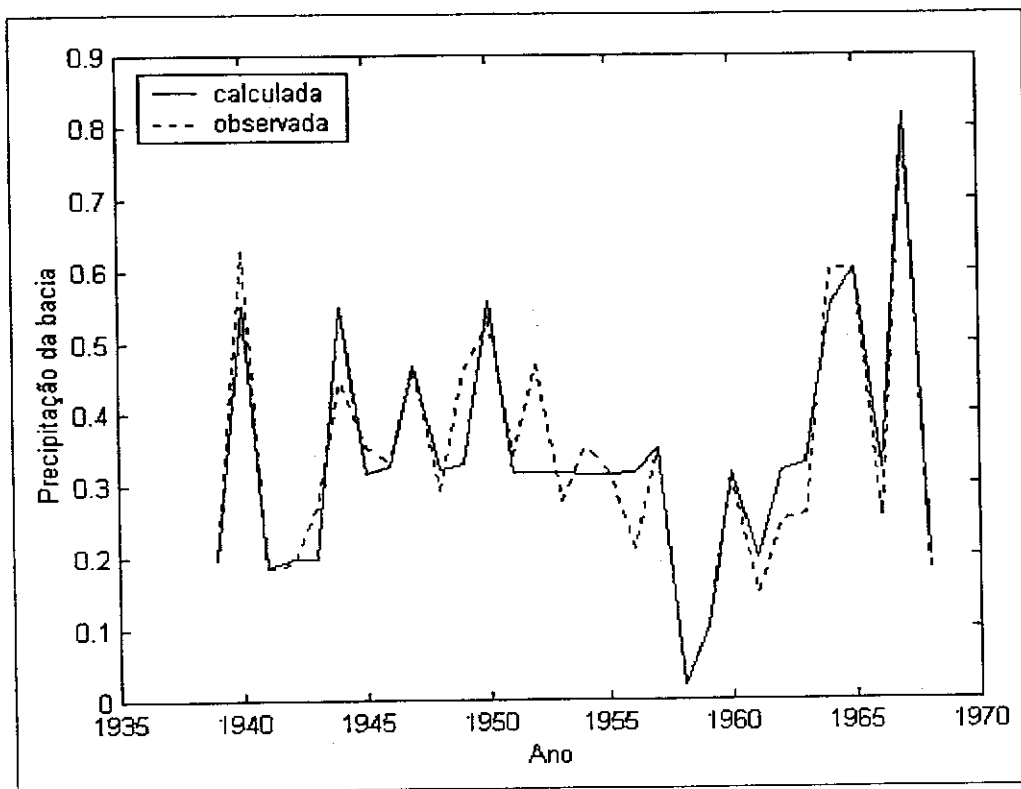


a) treinamento

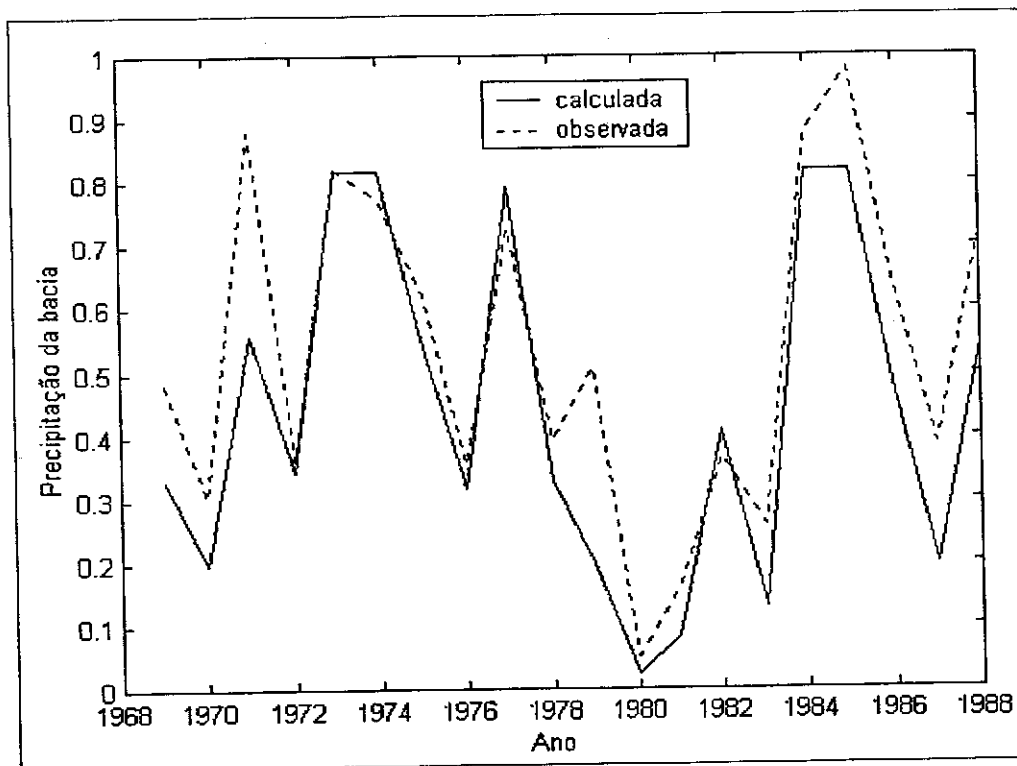


b) validação

Figura 4.29 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, no mês de Março, nos períodos de: a) treinamento e b) validação

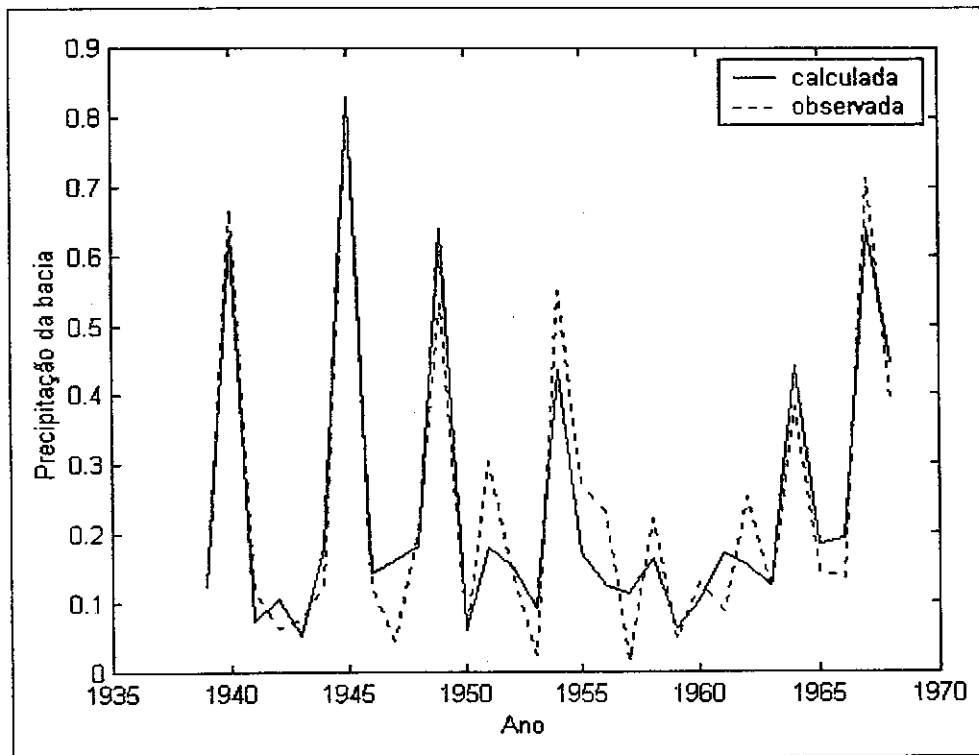


a) treinamento

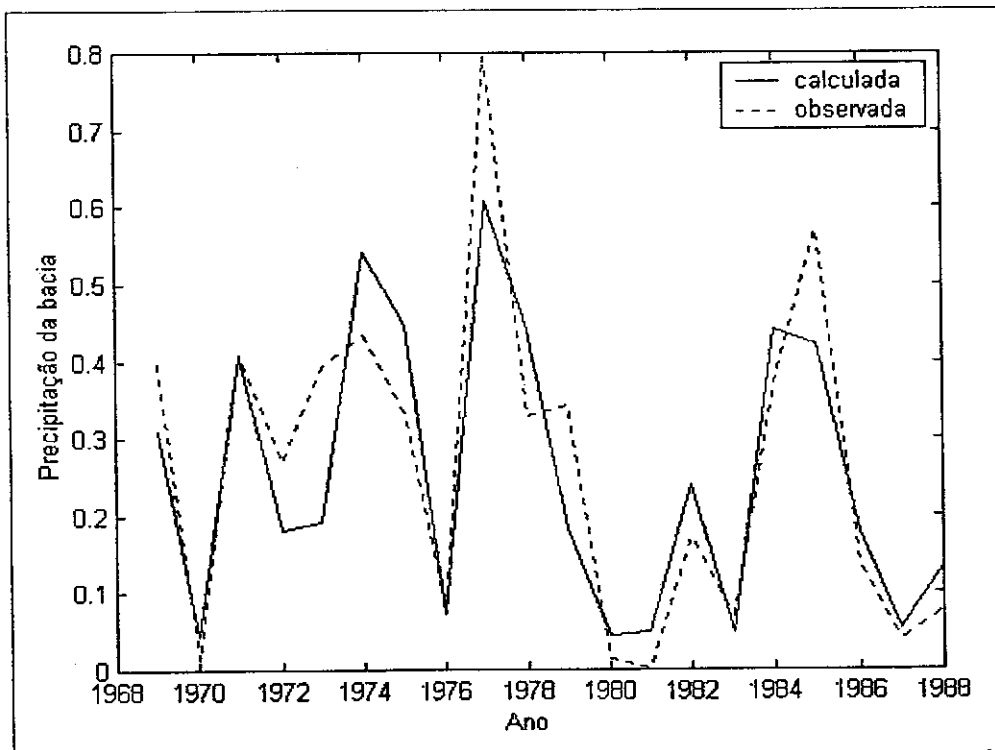


b) validação

Figura 4.30 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, no mês de Abril, nos períodos de: a) treinamento e b) validação



a) treinamento



b) validação

Figura 4.31 – Comparação entre as séries normalizadas observada e calculada pela Rede Neural, para a bacia do Piancó, no mês de Maio, nos períodos de: a) treinamento e b) validação

4.3.6. IMPORTÂNCIA DA ANÁLISE GRÁFICA

O principal motivo de se ter extraído a Rede 37 das Figuras 4.22 e 4.23, que apresentam o conjunto de redes com 4 neurônios na camada escondida, foi para se mostrar isoladamente o seu caso. Os coeficientes de correlação obtidos nas fases de treinamento e de validação dessa rede foram bons e são da mesma ordem de grandeza dos coeficientes das demais (Tabela 4.6). No entanto, apesar disso, observou-se nos gráficos obtidos nessas fases (Figura 4.32), que a rede gerou um mesmo valor para a precipitação na bacia, relacionando-o a vários dados diferentes da célula. Como a rede atribuiu à precipitação um comportamento irreal, não conseguindo um ajuste satisfatório com os dados observados da bacia, não se fez adequada a sua adesão ao conjunto de redes bens treinadas. Este exemplo confirma a importância de se utilizar também a análise gráfica como critério de seleção da rede, e não apenas a verificação dos coeficientes de correlação calculados.

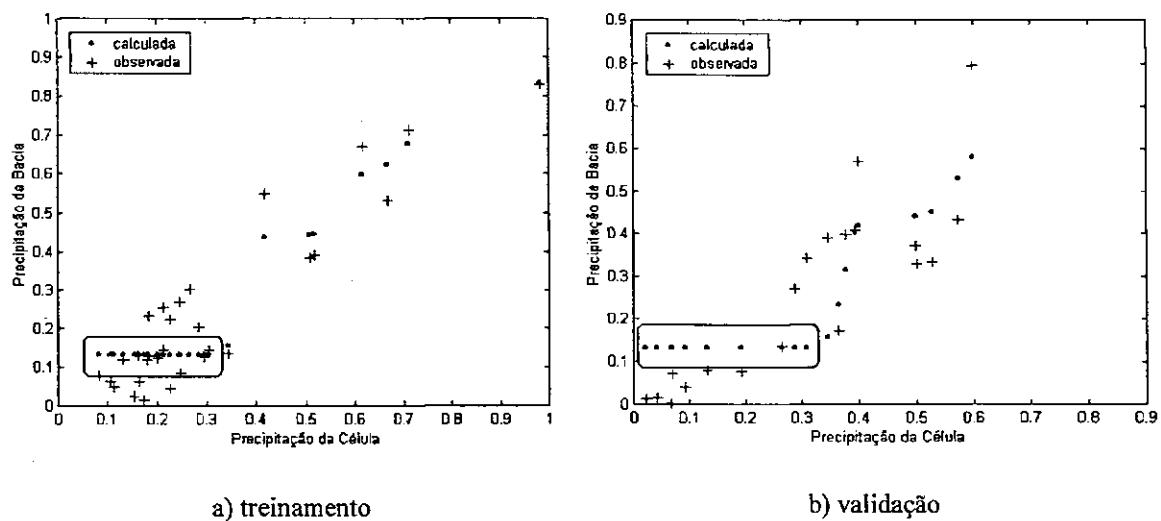
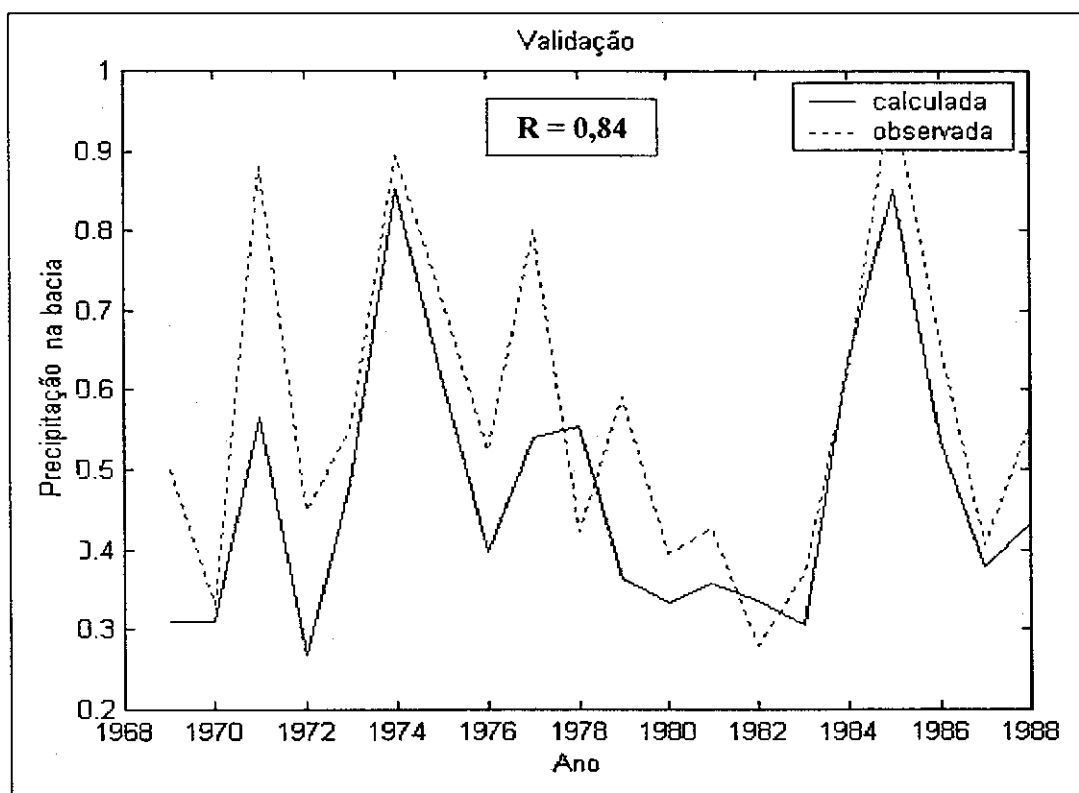


Figura 4.32 – Relações resultantes da Rede 37 entre as séries normalizadas da célula e da bacia, referentes ao mês de Maio, no período de: a) treinamento; e b) validação

O elemento responsável por se obter redes que forneçam resultados tão diferentes de uma estrutura de rede, é o algoritmo de treinamento, e é por isso que, apesar de se ter obtido grandes avanços na formulação desses algoritmos, várias pesquisas continuam a ser realizadas com o objetivo de promover correções e proporcionar melhorias a esses elementos.

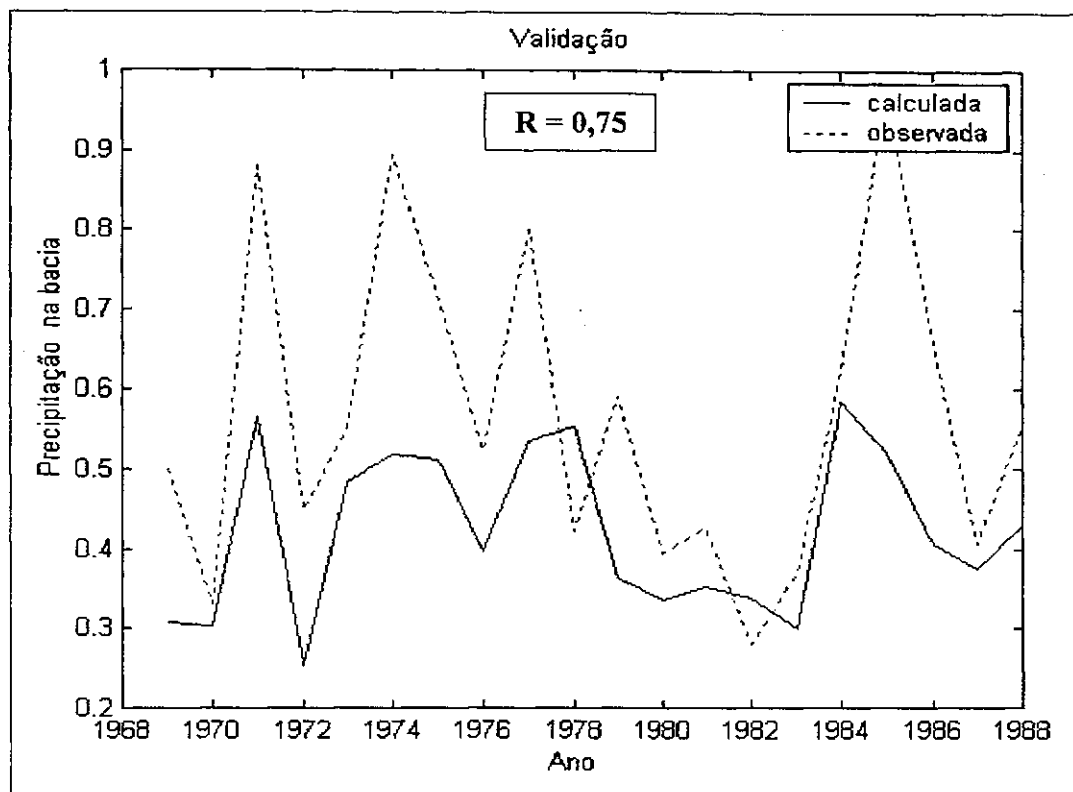
4.3.7. OCORRÊNCIA DE SUPERAJUSTAMENTO

Ao se fazer várias simulações com a rede que possui 9 neurônios na camada escondida, projetada para a escala sazonal, constatou-se que a cada aumento de 0,0001 na correlação entre as séries observada e prevista no treinamento, que não provoca mudanças visíveis no ajuste, os gráficos gerados na validação ficaram bastante diferentes uns dos outros (Figura 4.33), com correlações variando de 0,84 a -0,12. Isso é o resultado de uma rede instável devido à ocorrência de um superajustamento nos dados do período de treinamento, ou seja, apesar de fornecer resultados com altas correlações na fase de treinamento, a rede apenas memorizou os dados, perdendo sua boa capacidade de generalização.

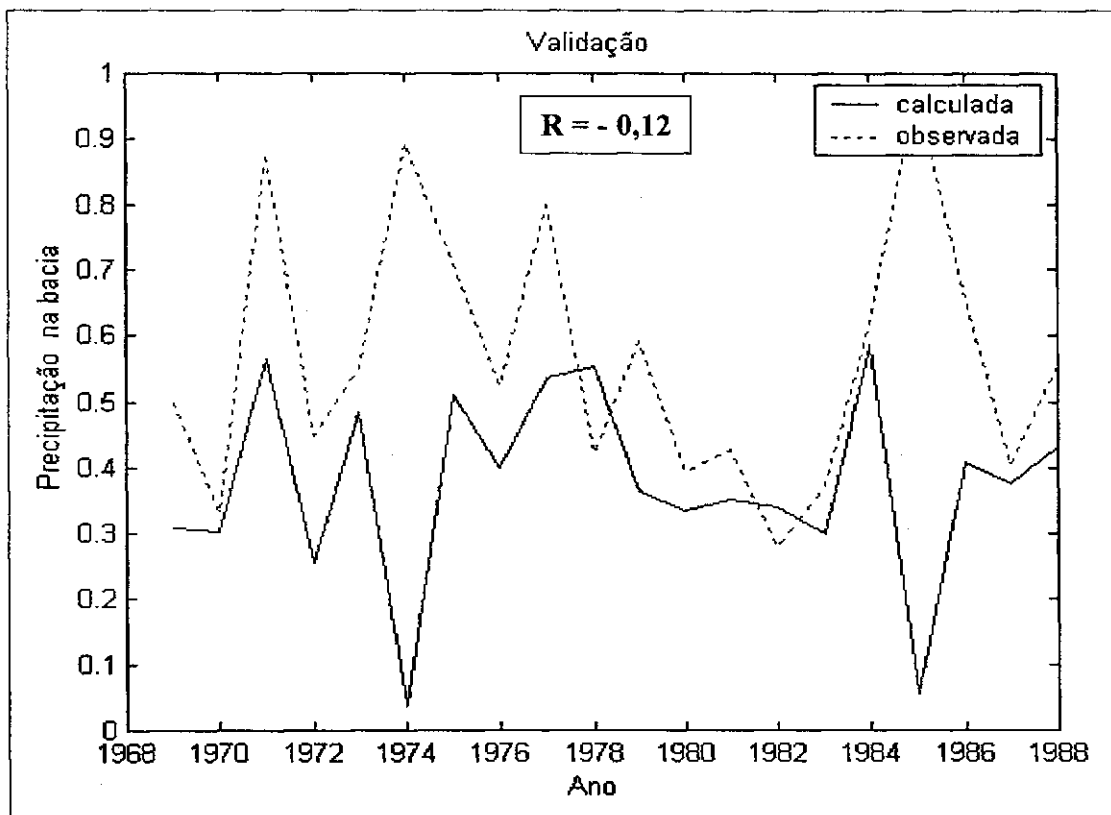


a) coeficiente de correlação do treinamento = 0.9846

Figura 4.33 – Validação de redes neurais com 9 neurônios na camada escondida correspondentes a treinamentos que obtiveram coeficientes de correlação de: a) 0,9846; b) 0,9849; e c) 0,9850



b) coeficiente de correlação do treinamento = 0.9849



c) coeficiente de correlação do treinamento = 0.9850

Figura 4.33 – Validação de redes neurais com 9 neurônios na camada escondida correspondentes a treinamentos que obtiveram coeficientes de correlação de: a) 0,9846; b) 0,9849; e c) 0,9850 (continuação)

A Figura 4.33 mostra um típico exemplo de superajustamento numa rede. Isto acontece quando ela contém um número elevado de neurônios, resultando em bons ajustes no treinamento, e péssimos resultados na validação. Diante de tal situação, o fluxograma proposto por Hines (1997) é questionado (seção 3.5.2 e Figura 3.6), pois para o caso de não se obter ajuste satisfatório na validação, ele aponta como causas ou uma fraca generalização do conjunto de treinamento, que deverá ser resolvida com uma nova seleção do grupo, ou a uma má representatividade do conjunto de validação, indicando que este necessita de outros dados. Hines (1997) esquece de mencionar, no entanto, algo que Valença e Ludermir (1999b) consideram como um dos aspectos fundamentais a ser analisado quando do treinamento de uma rede neural: a ocorrência de superajustamento (*overfitting*). Esquece que uma “fraca generalização do conjunto de treinamento” pode ser o resultado de um número excessivo de neurônios, o que seria resolvido com a redução deste número. Por isso, propõe-se aqui, uma pequena modificação neste fluxograma (Figura 4.34), onde se passa então a considerar uma das maiores causas de ajuste não satisfatório na validação: o superajustamento de uma rede.

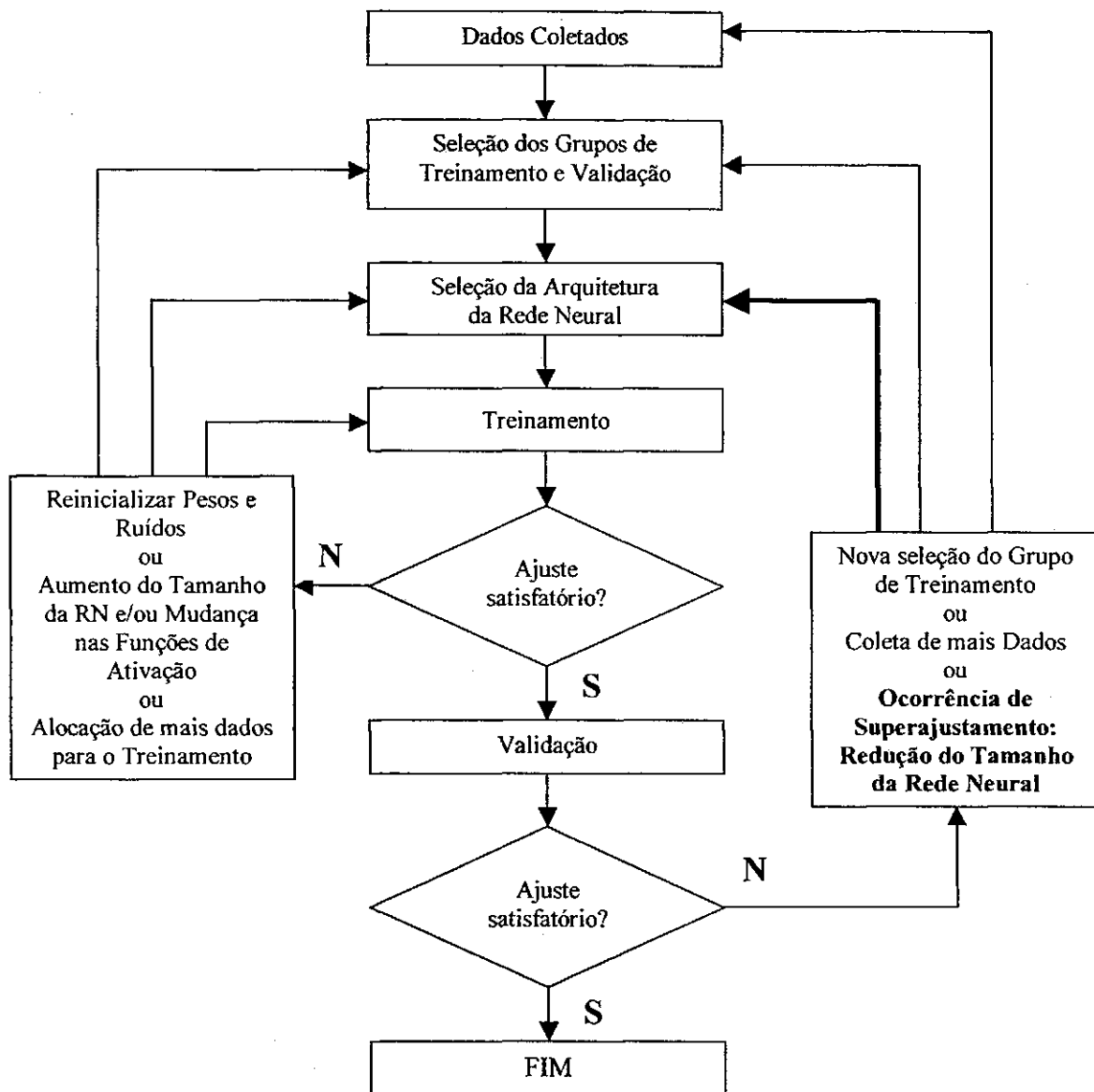


Figura 4.34 – Fluxograma do processo de treinamento e validação de uma RNA com aprendizado supervisionado (modificado de Hines, 1997)

4.3.8. COMENTÁRIOS FINAIS SOBRE O MÉTODO DE REDES NEURAIS

O método das redes neurais correspondeu às expectativas criadas com base na literatura (Sarle, 1994; Hewitson e Crane, 1996; Cavazos, 1997; entre outros), conseguindo capturar bem a não linearidade da relação entre a precipitação nas duas escalas, mostrando poder ser usado como uma técnica de desagregação da precipitação para a bacia do Piancó, tanto no nível sazonal quanto mensal. Devido à existência de relações aproximadamente lineares entre dados de precipitação da célula e da bacia, obteve-se redes com estruturas pequenas, possuindo, no máximo, 4 neurônios na camada escondida.

Para a implementação desse método, deve-se, no entanto, considerar especial atenção para um dos aspectos mais importantes: o excesso de neurônios numa rede, e o conseqüente superajustamento dos dados. Num primeiro ensaio deste estudo, apontou-se como as melhores redes para a desagregação, estruturas com até 8 neurônios na camada escondida. Só depois de se efetuar as dez calibrações de cada estrutura para a verificação da estabilidade da rede, e de se levar em consideração o princípio da parcimônia, verificou-se que se tratava de redes muito grandes (para o caso específico da bacia do Piancó), e por isso muito instáveis.

4.4. COMPARAÇÃO ENTRE OS MÉTODOS ADOTADOS PARA A DESAGREGAÇÃO

Apesar da existência de relações aproximadamente lineares entre dados de precipitação da célula e da bacia, para as escalas sazonal e mensal (Figuras 4.5 e 4.10), o método das redes neurais artificiais obteve melhores resultados nas fases de treinamento e de validação, em comparação aos resultados obtidos com a regressão linear simples (Tabela 4.7). Isto se deve, principalmente, à capacidade que um RNA tem de capturar a não linearidade da relação da precipitação nas duas escalas.

Através dos coeficientes de correlação apresentados na Tabela 4.7, observou-se que a diferença no desempenho dos dois métodos, apesar de considerável, é pequena, devido às “relações aproximadamente lineares”. Com isto, questionamentos sobre os reais benefícios que uma RNA poderia trazer para a desagregação da precipitação sobre a região em estudo, podem ser levantados. No entanto, deve-se lembrar que, se no caso específico da

bacia do Piancó, que apresenta relações entre dados relativamente fáceis de ser modeladas por outros métodos, a RNA ainda assim se mostrou superior; em outras regiões, com qualquer outro tipo de relação entre os dados, sua superioridade poderia ser ainda maior.

Tabela 4.7 – Coeficientes de correlação obtidos entre a precipitação observada da bacia do Piancó e os resultados da Regressão linear e das Redes Neurais, nos períodos de treinamento e validação

Escala temporal		Regressão linear		Redes Neurais	
		Treinamento	Validação	Treinamento	Validação
Sazonal		0,95	0,86	0,97	0,87
Mensal	Fevereiro	0,87	0,83	0,90	0,85
	Março	0,95	0,73	0,97	0,78
	Abril	0,92	0,93	0,94	0,92
	Maior	0,92	0,86	0,95	0,88

5. CONCLUSÃO E RECOMENDAÇÕES

Diante da carência de água existente no Nordeste do Brasil, previsões de longo prazo de precipitação são de grande valia para o manejo dos recursos hídricos da região. No entanto, existe o problema da diferença entre a escala para a qual são realizadas as previsões (malha dos modelos de circulação global atmosférica – MCGs), e a escala da região onde se deseja empregá-las (bacia hidrográfica, neste estudo). Por isso, os objetivos desta dissertação, já citados no Capítulo de Introdução, para uma região do Nordeste do Brasil, foram: (i) melhorar a desagregação espacial de precipitação sazonal em relação a trabalhos anteriores, (ii) realizar a desagregação espacial de precipitação mensal, referente aos meses da estação chuvosa e (iii) avaliar o comportamento do método de Redes Neurais Artificiais (RNAs) como técnica de desagregação espacial.

Como os dados observados nas duas escalas espaciais consideradas neste estudo (malha de um MCG e bacia do Rio Piancó) apresentaram relações sazonais e mensais aproximadamente lineares, utilizou-se o método de Regressão Linear para também se realizar a desagregação da precipitação, com o objetivo de se comparar e avaliar a eficiência do método das RNAs. Apesar da Regressão Linear ter fornecido bons resultados na desagregação, as RNAs forneceram resultados ainda melhores, obtendo altas correlações com os dados observados. Atribui-se a isto a capacidade que uma rede tem de conseguir capturar a não linearidade da relação entre a precipitação nas duas escalas, mesmo se tratando de relações aproximadamente lineares. Entretanto, verificou-se também que, em todos os meses considerados (fevereiro a maio) e no nível sazonal, o desempenho das redes neurais foi menor nas fases de validação, sendo esta queda atribuída à existência de grandes diferenças entre a precipitação observada nas duas escalas em alguns anos (1971, 1977, 1978 e 1979) do grupo de dados selecionados para esta fase da modelagem. Uma seleção mais adequada desses dados, obtida talvez através de validação cruzada (Valença e Ludermir, 1999b), poderia resultar no aumento do desempenho das redes, com uma melhoria das correlações também na fase de validação.

Constatou-se a importância de se efetuar a normalização dos dados de entrada da rede neural antes de utilizá-los no seu treinamento (calibração), e também a necessidade de se atribuir critérios de seleção para a escolha da melhor rede, de modo a se evitar um número excessivo de parâmetros, que pode causar problemas na utilização da mesma. Verificou-se, também, a importância de se utilizar a análise gráfica como critério de seleção da rede, e não apenas a verificação dos coeficientes de correlação, diante da alta sensibilidade das redes aos pesos e ruídos iniciais.

Portanto, sintetizando os resultados expostos acima, tem-se como conclusões deste estudo:

- Houve melhoria da desagregação espacial da precipitação sobre a bacia do Piacó, no nível sazonal, em comparação ao estudo de Galvão (1999);
- A precipitação pode ser desagregada também no nível mensal, com mesmo desempenho que no nível sazonal;
- Apesar da verificação de que as relações sazonais e mensais para a região são aproximadamente lineares, o método das RNAs forneceu resultados ainda melhores que o método da Regressão Linear;
- É importante que se efetue a normalização dos dados de entrada da rede antes de utilizá-los na sua calibração;
- É necessário o uso de vários critérios para a seleção da melhor rede;
- É importante que se utilize a análise gráfica como critério de seleção.

Conclui-se, portanto, que o método das Redes Neurais Artificiais pode e deve ser utilizado como técnica de desagregação espacial de precipitação, tanto no nível sazonal quanto mensal, devendo-se, no entanto, buscar o aperfeiçoamento deste através de alterações nos elementos que nele interferem (algoritmo de treinamento, por exemplo), para que se possa ter resultados ainda melhores numa desagregação de precipitação.

Com base nos conhecimentos adquiridos nesta pesquisa, além de sugestões lançadas até aqui, faz-se ainda algumas recomendações a trabalhos posteriores:

- Verificar se a utilização de outros nós da malha do MCG próximos ao nó utilizado acarretaria em melhorias na desagregação da precipitação;
- Desagregar a precipitação para a escala de posto pluviométrico;
- Verificar a previsibilidade do MCG para o nível mensal.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALVES, J. M. B. e REPELLI, C. A. (1992). A variabilidade pluviométrica no setor norte do Nordeste e os eventos El Niño – Oscilação Sul (ENOS). *Revista Brasileira de Meteorologia*. Rio de Janeiro, v.7, n.2, p.583-592.
- ALVES, J. M. B. e REPELLI, C. A. MELLO, N. S. e FERREIRA, A. G. (1994). Um estudo da variabilidade pluviométrica no setor norte do nordeste ao longo deste século. *In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 8., 1994, Belo Horizonte. Anais: ... Belo Horizonte: SBMET. v.2, p.111-114.*
- AZOFF, E. M. (1994). *Neural network time series forecasting of financial markets*. Chichester: John Wiley & Sons.
- BEALE, R. e JACKSON, T. (1990). *Neural computing: an introduction*. New York: Adam Hilger.
- BECKER, B. D. (1997). *The ECMWF ensemble simulation CD-ROMs*. Reading: European Centre for Medium-Range Weather Forecasts. Em CD-ROM.
- BRANKOVIC, C. e PALMER, T. N. (1998). *Seasonal skill and predictability of ECMWF PROVOST ensembles*. Reading: ECMWF. (ECMWF Technical Memorandum No. 263).
- BRANKOVIC, C., PALMER, T. N. e FERRANTI, L. (1994). Predictability of seasonal atmospheric variations. *Journal of Climate*. Boston, v.7, n.2, p.217-237.
- BRITO, J. I. B., NOBRE, C. A. e ZARANZA, A. R. (1991). A precipitação da pré-estação chuvosa e a previsibilidade da estação chuvosa do norte do Nordeste. *Climanálise*. São José dos Campos, v.6, n.6, p.39-55.
- CADIER, E. (1994). *Hidrologia das pequenas bacias do Nordeste semi-árido: transposição hidrológica*. Recife: SUDENE.
- CAUDILL, M. (1997). *Neural Networks Primer*. 3 Ed. San Francisco: Miller Freeman Inc.

- CAVALCANTI, I. F.A., PEZZI, L. P., NOBRE, P., SAMPAIO, G. e CAMARGO JR., H. (1999). Climate prediction of precipitation in Brazil for the Nordeste rainy season (MAM). *Experimental Long-Lead Forecast Bulletin*, v.8, n.1. Disponível on-line em <http://grads.iges.org/ellfb/Mar99/cptec.htm> . 18 Nov 1999.
- CAVAZOS, T. (1997). Downscaling large-scale circulation to local winter rainfall in north-eastern Mexico. *International Journal of Climatology*. Chichester, v.17, p.1069-1082.
- COLMAN, A., DAVEY, M., HARRISON, M. e EVANS, A. (1998). Prediction of march-april-may 1998 rainfall in Northeast Brazil using input from multiple regression, discriminant analysis and an atmospheric global circulation model. *Experimental Long-lead Forecast Bulletin*. v.7, n.1. <http://grads.iges.org/ellfb/Mar98/co2.html>.
- COSTA, W. D. e COSTA, W. D. (1997). Disponibilidades hídricas subterrâneas na região Nordeste do Brasil. *A Água em Revista*. anoV, n.9, p.47-59.
- DEMUTH, H. e BEALE, M. (1998). *Neural Network Toolbox for use with MATLAB – User's guide*, version 3. The MathWorks, Inc.
- EKLUND, P. (1994). Network size versus preprocessing. In: Yager, R. R., Zadeh, L. A. (Ed.). *Fuzzy sets, neural networks, and soft computing*. New York: Van Nostrand Reinhold. p.250-264.
- ECMWF – EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS (2000). *ECMWF's seasonal forecast project: introduction*. Disponível on-line em <http://www.ecmwf.int/services/seasonal/info/info.html> . 02 Mai 2000.
- EVANS, T. e GRAHAM, R. (1999). A dynamical one-month lead seasonal rainfall prediction for march to may 1998 for the North-Eastern area of South America. *Experimental Long-Lead Forecast Bulletin*, v.8, n.1. Disponível on-line em <http://grads.iges.org/ellfb/Mar99/evens.htm> . 18 Nov 1999.
- FOLLAND, C. K., OWEN, J., WARD, N. e COLMAN, A. (1991). Prediction of seasonal rainfall in the Sahel region using empirical and dynamical methods. *Journal of Forecasting*. Chichester, v.10, p.21-56.
- FONTANA, D. C. e BERLATO, M. A. (1997). A influência do El Niño Oscilação Sul sobre a precipitação pluvial no Estado do Rio Grande do Sul. *Revista Brasileira de Agrometeorologia*. Santa Maria, v.5, n.1, p.127-132.

- FREDERICK, K. D., MAJOR, D. C. e STAKHIV, E. Z. (1997). Water resources planning principles and evaluation criteria for climate change: summary e conclusions. *Climatic Change*, Dordrecht, v.37, p.291-313.
- FREITAS, M. A. V. e SANTOS, A. H. M. (1999). Importância da água e da informação hidrológica. In: Freitas, M. A. V. (Org.). *O estado das águas no Brasil*. Brasília: ANEEL, MMA/SRH, MME. p.13-16.
- GALVÃO, C. O. (1999). *Aplicabilidade em Recursos Hídricos da previsão de precipitação de longo prazo no nordeste do Brasil*. Porto Alegre. Tese (Doutor em Engenharia) – Programa de Pós-Graduação em Engenharia de Recursos Hídricos e Saneamento Ambiental. Universidade Federal do Rio Grande do Sul, Porto Alegre.
- GALVÃO, C. O. e TRIGO, R. M. (1999). Downscaling da precipitação e temperatura de modelos de circulação global para a escala de bacia hidrográfica. In: IV SIMPÓSIO DE HIDRÁULICA E RECURSOS HÍDRICOS DOS PAÍSES DE LÍNGUA OFICIAL PORTUGUESA. Coimbra: Portugal. Em CD-ROM.
- GARDNER, M. W. e DORLING, S. R. (1998). Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences. *Atmospheric Environment*. v.32, p.2627-2636.
- GATELY, E. (1996). *Neural networks for financial forecasting*. New York: John Wiley & Sons.
- HAGAN, M. T. e MENHAJ, M. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*. v. 5, n.6, p.989-993.
- HANSELMAN, D. e LITTLEFIELD, B. (1999). *MATLAB 5 Versão do Estudante – Guia do Usuário*. São Paulo: MAKRON Books.
- HARZALLAH, A., ARAGÃO, J. O. R. e SADOURNY, R. (1996). Interannual rainfall variability in North-East Brazil: observation and model simulation. *International Journal of Climatology*. Chichester, v.16, p.861-878.
- HERDIES, D. L., PEZZI, L. P. e ARAVÉQUIA, J. A. (1994). Análise da estação chuvosa do leste do Nordeste. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 8., 1994, Belo Horizonte. *Anais: ...* Belo Horizonte: SBMET. v.2, p.155-157.
- HEWITSON, B. C. e CRANE, R. G. (1996). Climate downscaling: techniques and application. *Climate Research*, v.7. p.85-95.
- HINES, J. W. (1997). *MATLAB Supplement to fuzzy and neural approaches in engineering*. Chichester: John Wiley & Sons.

- KANE, R. P. (1997). Prediction of droughts in North-East Brazil: role of ENSO and use of periodicities. *International Journal of Climatology*. Chichester, v.17, p.655-665.
- KAYANO, M. T. e KOUSKY, V. E. (1994). Análises regionais da variabilidade interanual. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 8., 1994, Belo Horizonte. *Anais: ... Belo Horizonte: SBMET*. v.2, p.7-9.
- KOUSKY, V. E. (1985). Atmospheric circulation changes associated with rainfall anomalies over tropical Brazil. *Monthly Weather Review*. Boston, v.113, p.1951-1957.
- LAWRENCE, J. (1991). Data preparation for a neural network. *AI Expert*. November, p.34-41.
- LIMA, M. C. de e RAO, V. B. (1994). Estudo de casos extremos (secas/enchentes) no leste da região Nordeste do Brasil. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 8., 1994, Belo Horizonte. *Anais: ... Belo Horizonte: SBMET*. v.2, p.46-48.
- MICROSOFT CORPORATION (1997). *Excel 97*. Microsoft Co.
- MMA – MINISTÉRIO DO MEIO AMBIENTE (2000). *PROÁGUA semi-árido*. Disponível on-line em <http://www.proagua.gov.br/htm/default.htm>. 14 Out 2000.
- MOLINIER, M., ALBUQUERQUE, C. H. C. e CADIER, E. (1994). *Análise da pluviometria e isoietas homogeneizadas do Nordeste brasileiro*. Recife: SUDENE.
- NIMER, E. (1989). *Climatologia do Brasil*. Rio de Janeiro: IBGE.
- NOBRE, C. A. e MOLION, L. C. B. (1988). The climatology of droughts and drought prediction. In: PARRY, M. L., CARTER, T. R. e KONIJN, N. T. (Ed.) *the Impact of Climate Variations on Agriculture*. Dordrecht: Kluwer: IIASA:UNEP. v.2, p.305-323.
- PALUTIKOF, J. P., WINKLER, J. A., GOODESS, C. M. e ANDRESEN, J. A. (1997). The simulation of daily temperature time series from GCM output. Part I: comparison of model data with observations. *Journal of Climate*. Boston, v.10, p.2497-2513.
- PIELKE, R. A., COTTON, W. R., WALKO, R. L., TREMBACK, C. J., LYONS, W. A., GRASSO, L. D., NOCHOLLS, M. E., MORAN, M. D., WESLEY, D. A., LEE, T. J. e COPELAND, J. H. (1992). A comprehensive meteorological modeling system – RAMS. *Meteorology Atmospheric Physics*, v.49, p.69-91.
- RAMOS, R. P. L. (1975). Precipitation characteristics in the Northeast Brazil dry region. *J. Geophys. Res.*, v.80, p.1665-1678.

- RAO, V. B., LIMA, M. C. e FRANCHITO, S. H. (1993). Seasonal and interannual variations of rainfall over eastern Northeast Brazil. *Journal of Climate*. Boston, v.6, p.1754-1763.
- REPELLI, C. A. e ALVES, J. M. B. (1994). O estado da arte da previsibilidade operacional da quadra chuvosa do setor norte do nordeste do Brasil. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 8., 1994, Belo Horizonte. *Anais: ...* Belo Horizonte: SBMET. v.3, p.438-442.
- ROBINSON, R. (1991). Neural networks offer an alternative to traditional regression. *Geobyte*. February, p.14-19.
- ROSENBLAT, F. (1958). The Perceptron: a probabilistic model for information storage and organization in the human brain. *Psychological Review*, v. 65, p.386-408.
- ROUCOU, P., ARAGÃO, J. O. R., HARZALLAH, A., FONTAINE, B. e JANICOT, S. (1996). Vertical motion changes related to North-East Brazil rainfall variability: a GCM simulation. *International Journal of Climatology*. Chichester, v.16, p.879-891.
- SARLE, W. S. (1994). Neural networks and statistical models. *Proceedings of the Nineteenth Annual SAS Users Group International Conference*. Cary, NC: SAS Institute.
- SCHUBERT, S. (1998). Downscaling local extreme temperature changes in south-eastern Australia from the CSIRO Mark2 GCM. *International Journal of Climatology*. Chichester, v.18, p.1419-1438.
- SEMENOV, M. A. e BARROW, E. M. (1997). Use of a stochastic weather generator in the development of climate change scenarios. *Climatic Change*, Dordrecht, v.35, p.397-414.
- SKELLY, W. C. e HENDERSON-SELLERS, A. (1996). Grid box ou grid point: what type of data do GCMs deliver to climate impacts researchers? *International Journal of Climatology*. Chichester, v.16, p.1079-1086.
- SOUZA, E. B., ALVES, J. M. B. e XAVIER, T. B. S. (1998). A estação chuvosa no Semi-Árido Nordestino nos anos de predominância de aquecimento ou de resfriamento em toda bacia do atlântico tropical. *Revista Brasileira de Recursos Hídricos*. v.3, n.2, p.47-55.

- SPERBER, K. R. e PALMER, T. N. (1996). Interannual tropical rainfall variability in general circulation model simulations associated with the atmospheric model intercomparison project. *Journal of Climate*. Boston, v.9, p.2727-2750.
- TRIGO, R. M. e PALUTIKOF, J. P. (1999). Simulation of daily maximum and minimum temperature over Portugal: a neural network model approach. *Climate Research*. V.13, p.45-59
- TUCCI, C. E. M. (1998). *Modelos hidrológicos*. Porto Alegre: Associação Brasileira de Recursos Hídricos.
- UVO, C. B. (1998). *Influence of sea surface temperature on rainfall and runoff in northeastern South America: analysis and modeling*. Department of Water Resources Engineering, Lund Institute of Technology, Lund University, Sweden. Tese (Doutorado)
- UVO, C. B., REPELLI, C. A., ZEBIAK, S. e KUSHNIR, Y. (1998). The relationships between tropical pacific and atlantic SST and the Northeast Brazil monthly precipitation. *Journal of Climate*. Boston, v.11, p.551-562.
- VALENÇA, M. J. S. e LUDERMIR, T. B. (1999a). Introdução às redes neurais. In: Galvão, C. O., Valença, M. J. S. (Org.). *Sistemas Inteligentes; aplicações a Recursos Hídricos e ciências ambientais*. Porto Alegre: Associação Brasileira de Recursos Hídricos. p.19-59.
- VALENÇA, M. J. S. e LUDERMIR, T. B. (1999b). Aplicações de redes neurais. In: Galvão, C. O., Valença, M. J. S. (Org.). *Sistemas Inteligentes; aplicações a Recursos Hídricos e ciências ambientais*. Porto Alegre: Associação Brasileira de Recursos Hídricos. p.61-96.
- WARD, M. N. e FOLLAND, C. K. (1991). Prediction of seasonal rainfall in the North Nordeste of Brazil using eigenvectors of sea surface temperatures. *International Journal of Climatology*. Chichester, v.11, p.711-743.
- WILBY, R. L., WIGLEY, T. M. L., CONWAY, D., JONES, P. D., HEWITSON, B. C., MAIN, J. e WILKS, D. S. (1998). Statistical downscaling of general circulation model output: a comparison of methods. *Water Resources Research*, Washington, v.34, n.11, p.2995-3008.
- WMO – WORLD METEOROLOGICAL ORGANIZATION (1997). *Comprehensive assessment of the freshwater resources of the World*. Genebra: WMO.