

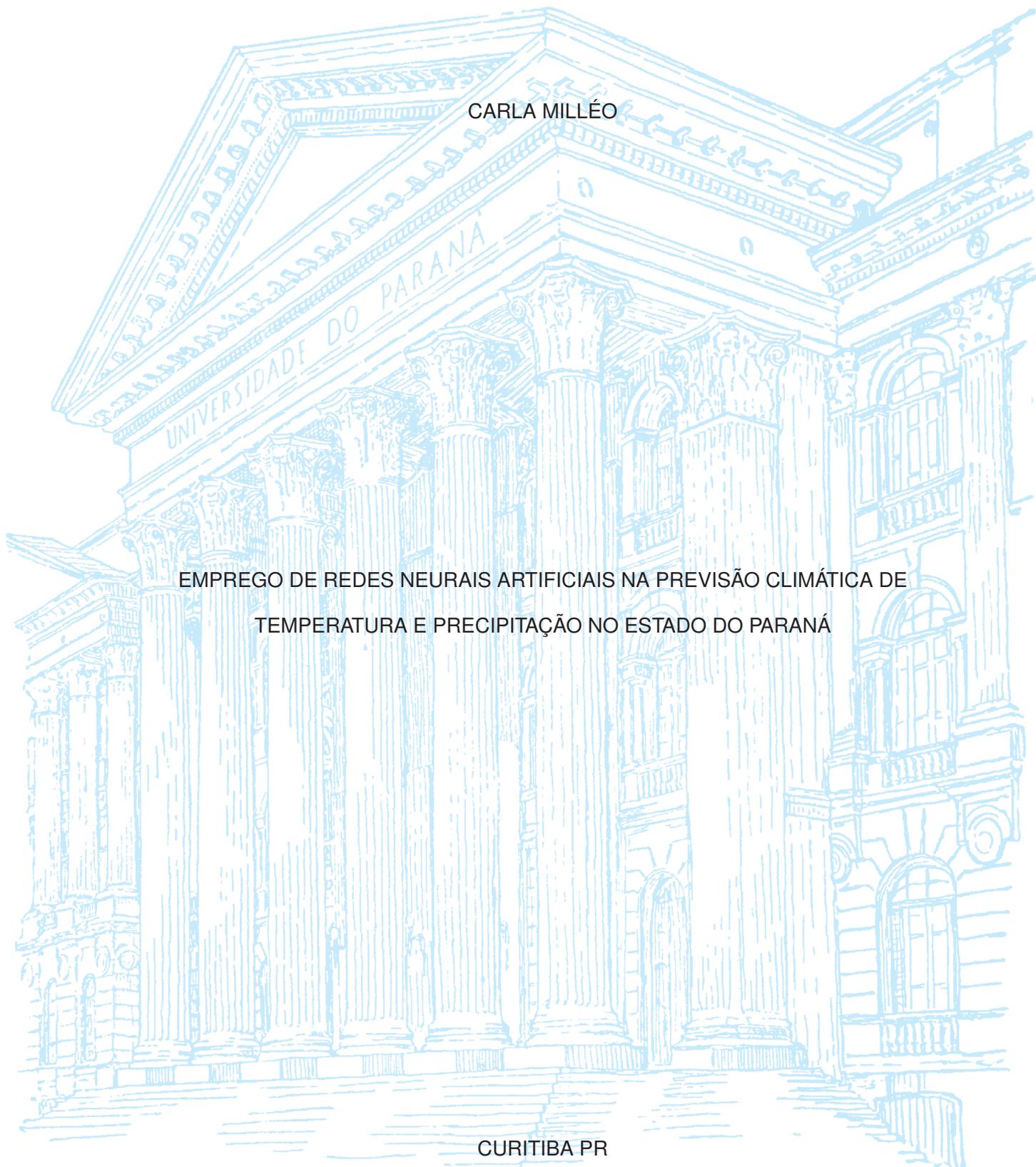
UNIVERSIDADE FEDERAL DO PARANÁ

CARLA MILLÉO

EMPREGO DE REDES NEURAIS ARTIFICIAIS NA PREVISÃO CLIMÁTICA DE
TEMPERATURA E PRECIPITAÇÃO NO ESTADO DO PARANÁ

CURITIBA PR

2020



CARLA MILLÉO

EMPREGO DE REDES NEURAIS ARTIFICIAIS NA PREVISÃO CLIMÁTICA DE
TEMPERATURA E PRECIPITAÇÃO NO ESTADO DO PARANÁ

Dissertação apresentada como requisito parcial à obtenção do título de Mestre em Engenharia Ambiental, no Programa de Pós-Graduação em Engenharia Ambiental, Setor de Tecnologia, da Universidade Federal do Paraná..

Área de concentração: *Atmosfera e Mudanças Globais*.

Orientador: Prof. Dr. Ricardo Carvalho de Almeida.

CURITIBA PR

2020

Catálogo na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

M646e

Milléo, Carla

Emprego de redes neurais artificiais na previsão climática de temperatura e precipitação no estado do Paraná [recurso eletrônico] / Carla Milléo. – Curitiba, 2020.

Dissertação - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-Graduação em Engenharia Ambiental, 2020.

Orientador: Ricardo Carvalho de Almeida

1. Meteorologia. 2. Temperatura atmosférica. 3. Precipitação (Meteorologia) – Paraná. 4. Redes neurais (Computação). I. Universidade Federal do Paraná. II. Almeida, Ricardo Carvalho de. III. Título.

CDD: 551.5

Bibliotecário: Elias Barbosa da Silva CRB-9/1894



MINISTÉRIO DA EDUCAÇÃO
SETOR DE TECNOLOGIA
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO ENGENHARIA
AMBIENTAL - 40001016075P3

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA AMBIENTAL da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **CARLA MILLÉO** intitulada: **EMPREGO DE REDES NEURAIAS ARTIFICIAIS NA PREVISÃO CLIMÁTICA DE TEMPERATURA E PRECIPITAÇÃO NO ESTADO DO PARANÁ**, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua **APROVAÇÃO** no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 26 de Agosto de 2020.

Assinatura Eletrônica

27/08/2020 17:13:55.0

RICARDO CARVALHO DE ALMEIDA

Presidente da Banca Examinadora (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

31/08/2020 11:59:22.0

JOSÉ EDUARDO GONÇALVES

Avaliador Externo (INSTITUTO TECNOLÓGICO SIMEPAR)

Assinatura Eletrônica

27/08/2020 17:35:20.0

EMILIO GRACILIANO FERREIRA MERCURI

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Ricardo Carvalho de Almeida, por me despertar o interesse na área das ciências atmosféricas, sobretudo pela dedicação, atenção e paciência com seus alunos.

Aos professores Emílio Mercuri e José Eduardo Gonçalves que aceitaram o convite de compor a minha banca, dedicando tempo e atenção, além de também enviarem suas contribuições para o trabalho.

Aos professores e professoras do mestrado que tive a honra de poder compartilhar de seus conhecimentos e entusiasmo pela ciência.

Ao Programa de Pós-Graduação em Engenharia Ambiental da UFPR e à CAPES, pela concessão da bolsa de estudos, que permitiram que esta pesquisa fosse realizada, contribuindo para meu crescimento pessoal e profissional.

Aos meus pais, Marilyn e Péricles, e a minha irmã, Kariny, pela dedicação, amor e carinho em todos os anos da minha vida. Também agradeço meu namorado, Evandro, por todo apoio, motivação e por toda ajuda ofertada (inclusive com o servidor Linux para rodar os programas necessários desta dissertação).

Meu muito obrigada aos meus amigos, pelo suporte e por tornarem tudo mais leve. As minhas amigas da graduação, em especial à Geovana, que dividiram comigo momentos de alegria e angustia dessa trajetória acadêmica. Aos meus novos amigos do Instituto Água e Terra que me receberam de braços abertos e que me trazem muitas alegrias. As minhas amigas de longa data, Ana e Ane, que entendem meus momentos de ausência e que eu sei que posso contar com elas para tudo. E a todos os amigos e colegas do PPGEA que me acompanharam durante disciplinas, seminários, almoços no RU e cafés da tarde.

Agradeço a todos aqueles que, diretamente ou indiretamente, me ajudaram a chegar até aqui.

"All models are wrong, but some are useful."

GEORGE E. P. BOX

RESUMO

O objetivo deste estudo é avaliar o desempenho das Redes Neurais Artificiais (RNA) de contra-propagação (CPN) e de função de base radial (RBF) na previsão climática de temperatura e precipitação no estado do Paraná. Uma etapa importante no desenvolvimento de um modelo de RNA é a seleção de variáveis de entrada. Para isso, buscou-se na literatura os principais fenômenos atmosféricos que produzem mudanças nos quadros médios de temperatura e precipitação no Paraná. Após selecionar os possíveis preditores das variações climáticas, foi utilizado o método de regressão *stepwise*, passo *forward*, para identificar o melhor conjunto de variáveis de entrada das redes neurais. A rede neural CPN, com 150 neurônios na camada oculta, e RBF, com 50 neurônios na camada oculta, foram desenvolvidas para cada variável a ser prevista. Os resultados mostraram que o desempenho da rede do tipo RBF foi superior ao da rede de CPN em todos os experimentos de previsões. Os melhores desempenhos foram obtidos para as previsões de variáveis filtradas por uma média móvel de três meses, em virtude de dados que possuem muito ruído são mais difíceis de serem simulados por RNA. As técnicas de RNAs mostraram melhor desempenho nas estimativas de temperatura e precipitação do Paraná que a técnica de persistência.

Palavras-chave: Previsão de temperatura e precipitação. Redes Neurais Artificiais. Seleção de variáveis de entrada. Redes Neurais de Contra-Propagação. Redes Neurais RBF.

ABSTRACT

The objective of this study is to evaluate the performance of counter-propagation (CPN) and radial base function (RBF) Artificial Neural Networks (ANN) in the climatic forecast of temperature and precipitation in the state of Paraná. An important step in the development of an ANN model is the selection of input variables. For that, the main atmospheric phenomena that produced changes in the average temperature and precipitation conditions in Paraná were searched in the literature. After selecting the possible predictors of climatic variations, the stepwise regression method, step forward, was used to identify the best set of input variables for neural networks. The CPN neural network, with 150 neurons in the hidden layer, and RBF, with 50 neurons in the hidden layer, were developed for each variable to be predicted. The results showed that the performance of the RBF network was superior to that of the CPN network in all forecasting experiments. The best performances were obtained for the predictions of variables filtered by a three-month moving average, due to data that have a lot of noise are more difficult to be simulated by RNA. The ANN techniques performed better than simple persistence in the forecast of temperature and precipitation for state of Paraná.

Keywords: Temperature and precipitation forecasting. Artificial Neural Network. Selection of input variables. Counterpropagation Neural Networks. RBF Neural Networks.

LISTA DE FIGURAS

2.1	Sistemas atmosféricos na baixa e alta troposfera atuantes na América do Sul. Fonte: Adaptado de Satyamurty et al. (1998).	22
2.2	Anomalia da temperatura média do oceano Pacífico em dezembro de 1988 e dezembro de 1997 correspondendo a um evento de La Niña e El Niño, respectivamente. Fonte: Adaptado de World Meteorological Organization (2014) 23	
2.3	Representação de um neurônio biológico. Fonte: Elaborada pela autora.	25
2.4	Representação de um neurônio artificial. Fonte: Elaborada pela autora..	25
2.5	Arquitetura da rede de contra-propagação completa. Fonte: Adaptado de Fausett (1994)	27
2.6	Arquitetura da rede de contra-propagação somente adiante. Fonte: Adaptado de Fausett (1994).	29
2.7	Arquitetura da rede RBF. Fonte: Adaptado de Silva et al. (2010)	32
2.8	Função Gaussiana. Fonte: Elaborada pela autora..	32
3.1	Localização de ocorrência dos fenômenos atmosféricos que podem ocasionar a variação climática no estado do Paraná. Fonte do mapa cartográfico Wikipedia, 2019.	37
3.2	Médias climatológicas de temperatura e precipitação mensais no estado do Paraná.	39
3.3	Evolução temporal e histograma dos valores de anomalias de temperatura no Paraná.	40
3.4	Evolução temporal e histograma dos valores de anomalias de precipitação no Paraná.	41
3.5	Exemplo da configuração utilizada para a previsão sazonal.	43
3.6	Fluxograma do método de regressão <i>stepwise</i> com o passo <i>forward</i>	44
4.1	Exemplo do resultado de um ciclo do método <i>stepwise</i>	48
4.2	Resultado dos cinco ciclos do método de seleção de variáveis regressão <i>stepwise</i> para a temperatura mensal utilizando a RNA de CPN.	49
4.3	Resultado dos cinco ciclos do método de seleção de variáveis regressão <i>stepwise</i> para a temperatura sazonal utilizando a RNA de CPN.. . . .	50
4.4	Resultado dos cinco ciclos do método de seleção de variáveis regressão <i>stepwise</i> para a precipitação mensal utilizando a RNA de CPN.	51
4.5	Resultado dos cinco ciclos da segunda etapa do método de seleção de variáveis regressão <i>stepwise</i> para a precipitação mensal utilizando a RNA de CPN.	52
4.6	Resultado dos cinco ciclos do método de seleção de variáveis regressão <i>stepwise</i> para a precipitação sazonal utilizando a RNA de CPN.. . . .	53

4.7	Resultado dos cinco ciclos da segunda etapa do método de seleção de variáveis regressão <i>stepwise</i> para a precipitação sazonal utilizando a RNA de CPN.	53
4.8	Gráfico de dispersão entre os dados de reanálise e o resultado da rede de contra-propagação para a temperatura mensal.	54
4.9	Gráfico de dispersão entre os dados de reanálise e o resultado da rede de contra-propagação para a temperatura sazonal.	55
4.10	Gráfico de dispersão entre os dados de reanálise e o resultado da rede de contra-propagação para a precipitação mensal.	55
4.11	Gráfico de dispersão entre os dados de reanálise e o resultado da rede de contra-propagação para a precipitação sazonal.	56
4.12	Resultado dos cinco ciclos do método de seleção de variáveis regressão <i>stepwise</i> para a temperatura mensal utilizando a RNA RBF.	57
4.13	Resultado dos cinco ciclos da segunda etapa do método de seleção de variáveis regressão <i>stepwise</i> para a temperatura mensal utilizando a RNA RBF.	58
4.14	Resultado dos cinco ciclos do método de seleção de variáveis regressão <i>stepwise</i> para a temperatura sazonal utilizando a RNA RBF.	59
4.15	Resultado dos cinco ciclos do método de seleção de variáveis regressão <i>stepwise</i> para a precipitação mensal utilizando a RNA RBF.	60
4.16	Resultado dos cinco ciclos do método de seleção de variáveis regressão <i>stepwise</i> para a precipitação sazonal utilizando a RNA RBF.	61
4.17	Gráfico de dispersão entre os dados de reanálise e o resultado da rede RBF para a temperatura mensal.	62
4.18	Gráfico de dispersão entre os dados de reanálise e o resultado da rede RBF para a temperatura sazonal.	63
4.19	Gráfico de dispersão entre os dados de reanálise e o resultado da rede neural RBF para a precipitação mensal.	63
4.20	Gráfico de dispersão entre os dados de reanálise e o resultado da rede RBF para a precipitação sazonal.	64
A.1	Médias climatológicas da temperatura da superfície do mar na região do ASAS.	72
A.2	Anomalias da temperatura da superfície do mar na região do ASAS.	73
A.3	Médias climatológicas da pressão na superfície na região do ASAS.	73
A.4	Anomalias da pressão na superfície na região do ASAS.	74
A.5	Médias climatológicas da temperatura da superfície do mar na região do ASPS.	75
A.6	Anomalias da temperatura da superfície do mar na região do ASPS.	75
A.7	Médias climatológicas da pressão na superfície na região do ASPS.	76
A.8	Anomalias da pressão na superfície na região do ASPS.	76
A.9	Médias climatológicas da temperatura da superfície do mar na região do AMP.	77
A.10	Anomalias da temperatura da superfície do mar na região do AMP.	78
A.11	Médias climatológicas da pressão na superfície na região do AMP.	78
A.12	Anomalias da pressão na superfície na região do AMP.	79

A.13	Anomalias da TSM na região do El Niño 3.4 e Índice de Oscilação Sul de 1979 a 2018.	80
A.14	Médias climatológicas do vento meridional a 850 hPa na região do JBN.	81
A.15	Anomalias do vento meridional a 850 hPa na região do JBN.	81
A.16	Médias climatológicas do vento zonal a 250 hPa na região do JP.	82
A.17	Anomalias do vento zonal a 250 hPa na região do JP.	83
A.18	Médias climatológicas do vento zonal a 250 hPa na região do JS.	84
A.19	Anomalias do vento zonal a 250 hPa na região do JS.	84
A.20	Médias climatológicas da umidade específica a 700 hPa em duas áreas analisadas da ZCAS.	85
A.21	Anomalias da umidade específica a 700 hPa em duas áreas analisadas da ZCAS.	86
A.22	Médias climatológicas da velocidade vertical a 500 hPa em duas áreas analisadas da ZCAS.	87
A.23	Anomalias da velocidade vertical a 500 hPa em duas áreas analisadas da ZCAS.	88
A.24	Médias climatológicas da umidade específica a 700 hPa na região do PR.	89
A.25	Anomalias da umidade específica a 700 hPa na região do PR.	89
A.26	Médias climatológicas da velocidade vertical a 500 hPa na região do PR.	90
A.27	Anomalias da velocidade vertical a 500 hPa na região do PR.	91
A.28	Médias climatológicas de geopotencial a 200, 500 e 850 hPa na região do PR.	92
A.29	Histograma das anomalias de geopotenciais a 200, 500 e 850 hPa na região do PR.	93
B.1	Resultado do primeiro ciclo do método <i>stepwise</i> para a temperatura mensal: 12 variáveis selecionadas.	94
B.2	Resultado do segundo ciclo do método <i>stepwise</i> para a temperatura mensal: 10 variáveis selecionadas.	94
B.3	Resultado do terceiro ciclo do método <i>stepwise</i> para a temperatura mensal: 15 variáveis selecionadas.	95
B.4	Resultado do quarto ciclo do método <i>stepwise</i> para a temperatura mensal: 11 variáveis selecionadas.	95
B.5	Resultado do quinto ciclo do método <i>stepwise</i> para a temperatura mensal: 14 variáveis selecionadas.	95
B.6	Resultado do primeiro ciclo do método <i>stepwise</i> para a temperatura sazonal: 16 variáveis selecionadas.	96
B.7	Resultado do segundo ciclo do método <i>stepwise</i> para a temperatura sazonal: 11 variáveis selecionadas.	96
B.8	Resultado do terceiro ciclo do método <i>stepwise</i> para a temperatura sazonal: 11 variáveis selecionadas.	96
B.9	Resultado do quarto ciclo do método <i>stepwise</i> para a temperatura sazonal: 15 variáveis selecionadas.	97

B.10	Resultado do quinto ciclo do método <i>stepwise</i> para a temperatura sazonal: 15 variáveis selecionadas.	97
B.11	Resultado do primeiro ciclo do método <i>stepwise</i> para a precipitação mensal: 17 variáveis selecionadas.	97
B.12	Resultado do segundo ciclo do método <i>stepwise</i> para a precipitação mensal: 17 variáveis selecionadas.	98
B.13	Resultado do terceiro ciclo do método <i>stepwise</i> para a precipitação mensal: 17 variáveis selecionadas.	98
B.14	Resultado do quarto ciclo do método <i>stepwise</i> para a precipitação mensal: 17 variáveis selecionadas.	98
B.15	Resultado do quinto ciclo do método <i>stepwise</i> para a precipitação mensal: 17 variáveis selecionadas.	99
B.16	Resultado do primeiro ciclo da segunda etapa do método <i>stepwise</i> para a precipitação mensal: 10 variáveis selecionadas..	99
B.17	Resultado do segundo ciclo da segunda etapa do método <i>stepwise</i> para a precipitação mensal: 10 variáveis selecionadas..	100
B.18	Resultado do terceiro ciclo da segunda etapa do método <i>stepwise</i> para a precipitação mensal: 10 variáveis selecionadas..	100
B.19	Resultado do quarto ciclo da segunda etapa do método <i>stepwise</i> para a precipitação mensal: 10 variáveis selecionadas..	100
B.20	Resultado do quinto ciclo da segunda etapa do método <i>stepwise</i> para a precipitação mensal: 10 variáveis selecionadas..	101
B.21	Resultado do primeiro ciclo do método <i>stepwise</i> para a precipitação sazonal: 17 variáveis selecionadas.	101
B.22	Resultado do segundo ciclo do método <i>stepwise</i> para a precipitação sazonal: 17 variáveis selecionadas.	101
B.23	Resultado do terceiro ciclo do método <i>stepwise</i> para a precipitação sazonal: 17 variáveis selecionadas.	102
B.24	Resultado do quarto ciclo do método <i>stepwise</i> para a precipitação sazonal: 17 variáveis selecionadas..	102
B.25	Resultado do quinto ciclo do método <i>stepwise</i> para a precipitação sazonal: 14 variáveis selecionadas.	102
B.26	Resultado do primeiro ciclo do método <i>stepwise</i> para a temperatura mensal: 15 variáveis selecionadas.	103
B.27	Resultado do segundo ciclo do método <i>stepwise</i> para a temperatura mensal: 15 variáveis selecionadas.	103
B.28	Resultado do terceiro ciclo do método <i>stepwise</i> para a temperatura mensal: 06 variáveis selecionadas.	104
B.29	Resultado do quarto ciclo do método <i>stepwise</i> para a temperatura mensal: 13 variáveis selecionadas.	104
B.30	Resultado do quinto ciclo do método <i>stepwise</i> para a temperatura mensal: 15 variáveis selecionadas.	104

B.31	Resultado do primeiro ciclo do método <i>stepwise</i> para a temperatura sazonal: 13 variáveis selecionadas.	105
B.32	Resultado do segundo ciclo do método <i>stepwise</i> para a temperatura sazonal: 08 variáveis selecionadas.	105
B.33	Resultado do terceiro ciclo do método <i>stepwise</i> para a temperatura sazonal: 10 variáveis selecionadas.	105
B.34	Resultado do quarto ciclo do método <i>stepwise</i> para a temperatura sazonal: 15 variáveis selecionadas.	106
B.35	Resultado do quinto ciclo do método <i>stepwise</i> para a temperatura sazonal: 12 variáveis selecionadas.	106
B.36	Resultado do primeiro ciclo do método <i>stepwise</i> para a precipitação mensal: 08 variáveis selecionadas.	106
B.37	Resultado do segundo ciclo do método <i>stepwise</i> para a precipitação mensal: 15 variáveis selecionadas.	107
B.38	Resultado do terceiro ciclo do método <i>stepwise</i> para a precipitação mensal: 14 variáveis selecionadas.	107
B.39	Resultado do quarto ciclo do método <i>stepwise</i> para a precipitação mensal: 08 variáveis selecionadas.	107
B.40	Resultado do quinto ciclo do método <i>stepwise</i> para a precipitação mensal: 08 variáveis selecionadas.	108
B.41	Resultado do primeiro ciclo do método <i>stepwise</i> para a precipitação sazonal: 11 variáveis selecionadas.	108
B.42	Resultado do segundo ciclo do método <i>stepwise</i> para a precipitação sazonal: 08 variáveis selecionadas.	108
B.43	Resultado do terceiro ciclo do método <i>stepwise</i> para a precipitação sazonal: 10 variáveis selecionadas.	109
B.44	Resultado do quarto ciclo do método <i>stepwise</i> para a precipitação sazonal: 09 variáveis selecionadas.	109
B.45	Resultado do quinto ciclo do método <i>stepwise</i> para a precipitação sazonal: 11 variáveis selecionadas.	109

LISTA DE TABELAS

3.1	Localização e variáveis preditoras possíveis para previsão de variação climática do Paraná..	38
3.2	Estatísticas calculadas a partir dos dados de precipitação e temperatura mensais.	41
3.3	Meses selecionados que tem os maiores valores absolutos de correlação cruzada com a temperatura e precipitação mensal.	42
3.4	Meses selecionados que tem os maiores valores absolutos de correlação cruzada com a temperatura e precipitação médias trimestrais.	43
4.1	Desempenho da previsão climática por persistência..	47
4.2	Desempenho da previsão climática por redes neurais de contra-propagação. . . .	54
4.3	Desempenho da previsão climática por redes neurais RBF.	62
4.4	Comparação dos resultados obtidos com diferentes técnicas para a previsão climática do Paraná..	64

LISTA DE ACRÔNIMOS

AMP	Anticiclone Móvel Polar
AS	Atlântico Sul
ASAS	Anticiclone Subtropical do Atlântico Sul
ASPS	Anticiclone Subtropical do Pacífico Sul
BCH	Baixa de Chaco
CCM	Complexo Convectivos de Mesoescala
CNS	Coeficiente de Nash-Sutcliffe
CP	Componentes Principais
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos
CV	Coeficiente de Variação
DP	Desvio Percentual
DPa	Desvio Padrão
ECMWF	<i>European Centre for Medium-Range Weather Forecasts</i>
EMA	Erro Médio Absoluto
ENOS	El Niño-Oscilação Sul
INMET	Instituto Nacional de Meteorologia
INPE	Instituto Nacional de Pesquisas Espaciais
IOS	Índice de Oscilação Sul
JBN	Jato de Baixo Nível
JP	Jato Polar
JS	Jato Subtropical
MCGA	Modelo de Circulação Geral da Atmosfera
MCTIC	Ministério da Ciência, Tecnologia, Inovação e Comunicação
MLP	<i>Multilayer Perceptron</i>
NCAR	<i>National Center for Atmospheric Research</i>
NCEP	<i>National Centers for Environmental Prediction</i>
NLPCA	<i>Nonlinear Principal Component Analysis</i>
PCA	<i>Principal Component Analysis</i>
OQB	Oscilação Quase-Bienal
RBF	<i>Radial Basis Function</i>
REQM	Raiz do Erro Quadrático Médio
RNA	Rede Neural Artificial
TSM	Temperatura da Superfície do Mar
ZCAS	Zona de Convergência do Atlântico Sul
ZCIT	Zona de Convergência Intertropical

LISTA DE SÍMBOLOS

Σ	Combinador linear da RNA
a	Taxa de aprendizado de saída da camada de clusterização
α	Taxa de aprendizado de entrada da camada de clusterização
b	Taxa de aprendizado de saída da camada de clusterização
β	Taxa de aprendizado de entrada da camada de clusterização
D_j	Distância Euclidiana entre os vetores de entrada e os pesos sinápticos
f	Função de ativação da RNA
gp	Altura geopotencial
J	Índice do neurônio vencedor
prec	Precipitação total
pres	Pressão na superfície
q	Umidade específica
r	Coeficiente linear de Pearson
t_{Ji}	Pesos sinápticos entre a camada de clusterização Z e a camada de saída X^*
temp2	Temperatura do ar a 2 metros
u	Vento zonal
u_{Jk}	Pesos sinápticos entre a camada de clusterização Z e a camada de saída Y^*
v	Vento meridional
v_{ij}	Pesos sinápticos entre a camada de entrada X e a camada de clusterização Z
w	Velocidade vertical
w_{kj}	Pesos sinápticos entre a camada de entrada Y e a camada de clusterização Z
x	Sinal de entrada da RNA
x^*	Sinal de entrada produzido pela RNA
x'	Vetor das componentes principais de x
X	Camada de entrada X
X^*	Camada de saída X
y	Variável de interesse
y^*	Sinal de saída da RNA
Y	Camada de entrada Y
Y^*	Camada de saída Y
Z	Camada de clusterização
λ	Variância de x'

SUMÁRIO

1	INTRODUÇÃO	17
2	FUNDAMENTAÇÃO TEÓRICA E REVISÃO BIBLIOGRÁFICA	19
2.1	CLIMATOLOGIA	19
2.2	CLIMA DO PARANÁ	19
2.3	PREVISÃO CLIMÁTICA	23
2.4	REDES NEURAIS ARTIFICIAIS	24
2.4.1	Redes de contra-propagação	27
2.4.2	Algoritmo da rede de contra-propagação	29
2.4.3	Redes de Função de Base Radial	31
2.4.4	Algoritmo da rede RBF	34
2.5	MÉTODO DE SELEÇÃO DE VARIÁVEIS: REGRESSÃO <i>STEPWISE</i>	35
3	MATERIAIS E MÉTODOS	36
3.1	BASE DE DADOS CLIMATOLÓGICOS	36
3.2	DESENVOLVIMENTO DO MODELO DE RNA	42
3.3	ESTRUTURAS DAS RNAs DE CONTRA-PROPAGAÇÃO E RBF	44
3.4	PREVISÃO POR PERSISTÊNCIA	45
3.5	AVALIAÇÃO DO DESEMPENHO DAS PREVISÕES	45
4	EXPERIMENTOS DE PREVISÃO DE TEMPERATURA E PRECIPITAÇÃO	47
4.1	PREVISÃO ATRAVÉS DE PERSISTÊNCIA	47
4.2	PREVISÃO ATRAVÉS DA REDE NEURAL DE CONTRA-PROPAGAÇÃO	47
4.2.1	Seleção das variáveis de entrada do modelo de temperatura mensal	47
4.2.2	Seleção das variáveis de entrada do modelo de temperatura sazonal	49
4.2.3	Seleção das variáveis de entrada do modelo de precipitação mensal	50
4.2.4	Seleção das variáveis de entrada do modelo de precipitação sazonal	52
4.2.5	Avaliação do desempenho das previsões pela rede neural de contra-propagação	54
4.3	PREVISÃO ATRAVÉS DA REDE NEURAL RBF	56
4.3.1	Seleção das variáveis de entrada do modelo de temperatura mensal	56
4.3.2	Seleção das variáveis de entrada do modelo de temperatura sazonal	58
4.3.3	Seleção das variáveis de entrada do modelo de precipitação mensal	59
4.3.4	Seleção das variáveis de entrada do modelo de precipitação sazonal	60
4.3.5	Avaliação do desempenho das previsões pela rede neural RBF	61
4.4	DISCUSSÃO DOS RESULTADOS	64

5	CONCLUSÃO	66
6	RECOMENDAÇÕES PARA TRABALHOS FUTUROS.	68
	REFERÊNCIAS	69
	APÊNDICE A – MÉDIAS CLIMATOLÓGICAS E ANOMALIAS DAS VARIÁ- VEIS PREDITORAS DA RNA.	72
A.1	ANTICICLONE SUBTROPICAL DO ATLÂNTICO SUL	72
A.2	ANTICICLONE SUBTROPICAL DO PACÍFICO SUL	74
A.3	ANTICICLONE MÓVEL POLAR	77
A.4	EL NIÑO OSCILAÇÃO SUL	79
A.5	JATO DE BAIXOS NÍVEIS.	80
A.6	JATO POLAR	82
A.7	JATO SUBTROPICAL	83
A.8	ZONA DE CONVERGÊNCIA DO ATLÂNTICO SUL	85
A.9	PARANÁ.	88
	APÊNDICE B – RESULTADOS DO MÉTODO <i>STEPWISE</i>	94
B.1	REDE NEURAL DE CONTRA-PROPAGAÇÃO	94
B.2	REDE NEURAL RBF	103

1 INTRODUÇÃO

As previsões climáticas são de grande importância para diversos setores da sociedade, principalmente para aqueles sensíveis às variações atmosféricas, como a agricultura, geração de energia, disponibilidade hídrica, etc. Além disso, a possibilidade de prever eventos extremos (por exemplo, secas drásticas, tempestades e deslizamentos de terra) permite que decisões possam ser tomadas para que os impactos de tais eventos sejam minimizados.

Neste ano, o Paraná vem sofrendo com uma severa estiagem que trouxe diversos transtornos à população e mudou o cenário de um dos principais pontos turísticos do estado, as Cataratas do Iguaçu. Já em 2011, as chuvas acima da média no verão favoreceram deslizamentos de terras na Serra do Mar e enchentes em cinco cidades do Litoral paranaense. Além disso, a economia do Paraná, fortemente relacionada à agricultura, faz com que sejam necessárias previsões climáticas com uma boa acurácia.

Vários métodos estão disponíveis para a previsão climática, como modelos numéricos, métodos estatísticos e técnicas de aprendizado de máquina. Uma das técnicas de aprendizado de máquina são as Redes Neurais Artificiais (RNAs), que têm sido amplamente utilizadas na previsão climática em virtude da capacidade de identificar relações não-lineares entre as variáveis de entrada e saída sem a necessidade de entender a natureza dos processos físicos (Lee et al., 2018).

Diversos estudos sobre previsão climática foram publicados usando as RNAs. Nezhad et al. (2019) desenvolveram um modelo para prever a máxima temperatura no inverno em Teerã, Irã. Também na região de Teerã, Gholizadeh e Darand (2009) usaram uma RNA combinada com um algoritmo genético para prever a precipitação mensal do próximo ano utilizando uma série histórica de 53 anos. Bodri e Čermák (2000) empregaram as técnicas de aprendizado de máquina para a previsão de precipitação mensal e precipitação de verão para o próximo ano, usando dados de duas estações meteorológicas na República Checa. No Brasil, Anochi e de Campos Velho (2016) utilizaram uma rede neural autoconfigurada para as previsões mensal e sazonal de precipitação na região Sul do Brasil. Esses estudos mostram que vem crescendo o interesse da comunidade científica em utilizar as RNAs para prever as variáveis climáticas em várias escalas temporais e em regiões áridas a úmidas.

Um dos passos mais importantes no desenvolvimento de um modelo satisfatório de previsão por redes neurais é a seleção de variáveis de entrada apropriadas. Portanto, os trabalhos científicos possuem diferentes metodologias para selecionar os vetores de entrada. Anochi e de Campos Velho (2016) empregaram dados meteorológicos mensais da mesma região que foram realizados as previsões mensal e sazonal de chuvas, no sul do Brasil. As entradas utilizadas no modelo de RNA nesse trabalho foram vento zonal em 300, 500 e 850 hPa, vento meridional em 300, 500 e 850 hPa, temperatura na superfície, umidade específica e precipitação. Para realizar a previsão de chuvas no final da primavera e início do verão na bacia do rio Geum na Coreia do Sul, Lee et al. (2018) utilizaram índices climáticos que representam padrões de teleconexão (tais como Oscilação do Ártico, Oscilação do Pacífico Leste / Pacífico Norte, Oscilação do Atlântico Norte, Índice do Pacífico Oeste, Oscilação Decadal do Pacífico, Oscilação do Sul, Pressão no Nível do Mar em Darwin, entre outros) como entradas para o modelo. Outros estudos utilizam a própria variável que pretende-se prever como entrada para a RNA, como o trabalho de Khalili et al. (2016), que usaram como vetores de entrada a precipitação do mês anterior, do ano anterior para o mês estimado e a média de chuvas de longo prazo para o mês previsto.

Sabendo-se da importância da previsão climática e do crescente uso das RNA no campo dos estudos atmosféricos, este trabalho tem como objetivo geral avaliar o desempenho das redes neurais artificiais de contra-propagação (CPN) e de função de base radial (RBF), respectivamente, na previsão climática de temperatura e precipitação no estado do Paraná. Para a execução do objetivo principal do trabalho foram analisados os fenômenos atmosféricos que podem causar variação climática no Paraná e, também, foi avaliada a destreza dos modelos de RNAs para a previsão mensal e sazonal da precipitação e temperatura. Enquanto as estimativas mensais do clima são capazes de fazer projeções a mais curto prazo, a previsão sazonal é capaz de identificar as anomalias a mais longo prazo que são importantes para verificação da disponibilidade hídrica, por exemplo. Portanto, podem-se citar ainda como objetivos específicos do estudo:

- Identificar quais variáveis influenciam de maneira mais significativa no desempenho da rede neural na previsão climática do Paraná. Juntamente, determinar qual período dos fenômenos atmosféricos antecedentes são importantes para o bom desempenho da RNA;
- Avaliar a destreza dos modelos de RNAs para a previsão mensal e sazonal da precipitação e temperatura do Paraná e
- Realizar uma técnica simples de previsão, que é a persistência, para verificação do desempenho da utilização das RNAs de CPN e RBF.

Para a realização desses objetivos, este trabalho foi dividido em 5 capítulos incluindo esta introdução. O capítulo dois é dedicado à fundamentação teórica e revisão bibliográfica, onde são detalhados os conceitos da climatologia, previsão climática, redes neurais e apresentado o métodos de redução de dimensionalidade *stepwise*. No terceiro capítulo é descrita a base de dados utilizada, a estrutura e desenvolvimento do modelo de RNA, a previsão por persistência e as métricas para avaliação do desempenho das previsões. O capítulo quatro é dedicado ao experimentos de previsão de temperatura e precipitação. Por fim, o último capítulo apresenta as principais conclusões do estudo desenvolvido.

2 FUNDAMENTAÇÃO TEÓRICA E REVISÃO BIBLIOGRÁFICA

2.1 CLIMATOLOGIA

Na ciência atmosférica usualmente é feita uma distinção entre tempo e clima, e entre meteorologia e climatologia. O tempo (em inglês, *weather*) entende-se como o estado da atmosfera em determinado lugar, por um curto período de tempo. Refere-se então a tempo, não clima, de hoje ou da última semana. Por outro lado, clima refere-se ao estado médio da atmosfera e é usado em referência à variabilidade atmosférica em escalas de tempo maiores.

A meteorologia é definida como a ciência que estuda os fenômenos que ocorrem na atmosfera, e está relacionada ao estado físico, dinâmico e químico da atmosfera, as interações entre elas e a superfície terrestre subjacente. A climatologia é o estudo científico do clima. Enquanto o meteorologista emprega as leis da física clássica e as técnicas matemáticas em seu estudo de processos atmosféricos, o climatólogo utiliza, principalmente, técnicas estatísticas a partir das informações disponíveis sobre o tempo para se ter conhecimento a respeito do clima (Ayoade e dos Santos, 2002). Outra diferença fundamental entre meteorologia e climatologia está nas perspectivas de tempo que apresentam para o estudo do tempo e do clima. Enquanto os meteorologistas produzem previsões destinadas a antecipar as condições atmosféricas a curto e médio prazos, geralmente em um horizonte de 7 a 10 dias, os climatologistas, por sua vez, empregam uma perspectiva de longo prazo, desenvolvendo e analisando modelos projetados para prever mudanças nos padrões climáticos nos próximos meses, próxima estação ou nos próximos anos. A seção 2.3 será destinada a mostrar os principais métodos de previsão climática.

Em relação à escala espacial, o clima é classificado em macro, meso e microclima. O macroclima está relacionado aos aspectos climáticos de amplas áreas da Terra e com os movimentos atmosféricos em larga escala (por exemplo as ondulações planetárias com escala horizontal maior que 10^4 km e as perturbações sinóticas com extensão horizontal na ordem de 10^3 km). O mesoclima abrange áreas relativamente pequenas, entre 10 e 100 quilômetros de extensão (por exemplo, sistemas climáticos locais severos tais como os tornados e os temporais). Já o microclima refere-se ao clima próximo à superfície ou de áreas muito pequenas, com menos de 100 metros de extensão (Ayoade e dos Santos, 2002). O presente trabalho tem por objetivo estudar as anomalias do clima do estado do Paraná, que possui uma grande área territorial, portanto enquadrando-se no estudo do macroclima.

2.2 CLIMA DO PARANÁ

O Paraná encontra-se em uma região de transição climática entre o clima tropical e subtropical, no entanto, a maior parte do território do estado do Paraná encontra-se na zona subtropical, segundo a classificação de Köppen, com duas variações: clima subtropical com verões quentes (Cfa) e clima temperado com verões amenos (Cfb). Para Neto (2010), os dois tipos climáticos (Cfa e Cfb) são descritos por:

- Cfa - Clima subtropical - temperatura média no mês mais frio abaixo de 18°C e temperatura média no mês mais quente acima de 22°C , geadas pouco frequentes e tendência de concentração das chuvas nos meses de verão, contudo sem estação seca definidas. Localiza-se no planalto norte e centro-leste do Paraná;

- Cfb - Clima temperado - na classificação Cfb, a temperatura média no mês mais frio fica abaixo de 18°C e a estação seca não é definida, assim como na classificação Cfa. No entanto, a temperatura média no mês mais quente é inferior a 22°C, o que caracteriza um clima temperado com verões amenos. Esse clima ocorre na região do primeiro planalto, onde localiza-se Curitiba; na região do segundo planalto, onde se encontra Ponta Grossa; e nas regiões mais altas do terceiro planalto, como na cidade de Palmas.

Em seguida, serão apresentados alguns fatores que alteram os sistemas de pressão e ventos, produzindo mudanças nos quadros médios sazonais, que são as variações climáticas. Embora o interesse desta seção seja estudar os aspectos do clima do Paraná, são apresentados os campos atmosféricos para uma área maior, pois necessitam ser colocados em contexto mais amplo.

O continente sul-americano aparece entre dois sistemas de alta pressão quase estacionários do Atlântico Sul e do Pacífico Sul, aos quais estão associadas circulações anticiclônicas e subsidentes. O Anticiclone Subtropical do Atlântico Sul (ASAS) é a principal influência sobre os ventos de superfície na região sul do Brasil. Este padrão está presente em todas as estações do ano, no entanto é mais forte no inverno, quando está mais para o norte e para o oeste, entrando no continente. No verão, o centro de alta pressão está deslocado mais para leste e para o sul (Grimm, 2009).

Os movimentos sazonais do centro de alta pressão no Atlântico Sul determinam a maior ou menor incidência de ventos em baixos níveis na costa e, conseqüentemente, maior ou menor precipitação. No verão, devido ao afastamento do ASAS do continente, a entrada de ventos cresce na região sul, enquanto no inverno cresce mais na região Nordeste do Brasil, sendo insignificante no sul (Grimm, 2009).

Para Nimer (1989), outro centro de pressão anticiclônico é a alta polar, conhecido por anticlone polar móvel da América do Sul. Essa alta, ao invadir a região sul, traz consigo um estado de tempo instável, com a importante corrente perturbada que age sobre a região sul do Brasil, que é a frente polar.

Um importante sistema de pressão para o sul do Brasil é um centro de baixa pressão intermitente no noroeste da Argentina, Paraguai e sul da Bolívia, nomeado Baixa do Chaco (BCH). É um sistema quente, menos intenso no inverno, localizado na baixa troposfera (até 700 hPa) e com frequência é acompanhado por subsidência em altos níveis, que causa ausência de nebulosidade nesta região. A Baixa do Chaco estende-se e aprofunda-se no verão, fortalecendo o gradiente zonal subtropical de pressão e, com isso, a componente meridional do vento, ajudando a fortalecer os ventos de noroeste em baixos níveis que conectam os trópicos com a região sul (Grimm, 2009).

A intensificação da BCH parece ter um papel muito importante não só na modificação da circulação atmosférica na baixa troposfera, como também no desenvolvimento dos Jatos de Baixos Níveis (JBN) (Seluchi e Garreaud, 2012). Os JBNs são escoamentos intensos na baixa troposfera, cuja velocidade máxima ocorre entre 1 a 2 km de altura (aproximadamente 850 hPa). Os JBNs estão presentes em diferentes regiões do globo e na América do Sul encontra-se a oeste e leste da cordilheira dos Andes. Com relação aos JBNs a leste dos Andes, esses são caracterizados por um escoamento meridional que se dirige da Amazônia para os subtrópicos da América do Sul, podendo adentrar próximo da fronteira oeste da região sul, produzindo muita chuva e vento.

Na alta troposfera, os ventos médios são predominante de oeste e existe uma região onde a componente do vento zonal atinge valores máximos. Essa componente é intensificada pelos contrastes de temperatura da superfície que produzem maiores gradientes de pressão em altitude. Esses escoamentos de ar, caracterizados por valores máximos na alta troposfera,

são denominados correntes de jato, sendo que cada hemisfério possui duas correntes de jato, denominadas jato polar (JP) e jato subtropical (JS).

O JP está associado à frente polar, devido aos grandes contrastes de temperatura que ocorrem ao longo dessas frentes. Esse jato é mais forte durante o inverno e pode desviar-se em direção ao equador, até 30° de latitude. Durante o verão, sua posição média é usualmente em torno de 50° de latitude. Já o JS ocorre em torno de 25° de latitude, no extremo da célula de Hadley. Em termo médio, ele é mais forte e menos variável em latitude que o JP.

Na baixa troposfera ocorre a formação da Zona de Convergência do Atlântico Sul (ZCAS). A ZCAS é caracterizada por uma faixa de nebulosidade convectiva que se estende da região Amazônica até a região central do Atlântico Sul, podendo variar sua posição norte-sul. O período de ocorrência mais intensa é durante os meses de primavera e verão, sendo um sistema extremamente importante para a estação chuvosa do sudeste do Brasil, frequentemente associada a sérios transtornos à população, devido aos grandes volumes pluviométricos. Exemplos de atuação com maior intensidade da ZCAS que ocasionaram tragédias à população ocorreram em 2011 na região Serrana do Rio de Janeiro, com o deslizamento de encostas e em 2020 com chuvas fortes e volumosas na capital de Minas Gerais.

Em parte do Paraná a ZCAS pode causar chuvas, uma vez que esta região localiza-se próxima da sua posição climatológica, mas no restante do sul do Brasil inibe a precipitação devido ao seu ramo subsidente (Reboita et al., 2010).

Tanto durante o verão quanto nas estações de transição (ou seja, no semestre quente, de outubro a abril), os Complexos Convectivos de Mesoescala (CCM) são frequentes e respondem por grande parte da precipitação total, especialmente nas estações de transição, da região sul (Grimm, 2009). O CCM é um sistema com espessa cobertura de nuvens frias, que podem ser facilmente identificados em imagens de satélites pelo formato de nuvens circular e com um crescimento explosivo em um intervalo de tempo de 6 a 12 horas. O elemento principal para formação dos CCMs nos extratropicais da América do Sul (AS) é a presença de um jato em baixos níveis, úmido e quente, vindo de norte que interage com um jato subtropical em altos níveis vindo de oeste (Velasco e Fritsch, 1987 apud Reboita et al., 2010).

A Figura 2.1 apresenta, de forma ilustrativa, a localização dos sistemas atmosféricos atuantes na AS, que podem influenciar o tempo e o clima no Paraná, considerando-se a baixa e alta troposferas.

O regime de chuvas no sul do Brasil sofre grandes influências remotas da temperatura dos Oceanos Atlântico e Pacífico (Cera e Ferraz, 2015; Grimm et al., 1998). O El Niño / Oscilação Sul (ENSO) é um dos fenômenos naturais caracterizado por anomalias de Temperatura da Superfície do Mar (TSM) nos oceanos do Pacífico equatorial central e oriental, acoplado com mudanças na atmosfera. O El Niño / Oscilação Sul compreende três fases: El Niño, La Niña e neutro. Os eventos do El Niño geralmente começam no meio do ano com aquecimento em larga escala de águas superficiais no Oceano Pacífico equatorial central e oriental e mudanças na circulação atmosférica (ou seja, no padrões climáticos de ventos, pressão e chuvas). Em geral, El Niño atinge um pico nos meses de novembro a janeiro e depois decai durante a primeira metade do ano seguinte. Esse evento ocorre a cada dois a sete anos e pode durar até 18 meses. O oposto de El Niño é conhecido como La Niña, que refere-se ao resfriamento em larga escala da TSM do oceano na mesma região do Pacífico equatorial, juntamente com uma inversão das condições atmosféricas subjacentes. Durante as fases neutras, as temperaturas no oceano pacífico estão mais próxima das médias climatológicas, portanto, os padrões atmosféricos são mais controlados por outros fatores climáticos (World Meteorological Organization, 2014).

A Figura 2.2 mostra a anomalia da TSM de um evento de La Niña que ocorreu em dezembro de 1988, e de um evento de El Niño em dezembro de 1997.

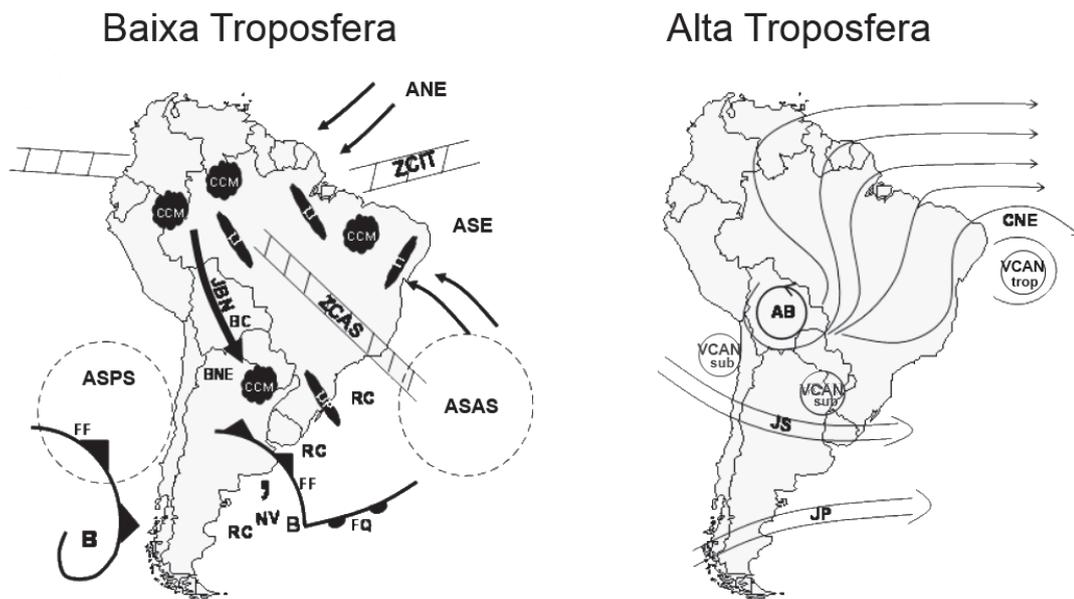


Figura 2.1: Sistemas atmosféricos na baixa e alta troposfera atuantes na América do Sul. Fonte: Adaptado de Satyamurty et al. (1998).

Em relação à baixa troposfera tem-se: **ANE** ventos alísios de nordeste, **ASE** ventos alísios de sudeste, **ASAS** anticiclone subtropical do Atlântico Sul, **ASPS** anticiclone subtropical do Pacífico Sul, **B** baixa pressão, **BC** baixa do Chaco, **BNE** região de baixas térmicas no noroeste argentino, **CCM** complexo convectivo de mesoescala, **FF** frente fria, **FQ** frente quente, **JBN** jato de baixos níveis a leste dos Andes, **LI** linha de instabilidade tropical, **LIP** linha de instabilidade pré-frontal, **NV** nuvem vírgula, **RC** regiões ciclogênicas, **ZCAS** zona de convergência do Atlântico Sul e **ZCIT** zona de convergência intertropical. Em relação à alta troposfera tem-se: **AB** alta da Bolívia, **CNE** cavado do nordeste do Brasil, **JS** jato subtropical, **JP** jato polar, **VCAN sub** vórtices ciclônicos de altos níveis subtropicais, **VCAN trop** vórtices ciclônicos de altos níveis tropicais.

As flutuações na temperatura do oceano durante o El Niño e a La Niña são acompanhadas por flutuações de larga escala na pressão do ar, conhecido por Oscilação Sul. Nos períodos mais quentes do oceano, a temperatura do mar é superior à temperatura do ar e favorece os fluxos de calor sensível e latente do oceano para a atmosfera. Nessas condições, ocorre a transferência de umidade para cima, e uma convecção vigorosa leva a umidade ao nível de condensação e são formadas nuvens de chuva. Em períodos muito mais longos de oceano frio, o ar é mais quente do que o mar e, em média, não existe fluxo ascendente de calor e umidade em toda a interface (Bjerknes, 1969). O fluxo ascendente no período de El Niño gera um centro de baixa pressão cobrindo o oceano Pacífico central e oriental e o centro de alta pressão desloca-se para a Indonésia, e essa diferença de pressão ao nível do mar entre o Pacífico Central (Taiti) e o Pacífico Oeste (Darwin/Austrália) é quantificada pelo Índice de Oscilação Sul (IOS). Em episódios de El Niño tem-se valores de IOS negativo, sendo denominados por fase negativa da Oscilação Sul. Já a fase positiva da Oscilação Sul ocorre durante os episódios de La Niña, quando há um centro de baixa pressão cobrindo a Indonésia e um centro de alta pressão sobre o oceano Pacífico central e oriental.

Grimm et al. (1998) avaliaram a influência dos eventos de El Niño e La Niña na anomalia de precipitação no sul do Brasil. Os autores verificaram picos positivos de anomalias durante a primavera, principalmente em novembro, sobre todo o sul do Brasil e associaram este aumento na precipitação à uma intensificação dos Complexos Convectivos de Mesoescala nesta região.

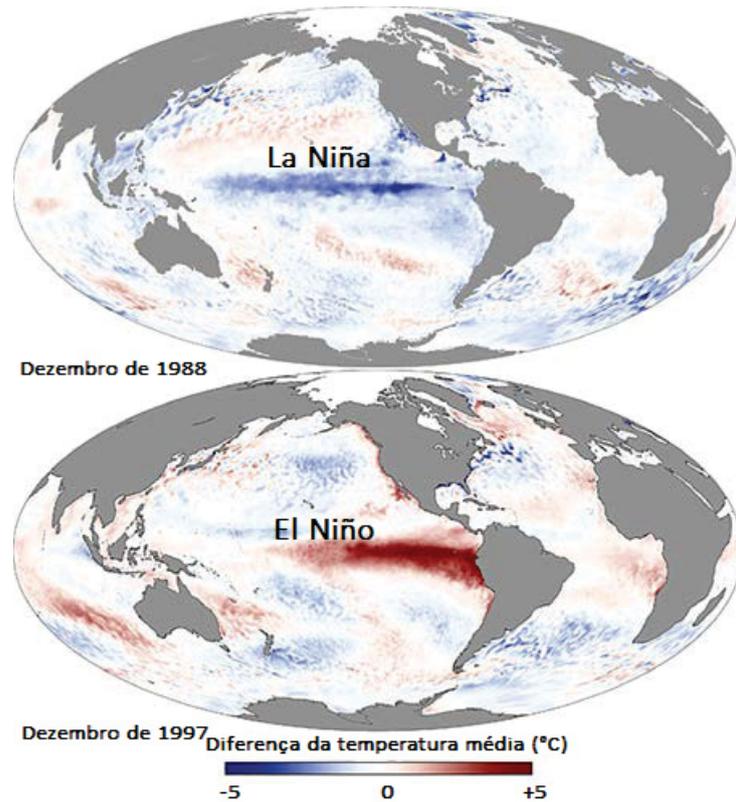


Figura 2.2: Anomalia da temperatura média do oceano Pacífico em dezembro de 1988 e dezembro de 1997 correspondendo a um evento de La Niña e El Niño, respectivamente.
 Fonte: Adaptado de World Meteorological Organization (2014)

2.3 PREVISÃO CLIMÁTICA

Dentre os avanços científicos que ocorreram no último século está a capacidade em simular sistemas físicos complexos, como a atmosfera, utilizando modelos numéricos que estabelecem uma representação matemática aproximada da realidade (Sampaio e da Silva Dias, 2014). Um dos métodos utilizados para prever alguns padrões de larga escala do clima sobre diferentes regiões do globo são os Modelos de Circulação Geral da Atmosfera (MCGAs). Os MCGAs podem simular os parâmetros climáticos apenas nos pontos de grade, com baixa resolução espacial, quando comparado com os modelos de previsão meteorológica de curto e médio prazo, sendo necessárias as técnicas de regionalização (*downscaling*) para operar modelos climáticos regionais (Alves et al., 2004).

Devido à natureza caótica da atmosfera, os MCGAs e os modelos climáticos regionais têm um limite teórico de duas semanas de confiabilidade nas previsões de tempo (Shukla et al., 2000). Entretanto, esses modelos também podem realizar previsões atmosféricas com meses de antecedência, ou seja, previsões climáticas. As previsões numéricas de clima derivam, principalmente, da previsibilidade das condições de contorno, sobretudo da temperatura da superfície do mar (TSM), ou seja, a variabilidade climática sazonal é controlada principalmente pelas lentas variações das temperaturas dos oceanos (Sampaio e da Silva Dias, 2014). Outro fator que permite utilizar os modelos dinâmicos para previsão a longo prazo, como no caso da previsão sazonal, é que não será avaliado em qual dia uma localidade acontecerá uma chuva, tempestade ou temperaturas extremas, por exemplo, mas serão estimados os desvios e anomalias no comportamento sazonal dos elementos do clima.

No entanto, mesmo se a TSM pudesse ser prevista sem erro, não seria possível prever exatamente a evolução das variáveis atmosféricas por causa da natureza caótica da dinâmica atmosférica (Branković e Palmer, 1997). Além disso, as incertezas podem ser decorrentes do próprio modelo, como parametrizações físicas, acurácia numérica e das limitações matemáticas na descrição dos fenômenos atmosféricos.

Uma forma de melhorar a previsão climática é utilizar a técnica de previsão por conjunto (em inglês, *ensemble forecasting*). Como a previsibilidade está relacionada às condições iniciais do modelo, a técnica de previsão por conjunto consiste em produzir um conjunto de previsões a partir de possíveis estados iniciais (Sampaio e da Silva Dias, 2014). A forma mais simples de utilizar as previsões por conjunto é através do conjunto médio, que representa a melhor estimativa do estado futuro da atmosfera. A previsão do conjunto médio apresenta um desempenho melhor em relação às previsões individuais do conjunto, visto que alguns erros individuais são cancelados na média (Tracton e Kalnay, 1993 apud Machado et al., 2010).

Embora o avanço tecnológico tenha permitido computadores mais potentes, ainda é muito caro para desenvolver, integrar e validar modelos numéricos. Uma alternativa para a modelagem numérica nos estudos diagnósticos e de previsão de variáveis climáticas são os modelos estatísticos (Alves et al., 2007). A modelagem estatística depende da qualidade e quantidade de observações históricas do oceano e da atmosfera, mas geralmente são menos dispendiosos de desenvolver e executar (Anderson et al., 1999).

No Brasil, por exemplo, o Instituto Nacional de Meteorologia (INMET) vem desenvolvendo e utilizando um modelo estatístico de previsão climática. O INMET produz, mensalmente, previsões sazonais de precipitação total e temperatura média para todo o país. As previsões são geradas utilizando os modelos de Regressão Linear Múltipla e Análise de Correlação Canônica, que consistem em técnicas de análise multivariadas. As previsões deste modelo têm sido fornecidas como subsídio para a Reunião Climática mensal organizada pelo Ministério da Ciência, Tecnologia, Inovação e Comunicação (MCTIC) (INMET, 2019). Nessas reuniões são analisados outros modelos numéricos e estatísticos de previsão climática sazonal e é formado a previsão por consenso, que fica disponibilizada no portal do INPE/CPTEC (<https://www.cptec.inpe.br/>). As previsões são representadas em mapas que determinam a probabilidade que a variável meteorológica tem de ficar em uma das três categorias previstas, a saber: acima, dentro ou abaixo da média histórica.

As técnicas de aprendizado de máquina, como Redes Neurais Artificiais (RNAs), têm sido amplamente utilizadas na previsão climática, substituindo os modelos estatísticos lineares, em virtude da capacidade desses sistemas em identificar relações não lineares altamente complexas entre variáveis de entrada e saída sem a necessidade de entender a natureza dos processos físicos (Lee et al., 2018). As RNAs, que foi o método utilizado para a previsão mensal e média trimestral no estado do Paraná, serão apresentadas na próxima seção.

2.4 REDES NEURAIAS ARTIFICIAIS

Redes Neurais Artificiais (RNAs) são modelos computacionais inspirados no sistema nervoso de seres vivos, que possuem a capacidade de “aprender” padrões através de treinamento. A principal vantagem de uma rede neural é a capacidade de aproximar relações funcionais, especialmente quando as relações não são bem definidas e/ou são não-lineares, o que torna difícil a utilização de métodos estatísticos convencionais para a tentativa de prever as variações futuras dessas relações. Portanto, RNAs são métodos úteis para resolver muitos tipos de problemas, especialmente nas questões que envolvem relações entre variáveis na natureza, que são quase sempre complexas e altamente não lineares.

A célula elementar do sistema nervoso central é o neurônio e seu papel se resume a conduzir impulsos (estímulos elétricos proveniente de reações físico-químicas) sob determinadas condições de operação. O neurônio, representado na Figura 2.3, pode ser dividido em três partes principais, que são, os dendritos, o corpo celular (também conhecido como soma) e o axônio.

A principal função dos dendritos é captar, de forma contínua, os estímulos vindos de outros neurônios (conectores) ou do próprio meio externo (neurônios sensitivos). O corpo celular é responsável por processar todas as informações vindas dos dendritos a fim de produzir um potencial de ativação que indicará se o neurônio deverá disparar um impulso elétrico ao longo de seu axônio. A função do axônio é conduzir os impulsos elétricos para outros neurônios conectores ou para aqueles que se conectam diretamente com o tecido muscular (neurônios efetadores). A terminação do axônio é formada por ramificações denominadas terminações sinápticas (Silva et al., 2010).

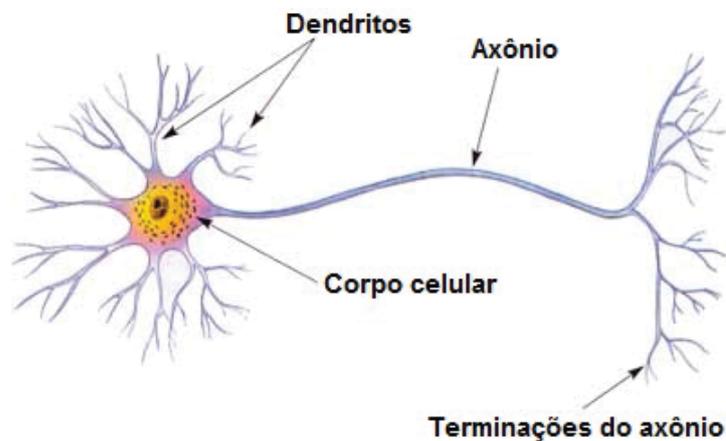


Figura 2.3: Representação de um neurônio biológico.
Fonte: Elaborada pela autora.

O neurônio artificial, representado na Figura 2.4, é um modelo simplificado do neurônio biológico, e é dividido em: sinais de entradas, pesos sinápticos, vies (*bias*), combinação linear, função de ativação e saída.

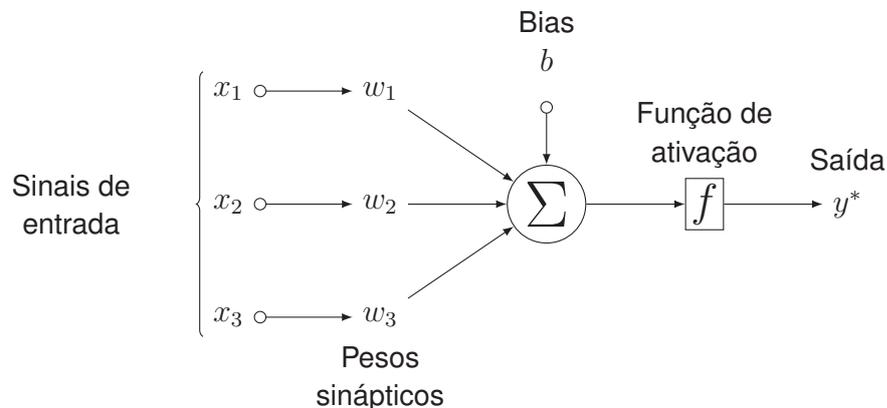


Figura 2.4: Representação de um neurônio artificial.
Fonte: Elaborada pela autora.

Os sinais de entrada, apresentados pelo conjunto (x_1 , x_2 e x_3), neste exemplo, são análogos aos impulsos elétricos externos captados pelos dendritos no neurônio biológico. Os valores dos pesos sinápticos (w_1 , w_2 e w_3) servem para ponderar cada uma das variáveis de

entrada da rede. O *bias* é um elemento que serve para aumentar o número de graus de liberdade dos ajustes dos pesos.

Os sinais de entrada que foram ponderados pelos respectivos pesos sinápticos junto com os *biases* são combinados e geram um sinal de saída da combinação linear (\sum). A função de ativação (f) limita a saída do neurônio dentro de um intervalo de valores que constituem seu conjunto de imagem. A saída (y^*) é o sinal de saída final produzido pelo neurônio artificial e é representado matematicamente pela Equação 2.1.

$$y^* = f\left(\sum_{i=1}^n w_i x_i + b\right) \quad (2.1)$$

O processo de aprendizado de uma rede neural é realizado na fase de treinamento, por meio do ajuste de seus pesos sinápticos. O treinamento pode ser classificado como supervisionado ou não supervisionado. No treinamento supervisionado são apresentados à RNA os valores de entrada e de saída (referência). Desta forma, os pesos sinápticos são ajustados pelo algoritmo de aprendizado a fim de atingir o valor de referência. No treinamento não-supervisionado não são apresentados os valores de saída, apenas de entrada. Assim, a rede modifica seus pesos sinápticos para conseguir associar os vetores de saída aos vetores de entrada, identificando subconjuntos (*clusters*) que contenham similaridade.

Em relação à organização das camadas, as redes neurais são classificadas como de camada simples ou multicamadas. Na arquitetura de RNAs de camada simples tem-se apenas uma camada de neurônios de entrada e uma camada de neurônios de saída, Assim, haverá apenas uma camada de conexões entre neurônios. Já a arquitetura de redes multicamadas são constituídas pela presença de uma ou mais camadas oculta ou escondidas, posicionadas entre as camadas entrada e saída. Quanto à propagação do sinal, as RNAs que enviam o sinal dos neurônios sempre adiante, para a próxima camada de neurônios, são conhecidas como *feedforward*, e as redes em que as saídas dos neurônios são realimentadas como sinais de entrada para outros neurônios são chamadas de recorrentes ou realimentadas (em inglês, *backward*).

Algumas das estruturas mais comuns são as seguintes:

- Perceptron: arquitetura *feedforward* de camada simples e processo de treinamento baseado na regra de Hebb;
- Perceptron multicamadas (*Multilayer Perceptron - MLP*): arquitetura *feedforward* de multicamadas e algoritmo de aprendizado baseado na regra delta generalizada;
- Redes de função de base radial (*Radial Basis Function - RBF*): arquitetura *feedforward* de multicamadas, com os métodos de aprendizado competitivo e regra delta generalizada;
- Redes de contra-propagação (*Counterpropagation Network - CPN*): arquitetura *feedforward* de multicamadas, com os métodos de aprendizado competitivo e arranjo de Grossberg e
- Redes de Hopfield: arquitetura multicamadas com realimentação entre neurônios situados em camadas distintas, cujo algoritmo de treinamento é baseado na minimização de funções energia.

A seguir, serão apresentadas em mais detalhes as redes de contra-propagação (CPN) e de função de base radial (RBF).

2.4.1 Redes de contra-propagação

A rede de contra-propagação (CPN), idealizada por Hecht-Nielsen (1987), pode ser utilizada tanto em classificação de padrões como em problemas relacionados à aproximação de funções. A arquitetura da rede CPN é do tipo *feed-forward* de multicamadas, sendo uma camada de entrada, uma camada oculta e uma de saída. O treinamento é constituído por duas etapas. A primeira etapa, associada com os ajustes dos pesos dos neurônios da camada oculta, adota um método de aprendizado competitivo (não-supervisionado), que é dependente apenas das características dos dados de entrada. A segunda etapa do treinamento, vinculado aos ajustes dos pesos dos neurônios da camada de saída, se processa de forma supervisionada, com o intuito de associar as respostas da rede com as respectivas saídas desejadas (Silva et al., 2010).

Como citado por Fausett (1994), há dois tipos de rede de contra-propagação: completa (*full*) e somente adiante (*forward-only*).

A Figura 2.5 ilustra a arquitetura da rede de contra-propagação completa.

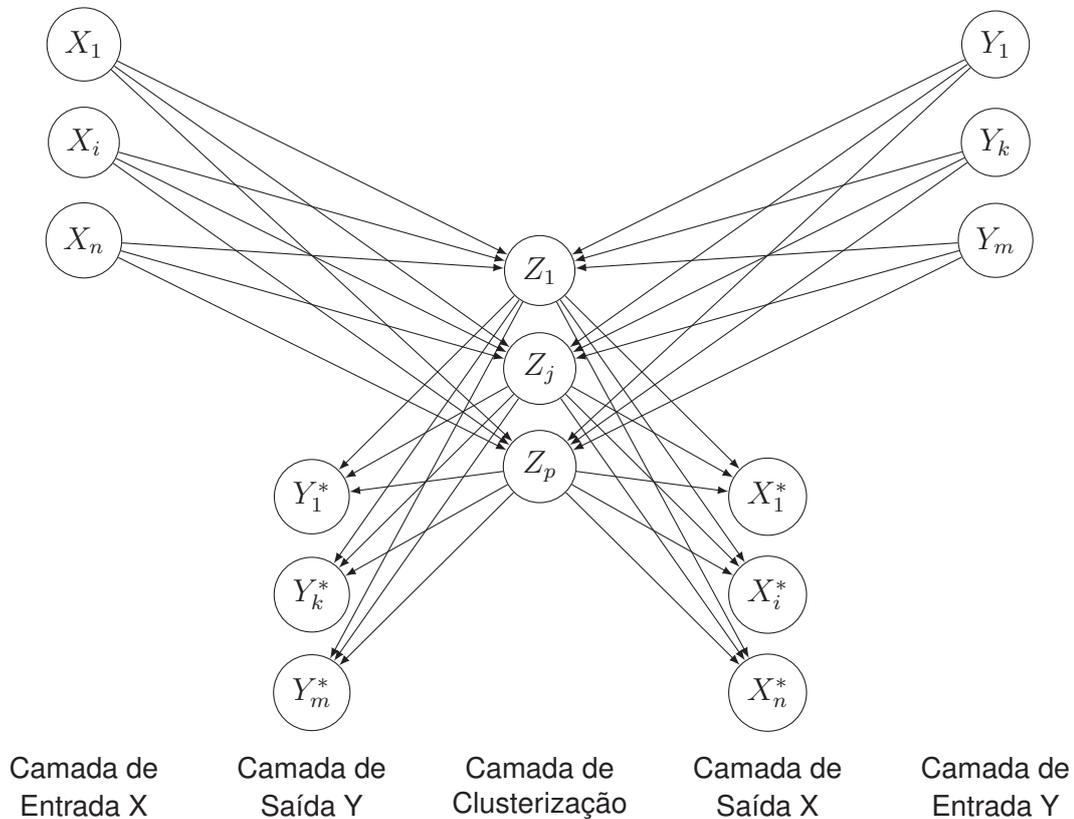


Figura 2.5: Arquitetura da rede de contra-propagação completa.

Fonte: Adaptado de Fausett (1994)

O treinamento da rede de contra-propagação completa ocorre em duas fases. Durante a primeira fase, a rede usa as camadas X e Y para formar os *clusters*. Assumindo nenhuma topologia para a camada de clusterização, apenas o neurônio vencedor é que aprende. O neurônio vencedor (representado pelo índice J) é definido pela menor distância Euclidiana (D_j), dada por:

$$D_j = \sum_{i=1} (x_i - v_{ij})^2 + \sum_k (y_k - w_{kj})^2 \quad (2.2)$$

onde x_i e y_i são os vetores de entrada; v_{ij} é o vetor dos pesos sinápticos entre a camada de entrada X e a camada de clusterização Z ; e w_{kj} é o vetor dos pesos sinápticos entre a camada de entrada Y e a camada de clusterização Z .

A regra de aprendizado de Kohonen para atualizações dos pesos do neurônio vencedor é:

$$v_{iJ}(\text{novoo}) = (1 - \alpha)v_{iJ}(\text{anterior}) + \alpha x_i, \quad i = 1, \dots, n; \quad (2.3)$$

$$w_{kJ}(\text{novoo}) = (1 - \beta)w_{kJ}(\text{anterior}) + \beta y_k, \quad k = 1, \dots, m. \quad (2.4)$$

onde α e β são as taxas de aprendizagem de entrada da camada de clusterização.

Durante a segunda fase do algoritmo de aprendizado, apenas os pesos do neurônio vencedor para as camadas de saída são ajustados a fim de que o vetor y^* seja uma aproximação do vetor de entrada y e que x^* seja uma aproximação de x . A atualização dos pesos para a camada Y^* e X^* são:

$$u_{Jk}(\text{novoo}) = (1 - a)u_{Jk}(\text{anterior}) + ay_k, \quad k = 1, \dots, m; \quad (2.5)$$

$$t_{Ji}(\text{novoo}) = (1 - b)t_{Ji}(\text{anterior}) + bx_i, \quad i = 1, \dots, n. \quad (2.6)$$

onde u_{Jk} é o vetor dos pesos sinápticos entre a camada de clusterização Z e a camada Y^* ; t_{Ji} é o vetor dos pesos sinápticos entre a camada de clusterização Z e a camada X^* ; e a e b são as taxas de aprendizagem de saída da camada de clusterização. A regra de aprendizado representada pelas Equações 2.5 e 2.6 constitui um arranjo de Grossberg, denominado também de estrutura *Outstar*. O treinamento da camada *Outstar* se processa de forma supervisionada, análogo à regra delta, em que o aprendizado ocorre para todos os neurônios da camada de saída (Fausett, 1994).

As redes de contra-propagação somente adiante, que são uma versão simplificada das redes de contra-propagação completa, usam apenas o vetor da camada de entrada para formar os *clusters* da camada de Kohonen durante a primeira fase do treinamento. A arquitetura dessa rede consiste em três camadas: camada de entrada, camada de clusterização e camada de saída. A arquitetura, incluindo a representação dos vetores de pesos, é ilustrada na Figura 2.6.

O primeiro passo do treinamento da rede de contra-propagação somente adiante é apresentar os vetores na camada de entrada. Em seguida, os neurônios competem na camada de clusterização e apenas o neurônio vencedor atualiza seus pesos. Depois de todo o conjunto de vetores de treinamento ter sido apresentado, a taxa de aprendizado é reduzida e os vetores são novamente apresentados à rede, até atingir algum critério de parada.

O neurônio vencedor é calculado pela menor distância Euclidiana entre os vetores x e v . O neurônio vencedor atualiza seus pesos pela seguinte regra de aprendizado:

$$v_{iJ}(\text{novoo}) = v_{ij} + \alpha(x_i - v_{iJ}) \quad (2.7)$$

onde v_{iJ} é o vetor de pesos entre a camada de entrada e de clusterização; x_i é o vetor de entrada; α é a taxa de aprendizado (Fausett, 1994).

A segunda fase do treinamento consiste em atualizar os pesos da camada de clusterização para a camada de saída da rede. Portanto, à medida que cada vetor de entrada de treinamento é apresentado à rede, deve ser apresentado um alvo associado à camada de saída.

A regra de aprendizado para atualização dos pesos entre a camada de clusterização e a camada de saída é dada por:

$$w_{Jk}(\text{novoo}) = (1 - a)w_{Jk} + ay_k \quad (2.8)$$

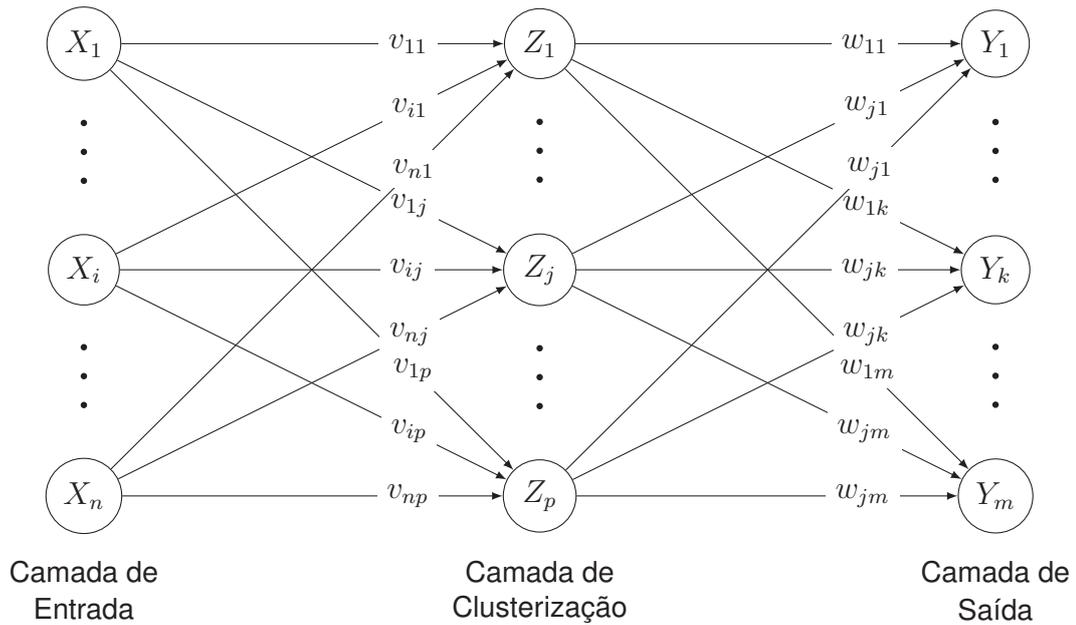


Figura 2.6: Arquitetura da rede de contra-propagação somente adiante.
 Fonte: Adaptado de Fausett (1994)

onde w_{Jk} é o vetor de pesos entre a camada de clusterização e a camada de saída, y_k é o alvo associado ao vetor de entrada e a é a taxa de aprendizado que varia entre 0 e 1 (Fausett, 1994).

O neurônio vencedor da camada de clusterização manda um sinal de ativação z_j igual a 1 para a camada de saída, outros neurônios não mandam sinal de ativação para a camada de saída. Assim a Equação 2.8 pode ser escrita na forma da regra delta:

$$w_{Jk}(\text{novoo}) = w_{Jk} + az_j(y_k - w_{Jk}). \quad (2.9)$$

2.4.2 Algoritmo da rede de contra-propagação

O algoritmo de treinamento da rede de contra-propagação completa é sintetizado no Quadro 2.1.

Quadro 2.1: Algoritmo de treinamento da rede de contra-propagação completa. Fonte: Fausett (1994).

Passo 0	Iniciar pesos e taxas de aprendizado.
Passo 1	Enquanto a condição de parada da fase 1 do treinamento for falsa, fazer Passos 2-7.
Passo 2	Para cada par de treinamento $x:y$, fazer Passos 3-5.
Passo 3	Definir a ativação da camada de entrada X para o vetor x . Definir a ativação da camada de entrada Y para o vetor y .
Passo 4	Encontrar o neurônio vencedor e nomear de índice J. O neurônio vencedor é o que possui a menor distância Euclidiana, calculada por: $D_j = \sum_i (x_i - v_{ij})^2 + \sum_k (y_k - w_{kj})^2;$
Passo 5	Atualizar pesos do neurônio vencedor pelas Eq. 2.3 e 2.4.
Passo 6	Reduzir as taxas de aprendizado α e β .
Passo 7	Testar condição de parada para fase 1 do treinamento.
Passo 8	Enquanto condição de parada da fase 2 do treinamento for falsa, fazer Passos 9-15.
Passo 9	Para cada par de treinamento $x:y$, fazer Passos 10-13.
Passo 10	Definir a ativação da camada de entrada X para o vetor x . Definir a ativação da camada de entrada Y para o vetor y .
Passo 11	Encontrar o neurônio vencedor e nomear de índice J.
Passo 12	Atualizar os pesos de entrada da camada de clusterização pelas Eq.s 2.3 e 2.4.
Passo 13	Atualizar os pesos da camada de clusterização para a camada de saída pelas Eq.s 2.5 e 2.6.
Passo 14	Reduzir as taxas de aprendizado a e b .
Passo 15	Testar condição de parada para fase 2 do treinamento.

Após a aplicação dos procedimentos de treinamento, a rede será capaz de obter as saídas x^* e y^* com os vetores de entrada x e y . O algoritmo dessa etapa, chamado de operação da rede, é apresentado no Quadro 2.2.

Quadro 2.2: Algoritmo de operação da rede de contra-propagação completa.

Passo 0	Iniciar pesos.
Passo 1	Para cada par x e y , fazer Passos 2-4.
Passo 2	Definir a ativação da camada de entrada X para o vetor x . Definir a ativação da camada de entrada Y para o vetor y .
Passo 3	Encontrar o neurônio que está mais perto dos pares de entrada pela menor distância Euclidiana.
Passo 4	Calcular a aproximação de x e y , por: $x_i^* = \sum_j z_j t_{ji}$ $y_k^* = \sum_j z_j u_{jk}$

O algoritmo de treinamento da rede de contra-propagação somente adiante, apresentado no Quadro 2.3, é muito similar ao algoritmo da rede de contra-propagação completa. Na rede somente adiante, o vetor de saída não é apresentado na primeira fase do treinamento, o que torna o algoritmo um pouco mais simples do que foi apresentado no Quadro 2.1.

Quadro 2.3: Algoritmo de treinamento da rede de contra-propagação somente adiante. Fonte: Fausett (1994).

Passo 0	Iniciar pesos e taxas de aprendizado.
Passo 1	Enquanto a condição de parada da fase 1 do treinamento for falsa, fazer Passos 2-7.
Passo 2	Para cada par de treinamento x , fazer Passos 3-5.
Passo 3	Definir a ativação da camada de entrada X para o vetor x .
Passo 4	Encontrar o neurônio vencedor e nomear de índice J . O neurônio vencedor é o que possui a menor distância Euclidiana, calculada por: $D_j = \sum_i (x_i - v_{ij})^2;$
Passo 5	Atualizar pesos do neurônio vencedor pela Eq. 2.7.
Passo 6	Reduzir a taxa de aprendizado α .
Passo 7	Testar condição de parada para fase 1 do treinamento.
Passo 8	Enquanto condição de parada da fase 2 do treinamento for falsa, fazer Passos 9-15.
Passo 9	Para cada par de treinamento $x:y$, fazer Passos 10-13.
Passo 10	Definir a ativação da camada de entrada X para o vetor x . Definir a ativação da camada de entrada Y para o vetor y .
Passo 11	Encontrar o neurônio vencedor e nomear de índice J .
Passo 12	Atualizar os pesos de entrada da camada de clusterização pela Eq. 2.7.
Passo 13	Atualizar os pesos da camada de clusterização para a camada de saída pela Eq. 2.8.
Passo 14	Reduzir a taxa de aprendizado a .
Passo 15	Testar condição de parada para fase 2 do treinamento.

O algoritmo da etapa de operação da rede *forward-only* é mostrado no Quadro 2.4. Nesse caso a saída y^* é obtida após a apresentação à rede do vetor x^* apenas.

Quadro 2.4: Algoritmo de operação da rede de contra-propagação somente adiante.

Passo 0	Iniciar pesos.
Passo 1	Apresentar o vetor x .
Passo 2	Encontrar o neurônio J mais perto do vetor x .
Passo 3	Calcular a aproximação de y , por: $y_k^* = \sum_j z_j w_{jk}$

2.4.3 Redes de Função de Base Radial

As redes denominadas funções de base radial, convencionalmente conhecidas como RBF (*radial basis function*), podem ser empregadas em diversos tipos de problemas, inclusive aqueles que envolvem aproximação de funções e classificação de padrões.

A arquitetura da rede RBF é do tipo *feedforward*, ou seja, as informações são passadas entre as camadas apenas em um sentido, não sendo retroalimentadas. A rede possui três camadas: uma camada de entrada, uma camada oculta, na qual as funções de ativação são de base radial, e uma camada de saída. A Figura 2.7 ilustra a arquitetura de uma rede RBF.

Na Figura 2.7, X representa os neurônios da camada de entrada, $W_{ij}^{(1)}$ são os pesos sinápticos entre a camada de entrada e a camada oculta, $g_j^{(1)}(u_j^{(1)})$ são as funções de ativação da camada oculta (a saber, funções de base radial), $W_{jk}^{(2)}$ são os pesos sinápticos que ponderam as ligações entre o neurônio da camada oculta e o neurônio de saída, $g_k^{(2)}(u_k^{(2)})$ são as funções de ativação da camada de saída (funções lineares) e y são os vetores de saída.

O treinamento é realizado em duas etapas. Na primeira etapa é adotado um método de aprendizado competitivo (não-supervisionado), nos quais os neurônios da camada oculta competem para estabelecer o neurônio vencedor e assim formar *clusters* a partir dos dados de

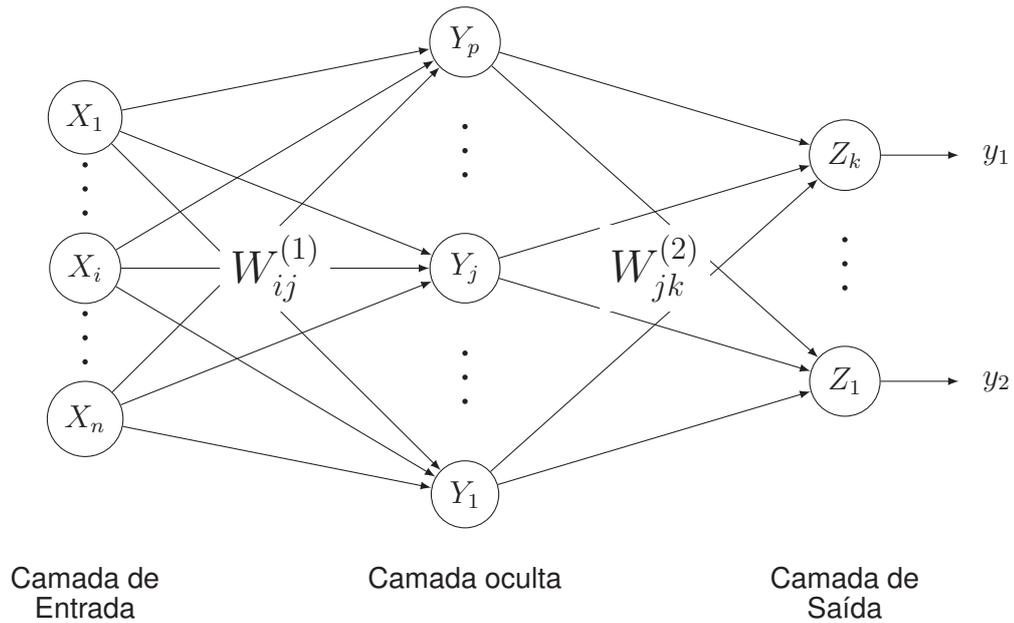


Figura 2.7: Arquitetura da rede RBF.
 Fonte: Adaptado de Silva et al. (2010)

entrada. A segunda etapa, relacionada ao ajuste dos pesos dos neurônios da camada de saída, utiliza o método de aprendizado da regra delta generalizada (Silva et al., 2010).

Os ajustes dos neurônios da camada oculta, chamado de primeira etapa do treinamento, são constituídos de funções de ativação do tipo base radiais, sendo que a função gaussiana é uma das mais empregadas. A expressão que define uma função de ativação do tipo gaussiana é representada por:

$$g(u) = \exp \frac{-(u - c)^2}{2\sigma^2} \quad (2.10)$$

onde c é o centro da função gaussiana e σ^2 é a variância da função (em que σ equivale ao desvio padrão), a qual indica o quão disperso está o potencial de ativação (u) em relação ao centro (c) (Silva et al., 2010).

A Figura 2.8 ilustra o formato geométrico de uma função gaussiana.

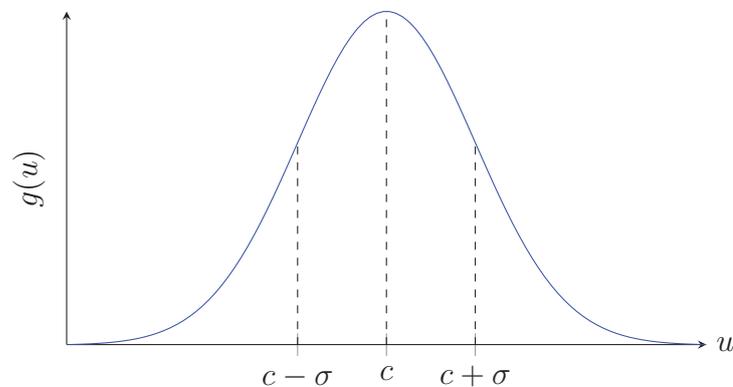


Figura 2.8: Função Gaussiana.
 Fonte: Elaborada pela autora.

Os parâmetros livres a serem ajustados em uma função gaussiana são a posição do centro c e o desvio padrão σ . Quanto maior o desvio padrão, mais alargada será a base da gaussiana.

A posição do centro c está diretamente associado aos pesos sinápticos da primeira etapa do treinamento e a entrada u_j^1 é o próprio vetor de entrada x . A saída de cada neurônio j da camada oculta é expressa por:

$$g_j^{(1)}(u_j^{(1)}) = \exp \frac{-\sum_{i=1}^p (x_i - w_{ij}^{(1)})^2}{2\sigma_j^2} \quad (2.11)$$

onde $j = 1, \dots, p$.

Quanto mais próxima esteja o vetor x do centro da gaussiana, mais significativa será o valor produzido pela função de ativação. Em tal condição, o neurônio produzirá respostas similares para todos aqueles padrões que estejam a uma mesma distância radial do centro.

Os valores dos centros (também conhecido por pesos sinápticos da primeira etapa do treinamento) são iniciados aleatoriamente. Quando um vetor de entrada é apresentado, calculam-se as distâncias Euclidianas entre este vetor e cada um dos centros. O que apresentar menor distância é o neurônio vencedor e o vetor passa a pertencer ao conjunto de amostras daquele neurônio vencedor. Após a classificação de todos os vetores, as posições dos centros das gaussianas são ajustadas levando-se em conta todas as amostras contidas nos conjuntos. Os passos de apresentação dos vetores e ajustes dos centros das gaussianas são repetidos de maneira iterativa até que a diferença das posições dos centros entre a iteração atual e a anterior seja inferior à um valor pré-estabelecido. Este valor deve ser pequeno a ponto de significar a quase estagnação do processo de ajuste. Essa é a primeira etapa do treinamento da rede RBF e a sequência computacional é apresentada no Quadro 2.5.

Portanto, o principal objetivo dos neurônios da camada oculta é posicionar os centros de suas gaussianas de forma mais apropriada possível. Um método muito utilizado para esta finalidade é chamado de *k-means* (*k*-médias), cujo propósito é posicionar os centros das *k*-gaussianas em regiões onde os padrões de entrada tenderão a se agrupar, sendo o parâmetro *k* igual ao número de neurônios da camada oculta (Silva et al., 2010).

A segunda etapa do treinamento da rede RBF refere-se ao ajuste dos pesos dos neurônios da camada de saída e deve ser executada somente após a finalização da primeira etapa do treinamento. Diferentemente da primeira etapa do treinamento, esta segunda etapa utiliza um processo de aprendizado supervisionado, ou seja, são apresentados à RNA os valores de entrada e de saída (referência), sendo que as entradas são as respostas produzidas pelas funções de ativação gaussianas dos neurônios da camada intermediária, ou seja:

$$u_k^{(2)} = \sum_{j=1}^p W_{jk}^{(2)} \times g_j^{(1)}(u_j^{(1)}) - \theta_k \quad (2.12)$$

onde $W_{jk}^{(2)}$ e θ_k são, respectivamente, os pesos e limiares referentes aos neurônios da camada de saída e $g_j^{(1)}(u_j^{(1)})$ é valor de saída do neurônio j da camada oculta, obtido pela equação 2.11 na primeira etapa do treinamento.

Finalmente, adotando-se a função linear como ativação para a camada de saída da RBF, as respostas produzidas pelos neurônios de saída serão dados por:

$$y_k = g_k^{(2)}(u_k^{(2)}). \quad (2.13)$$

2.4.4 Algoritmo da rede RBF

A sequência de procedimentos computacionais da primeira etapa do treinamento da rede RBF é sintetizada no Quadro 2.5.

Quadro 2.5: Algoritmo da primeira etapa do treinamento da rede RBF. Fonte:(Silva et al., 2010).

Passo 0	Determinar os dados para amostra de treinamento.
Passo 1	Iniciar o vetor de pesos dos neurônios da camada oculta $W_{ij}^{(1)}$ com valores aleatórios.
Passo 2	Repetir os Passos 3-7.
Passo 3	Para todas as amostras de treinamento, fazer Passos 4-5.
Passo 4	Calcular a distância Euclidiana entre o vetor de entrada e o vetor $W_{ij}^{(1)}$.
Passo 5	Selecionar o neurônio que apresenta a menor distância. Assim, o vetor de entrada passa a pertencer ao conjunto daquele neurônio.
Passo 6	Para todos os vetores de pesos $W_{ij}^{(1)}$, fazer Passo 7
Passo 7	Ajustar o vetor $W_{ij}^{(1)}$ para o valor médio de todos os vetores que pertencem ao conjunto.
Repetir Passo 2 até que a posição dos vetores $W_{ij}^{(1)}$ seja suficientemente pequena.	
Passo 8	Para todos $W_{ij}^{(1)}$, calcular as variâncias de cada uma das funções de ativação gaussianas.

Ao término da primeira etapa do treinamento, inicia-se a segunda etapa, em que são ajustados os pesos sinápticos da camada de saída para obter as respostas produzidas pela rede. A sequência computacional dessa etapa é apresentada no Quadro 2.6

Quadro 2.6: Algoritmo da segunda etapa do treinamento da rede RBF. Fonte:(Silva et al., 2010).

Passo 0	Determinar os dados para amostra de treinamento.
Passo 1	Determinar o vetor de saída desejada ($d_{\{k\}}$) para cada amostra.
Passo 2	Iniciar o vetor de pesos dos neurônios da camada de saída $W_{jk}^{(2)}$ com valores aleatórios.
Passo 3	Determinar a taxa de aprendizagem e o erro máximo.
Passo 4	Para todas as amostras de treinamento, fazer Passo 5.
Passo 5	Obter os valores de $g_j^{(1)}$ em relação ao vetor de entrada, conforme equação 2.11.
Passo 6	Iniciar o contador de número de épocas.
Passo 7	Repetir os Passos 8-10
Passo 8	O erro quadrático médio (E_m) é igual ao erro quadrático médio anterior (E_m^{ant}).
Passo 9	Ajustar $W_{jk}^{(2)}$ e θ_j para todos os pares de treinamento.
Passo 10	O E_m obtido na época atual é igual a E_m^{atual} .
Até que $ E_m^{atual} - E_m^{ant} $ seja menor que o erro máximo.	

Após a finalização do treinamento, a rede está apta para estimar as saídas do processo que foi mapeado, diante às novas amostras que são apresentadas apenas as suas entradas. A sequência computacional da fase de operação é apresentada no Quadro 2.7.

Quadro 2.7: Algoritmo da fase de operação da rede RBF. Fonte:(Silva et al., 2010).

Passo 0	Apresentar uma amostra.
Passo 1	Assumir os parâmetros $W_{ij}^{(1)}$, σ_j^2 , $W_{jk}^{(2)}$ e θ_k já ajustados no treinamento.
Passo 2	Executar os Passos 3-5.
Passo 3	Obter $g_j^{(1)}$, conforme equação 2.11;
Passo 4	Obter $u_k^{(2)}$, conforme equação 2.12;
Passo 5	Obter y_k , conforme equação 2.13.

2.5 MÉTODO DE SELEÇÃO DE VARIÁVEIS: REGRESSÃO *STEPWISE*

As RNAs, como outros modelos empíricos de previsão climática, exigem um conjunto de dados preditores de entrada. Há uma grande quantidade de variáveis de entrada potencialmente relevantes para a previsão de temperatura e precipitação a médio prazo. No entanto, deve-se selecionar o menor número possível de entradas que retenham o máximo de informação dos padrões que busca-se prever.

Os métodos de seleção de variáveis são importantes ferramentas para obter o melhor conjunto de dados preditores do problema. Esses métodos evitam a multicolinearidade, que é causada pela presença de muitas variáveis no mesmo modelo refletindo as mesmas informações. Outra vantagem da aplicação desses métodos é que o tempo computacional pode ser reduzido, aumentando a velocidade do modelo, pois os preditores redundantes são desconsiderados (Maheshwari, 2019).

Existem algumas técnicas de seleção de variáveis, entre elas estão a análise de componentes principais (do inglês, *Principal Component Analysis*, PCA), que transforma linearmente um conjunto de variáveis, possivelmente correlacionadas entre si, num conjunto substancialmente menor de variáveis não correlacionadas que contém a maior parte da informação do conjunto original. Embora a PCA tenha sido amplamente utilizado em oceanografia e meteorologia desde sua introdução, o fato de ser um método linear implica uma potencial simplificação excessiva dos conjuntos de dados analisados (Hsieh, 2001).

Assim, foram desenvolvidos alguns métodos para superar a dificuldade em se analisar dados com relações não lineares, por exemplo, a PCA não linear e o método de regressão *stepwise*. Esse último foi usado no presente trabalho para a seleção das variáveis de entrada das RNAs. A PCA não linear é baseada em uma rede neural perceptron multicamada com uma topologia autossociativa, também conhecida como *autoencoder*, rede replicadora ou gargalo. A rede autoassociativa executa um mapeamento de identidade, em que a saída é forçada a se igualar à entrada com grande acurácia, minimizando o erro quadrático (Scholz et al., 2008).

O método de regressão *stepwise* é feito de forma iterativa, adicionando (passo *forward*) ou removendo variáveis (passo *backward*), a partir de um critério de seleção. Um dos critérios de seleção mais usados é o teste estatístico F, mas também são utilizados o coeficiente de correlação linear múltipla e o erro quadrático total (Alves et al., 2013).

Para realizar o método de regressão *stepwise*, passo *forward*, deve-se escolher uma variável x_k e avaliar sua importância para o modelo através de algum teste estatístico. A seguir, outra variável x_i entra no modelo, se o teste estatístico for melhor que o anterior, x_i permanece no modelo, caso contrário x_i sai do modelo. Este processo é repetido até que não haja mais variáveis a serem incluídas no modelo. O método *backward* segue a mesma ideia, porém no sentido contrário, iniciando com todas as variáveis candidatas a preditores e removendo-as sucessivamente.

Marzouq et al. (2019) desenvolveram métodos de seleção automática das entradas da RNA que estima a radiação solar global diária na cidade de Fez (Marrocos). Dentre os métodos de seleção automática está a regressão *stepwise forward*. Os autores concluíram que os métodos de seleção automática, incluindo o *stepwise forward*, atingiram excelentes desempenhos na escolha de preditores para a estimativa da radiação solar global diária, reduzindo o tamanho das variáveis de entrada do modelo e melhorando o desempenho computacional.

3 MATERIAIS E MÉTODOS

3.1 BASE DE DADOS CLIMATOLÓGICOS

Para a realização da previsão climática de temperatura e precipitação no Paraná serão utilizados dados climatológicos extraídos da base de dados do *European Centre for Medium-Range Weather Forecasts* (ECMWF). O ECMWF é uma organização intergovernamental independente apoiada pela maioria dos países da Europa e está sediada em Shinfield Park, Reading, Reino Unido. A organização, que foi criada em 1975, é tanto um instituto de pesquisa quanto um serviço operando 24 horas por dia, 7 dias por semana, que produz e divulga previsões meteorológicas numéricas para seus Estados-membros. O Centro também oferece um catálogo de dados de previsão que podem ser comprados por empresas em todo o mundo e outros clientes comerciais (ECMWF, 2019).

O ECMWF usa periodicamente seus modelos de previsão e dados observados para criar um conjunto de dados de reanálise. Os dados de reanálise contêm estimativas pretéritas de parâmetros atmosféricos, tais como temperatura do ar, pressão e vento em diferentes altitudes e parâmetros de superfície, como precipitação, umidade do solo e temperatura da superfície do mar. As estimativas são produzidas para todo o globo e abrangem um longo período de tempo, que pode se estender por décadas ou mais. Os dados de reanálise do ECMWF estão disponíveis gratuitamente através do endereço eletrônico <https://www.ecmwf.int/en/forecasts/datasets/browse-reanalysis-datasets>.

Como os projetos de reanálise usam dados de observações meteorológicas, as incertezas dos produtos de reanálise não são iguais para todo o globo. Em geral, os dados são mais precisos em datas recentes em comparação com 30 anos atrás, quando menos observações estavam disponíveis, e possuem maiores incertezas em determinados locais, por exemplo, perto de ciclones tropicais ou nas trilhas de tempestades.

O projeto de reanálise do ECMWF utilizado neste estudo é o ERA5, cujos os dados abrangem o período de 1979 até a atualidade. Os dados de reanálise do ERA5 estão redistribuídos para uma grade regular com resolução espacial na horizontal (latitude e longitude) de $0,25^\circ \times 0,25^\circ$. Na vertical, os dados estão disponíveis em níveis de pressão ou em nível único. Para os níveis de pressão, há 16 variáveis atmosféricas em 37 níveis de 1.000 hPa (superfície) a 1 hPa (em torno do topo da estratosfera) (Copernicus, 2017).

Os dados utilizados para as previsões de temperatura e precipitação são médias espaciais e mensais e o período compreendido é de janeiro de 1979 até dezembro de 2018 (480 meses). Baseado no estudo climatológico apresentado na seção 2.2, foram selecionados alguns fenômenos que podem ocasionar variação no quadro médio climático do estado do Paraná e suas posições geográficas foram determinadas pela maior probabilidade de ocorrência desses fenômenos. As localizações nestes fenômenos são mostrados na Figura 3.1.

Com base nos fenômenos atmosféricos mostrados na Figura 3.1, foi selecionado um conjunto de variáveis preditoras possíveis para a variação climática do Paraná que representa o fenômeno estudado, por exemplo, a aproximação de um anticiclone modifica a TSM e a pressão na superfície na região, assim foram escolhidas essas duas variáveis para representar o ASAS, ASPS e AMP. Deste modo, foram selecionadas as variáveis possíveis para a previsão climática do Paraná que são apresentadas na Tabela 3.1.

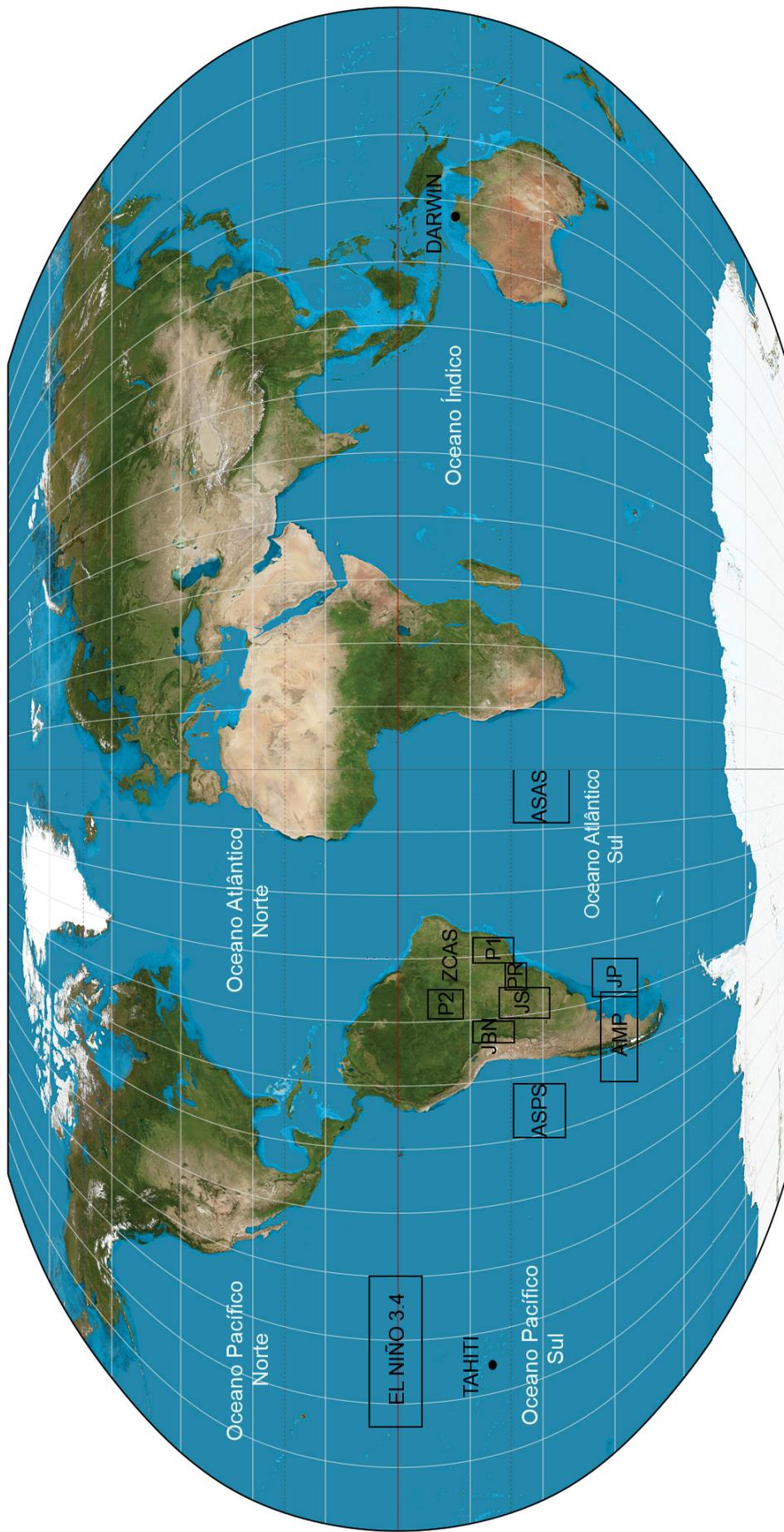


Figura 3.1: Localização de ocorrência dos fenômenos atmosféricos que podem ocasionar a variação climática no estado do Paraná. Fonte do mapa cartográfico Wikipedia, 2019. Na Figura tem-se: **ASAS** Anticiclone Subtropical do Atlântico Sul; **ASPS** Anticiclone Subtropical do Pacífico Sul; **AMP** Anticiclone Móvel Polar; **EL NIÑO 3.4** Região que ocorre o El Niño; **DARWIN e TAHITI** a diferença de pressão atmosférica das duas regiões é representada pelo Índice de Oscilação Sul (IOS); **JBN** Jato de Baixo Nível; **JP** Jato Polar; **JS** Jato Subtropical; **PR** estado do Paraná; **ZCAS** Zona de Convergência do Atlântico Sul, representada por duas áreas, P1 e P2.

Tabela 3.1: Localização e variáveis preditoras possíveis para previsão de variação climática do Paraná.

Regiões	coordenadas		variáveis	código
	latitude	longitude		
Anticiclone Subtropical do Atlântico Sul	27 - 32°S	0 - 10°W	temperatura da superfície do mar pressão na superfície	tsm_ASAS pres_ASAS
Anticiclone Subtropical do Pacífico Sul	27 - 32°S	80 - 90°W	temperatura da superfície do mar pressão na superfície	tsm_ASPS pres_ASPS
Anticiclone Móvel Polar	42,5 - 50°S	60 - 80°W	temperatura da superfície do mar pressão na superfície	tsm_AMP pres_AMP
El Niño 3.4	5°N - 5°S	120 - 170°W	temperatura da superfície do mar	tsm_ELNINO
Jato de Baixo Nível	16 - 20°S	60 - 64°W	vento meridional a 850 hPa	v850_JBN
Jato Polar	40 - 50°S	50 - 60°W	vento zonal a 250 hPa	u250_JP
Jato Subtropical	28 - 33°S	53 - 60°W	vento zonal a 250 hPa	u250_JS
Darwin	12,45°S	130,84°E	pressão na superfície	pres_DARWIN
Tahiti	17,6°S	149,45°W	pressão na superfície	pres_TAHITI
Zona de Convergência do Atlântico Sul	Área 1	16 - 20°S 42 - 46°W	velocidade vertical a 500 hPa umidade específica a 700 hPa	w500_ZCAS1 q700_ZCAS1
	Área 2	10 - 14°S 53 - 60°W	velocidade vertical a 500 hPa umidade específica a 700 hPa	w500_ZCAS2 q700_ZCAS2
Paraná			temperatura a 2m da superfície	temp2_PR
			precipitação total	prec_PR
	22 - 27°S	48 - 55°W	velocidade vertical a 500 hPa	w500_PR
			umidade específica a 700hPa	q700_PR
		geopotencial a 850, 500 e 200 hPa	gp850_PR, gp500_PR, gp200_PR	

Para compreender melhor o fenômeno que é objeto de estudo neste trabalho, são apresentadas graficamente as médias climatológicas e anomalias dos fenômenos atmosféricos usados como possíveis preditores das RNAs no Apêndice A.

As médias climatológicas são definidas por valores médios calculados para um período relativamente longo e uniforme, compreendendo no mínimo três décadas consecutivas e sua análise faz parte de uma das maneiras de compreender melhor um fenômeno meteorológico. Os valores médios dos fenômenos atmosféricos foram calculados com base nos dados coletados do projeto de reanálise ERA5 no período de janeiro de 1979 a dezembro de 2018. As médias climatológicas da temperatura e precipitação mensais para o estado do Paraná são apresentados na Figura 3.2.

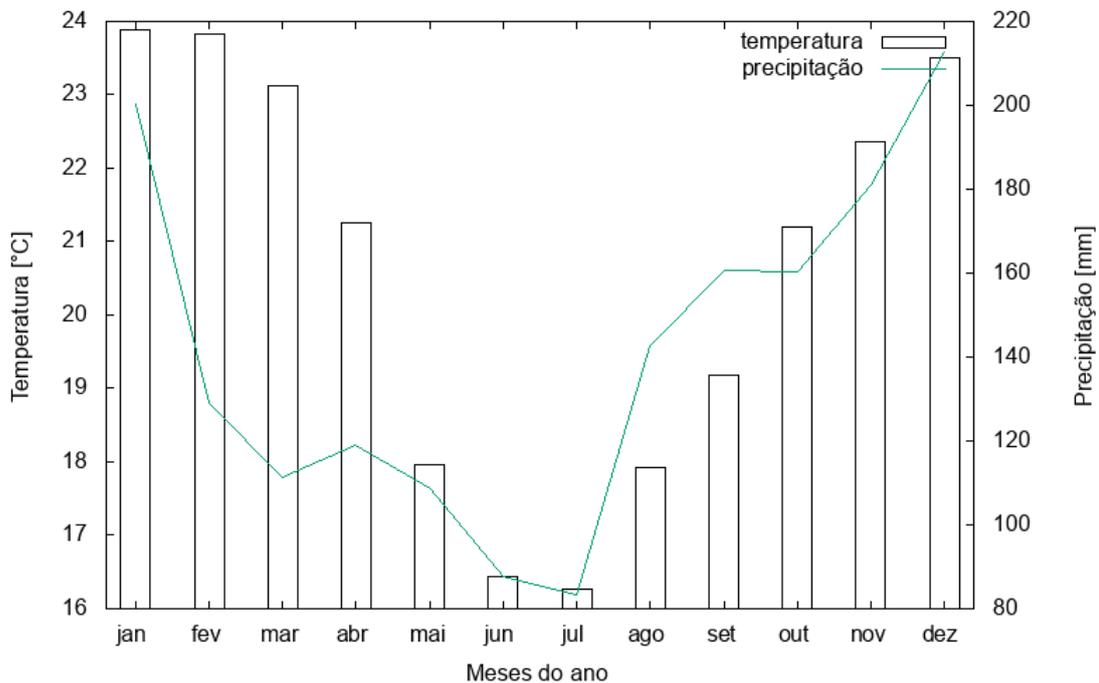


Figura 3.2: Médias climatológicas de temperatura e precipitação mensais no estado do Paraná.

Pela Figura 3.2 é possível notar que as temperaturas médias da base de dados do ERA5 são mais elevadas nos meses de dezembro a março quando as médias ficam acima de 23°C. Já as precipitações médias são maiores nos meses de novembro, dezembro e janeiro, quando as médias ficam acima de 180 mm.

As evoluções temporais e os histogramas das anomalias de temperatura e precipitação do período de janeiro de 1979 a dezembro de 2018 são apresentados nas Figuras 3.3 e 3.4, respectivamente.

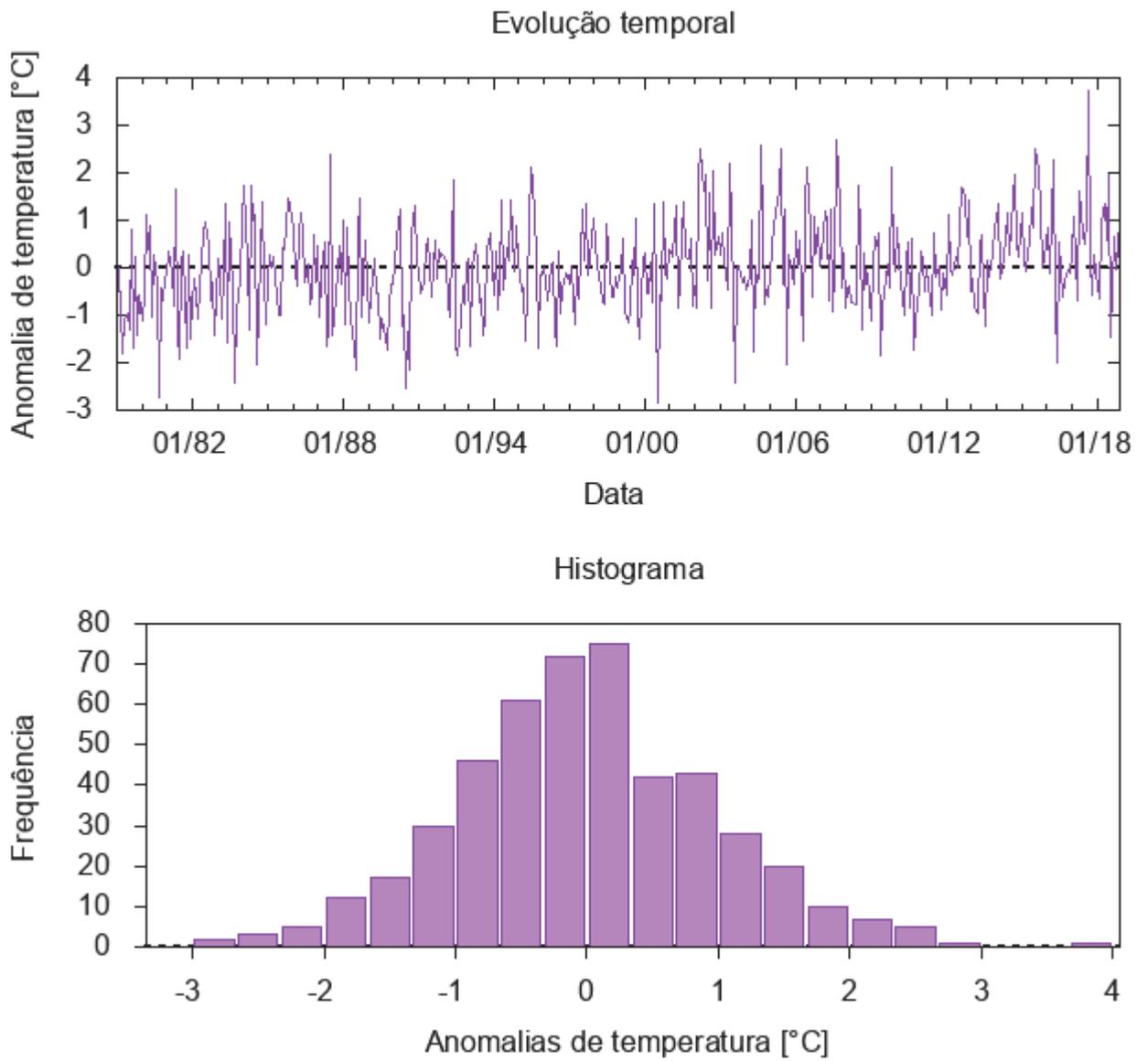


Figura 3.3: Evolução temporal e histograma dos valores de anomalias de temperatura no Paraná.

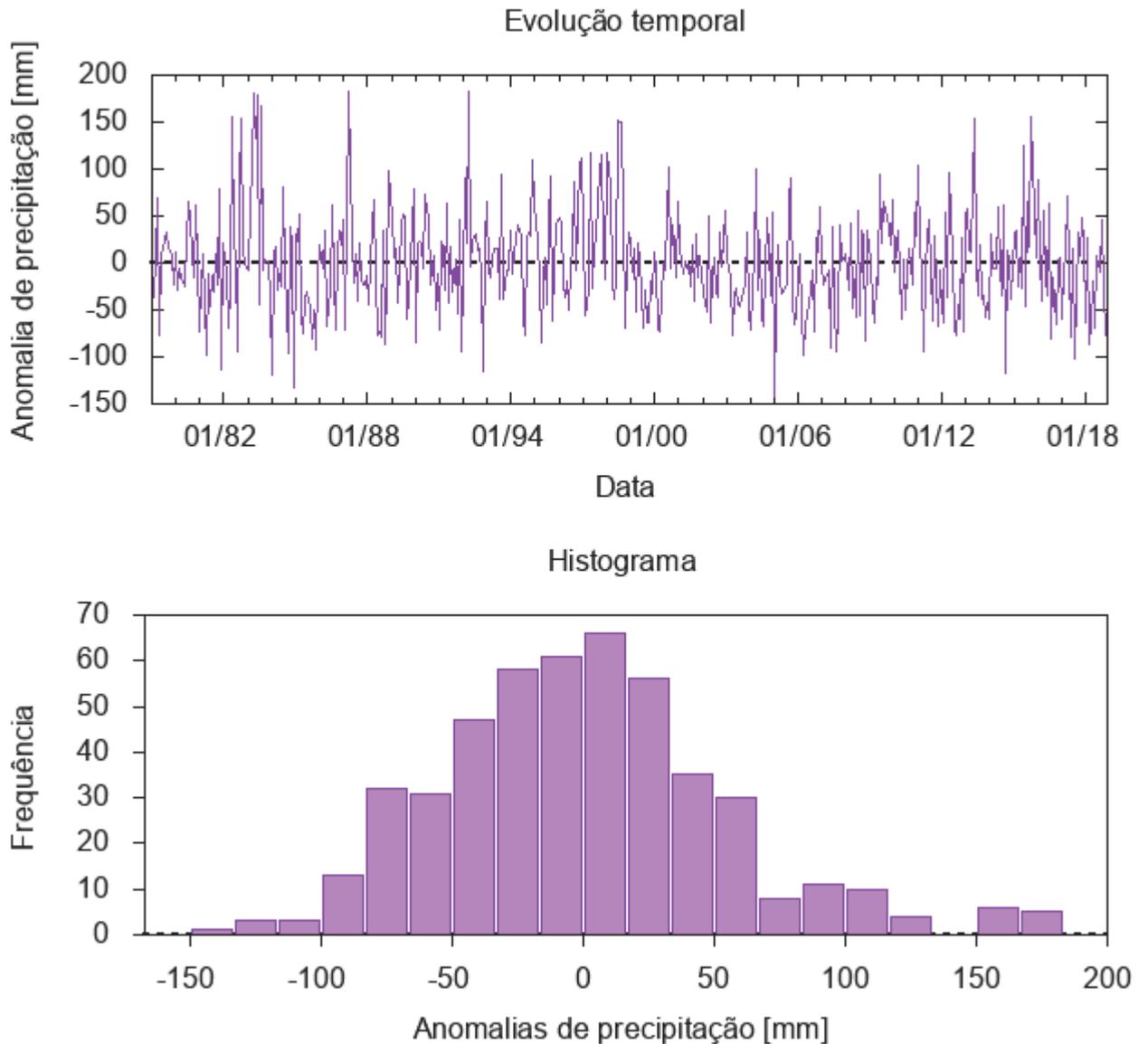


Figura 3.4: Evolução temporal e histograma dos valores de anomalias de precipitação no Paraná.

Pode-se observar pela Figura 3.3 que os valores de anomalias de temperatura estão entre -3 e 4°C . Já as anomalias de precipitação (Figura 3.4), apresentam valores entre -150 e 200 mm. Pelos gráficos das evoluções temporais das anomalias de temperatura e precipitação, apresentadas nas Figuras 3.3 e 3.4, nota-se que os dados são muito ruidosos, principalmente a precipitação, que o valor da anomalia pode corresponder ao valor da média climatológica.

As estatísticas calculadas a partir dos dados de precipitação e temperatura mensais do ERA5 são apresentadas na Tabela 3.2.

Tabela 3.2: Estatísticas calculadas a partir dos dados de precipitação e temperatura mensais.

	Média	Mediana	DPa	CV	máx	mês/ano	min	mês/ano
Temperatura [$^{\circ}\text{C}$]	20,5	21,2	2,9	14,3	25,5	2/1984	13,4	7/2000
Precipitação [mm]	141,2	132,5	68,1	48,2	323,6	12/1996	8,1	6/2017

DPa: Desvio Padrão; CV: Coeficiente de variação; máx: valor máximo mensal; min: valor mínimo mensal.

3.2 DESENVOLVIMENTO DO MODELO DE RNA

Para a construção de um modelo de RNA são necessárias algumas etapas importantes, entre elas estão a identificação de variáveis de entrada significativas e a otimização da estrutura da rede. Para selecionar as variáveis de entrada foram, primeiramente, identificados os meses que poderiam ser usados como entrada nas RNAs. Para isso, realizaram-se análises de correlação cruzada entre as variáveis apresentadas na Tabela 3.1 com atrasos de até 11 meses e a variável de interesse, que são a precipitação e temperatura, mensais e médias de 3 meses.

A Tabela 3.3 fornece uma lista dos possíveis preditores selecionados com atrasos no tempo e os valores de correlação cruzada que foram obtidos com a temperatura e precipitação mensal.

Tabela 3.3: Meses selecionados que tem os maiores valores absolutos de correlação cruzada com a temperatura e precipitação mensal.

	Temperatura mensal		Precipitação mensal	
	atraso (meses)	coef. de correlação	atraso (meses)	coef. de correlação
tsm_ASAS	4	-0,88	3	-0,52
pres_ASAS	4 e 11	0,63 e -0,66	2 e 9	0,36 e -0,39
tsm_ASAS	4	-0,89	3 e 9	-0,50 e 0,52
pres_ASAS	4	0,56	3 e 8	0,26 e -0,37
tsm_AMP	5	-0,91	3	-0,50
pres_AMP	5	0,21	3 e 9	0,10 e -0,14
tsm_ELNINO	2	-0,35	6	0,33
v850_JBN	6 e 11	-0,40 e 0,45	4	-0,31
u250_JP	3	0,10	7	-0,09
u250_JS	5	0,61	4 e 9	0,33 e -0,36
pres_DARWIN	6	0,84	5	0,53
pres_TAHITI	5 e 11	0,70 e -0,71	4 e 10	0,37 e -0,42
w500_ZCAS1	6	0,77	5	0,41
q700_ZCAS1	6	-0,77	5	-0,47
w500_ZCAS2	6	0,85	5 e 11	0,51 e -0,52
q700_ZCAS2	6 e 11	-0,73 e 0,76	5	-0,48
temp2_PR	6	-0,81	5	-0,48
prec_PR	7	-0,50	1	0,35
w500_PR	2	-0,19	6	0,21
q700_PR	6	-0,76	5 e 10	-0,44 e 0,46
gp850_PR	6 e 7	0,76 e 0,76	5 e 11	0,44 e -0,45
gp500_PR	5 e 11	-0,52 e 0,52	4 e 9	-0,26 e 0,30
gp200_PR	5 e 6	-0,76 e -0,76	4 e 10	-0,45 e 0,46

Para a temperatura e precipitação trimestrais, foram obtidas as médias móveis de 3 meses para todas as variáveis analisadas e calculado a correlação cruzada com atrasos de 3 a 11 meses. A Figura 3.5 apresenta um exemplo da configuração utilizada para a previsão sazonal e a seleção dos preditores com os atrasos.

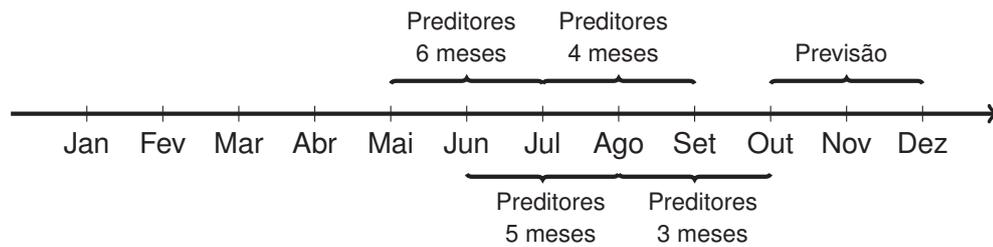


Figura 3.5: Exemplo da configuração utilizada para a previsão sazonal.

A Tabela 3.4 apresenta a lista de entrada com os atrasos no tempo e os valores de correlação cruzada que foram obtidos com a temperatura e precipitação médias trimestrais.

Tabela 3.4: Meses selecionados que tem os maiores valores absolutos de correlação cruzada com a temperatura e precipitação médias trimestrais.

	Temperatura trimestral		Precipitação trimestral	
	atraso (meses)	coef. de correlação	atraso (meses)	coef. de correlação
tsm_ASAS	4	-0,91	3 e 9	-0,67 e 0,63
pres_ASAS	4	0,77	3 e 9	0,56 e -0,54
tsm_ASPS	4 e 10	-0,92 e 0,91	3 e 9	-0,64 e 0,66
pres_ASPS	4 e 10	0,70 e -0,69	8	-0,54
tsm_AMP	5 e 11	-0,94 e 0,93	3 e 9	-0,63 e 0,63
pres_AMP	5	0,31	10	-0,21
tsm_ELNINO	8	0,36	6	0,42
v850_JBN	6 e 11	-0,54 e 0,55	4	-0,49
u250_JP	3	0,13	3	0,10
u250_JS	5 e 11	0,73 e -0,70	9 e 10	-0,52 e -0,52
pres_DARWIN	6	0,89	4	0,71
pres_TAHITI	5 e 11	0,76 e -0,76	4 e 10	0,48 e -0,56
w500_ZCAS1	6	0,87	5 e 11	0,58 e -0,60
q700_ZCAS1	6 e 11	-0,87 e 0,79	5	-0,63
w500_ZCAS2	6	0,92	5 e 11	0,66 e -0,67
q700_ZCAS2	6 e 11	-0,84 e 0,81	5	-0,64
temp2_PR	6	-0,89	5	-0,64
prec_PR	7	-0,65	6	-0,44
w500_PR	9	0,23	7	0,23
q700_PR	6	-0,86	5 e 10	-0,60 e 0,61
gp850_PR	6 e 7	0,86 e 0,86	5 e 11	0,61 e -0,60
gp500_PR	5 e 11	-0,62 e 0,65	4 e 10	-0,41 e 0,43
gp200_PR	6 e 11	-0,84 e 0,83	4 e 10	-0,60 e 0,60

Após selecionar os possíveis preditores e identificar os meses que possuem as maiores correlações lineares com as variáveis que pretendiam-se prever, foi aplicado o método de regressão *stepwise* com o intuito de diminuir a quantidade de vetores de entrada da rede e desconsiderar os preditores redundantes. Neste trabalho foi adotado o método de regressão *stepwise* com o passo *forward*, que começa com o treinamento o modelo de RNA com apenas um neurônio na camada de entrada e seleciona a variável que maximiza o desempenho do modelo, baseado no critério da menor Raiz do Erro Quadrático Médio (REQM). Em seguida,

é adicionada outra variável e avaliado o desempenho, até a última variável do conjunto de preditores possíveis. O método *stepwise* foi realizado de forma automática, por um programa desenvolvido na linguagem de programação Fortran e executado em sistema Linux. O fluxograma do método de redução de variáveis regressão *stepwise* com o passo *forward* é mostrado na Figura 3.6.

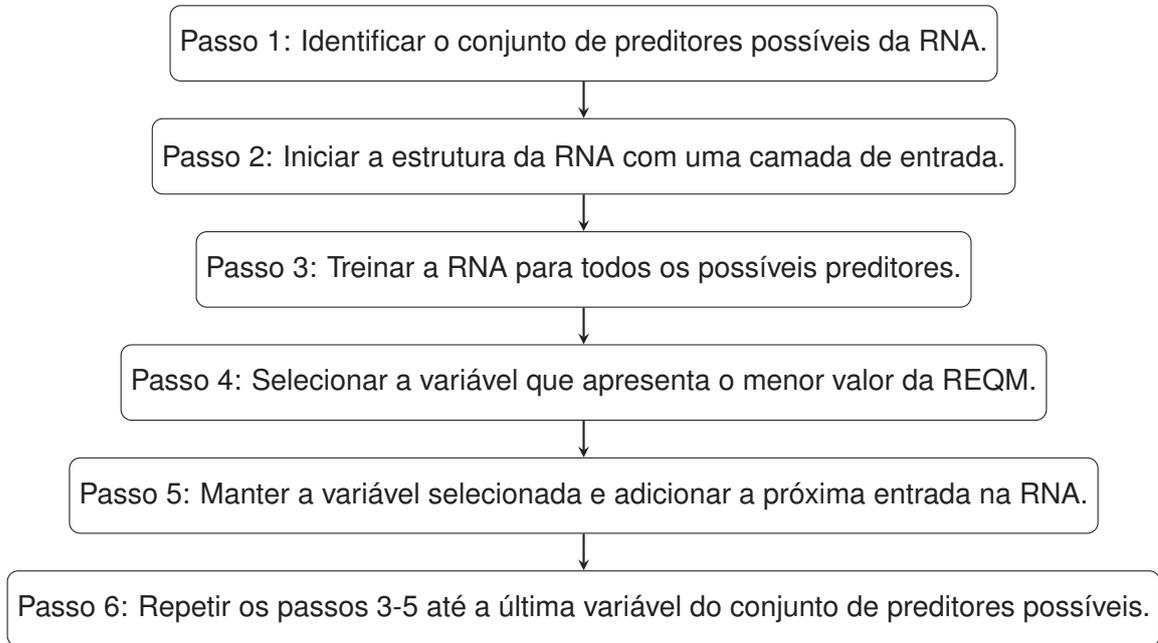


Figura 3.6: Fluxograma do método de regressão *stepwise* com o passo *forward*

A identificação do conjunto de preditores possíveis da RNA, indicado no passo 1 do fluxograma da Figura 3.6, já foram apresentados nas Tabelas 3.3 e 3.4. Para os passos seguintes da aplicação do método de regressão *stepwise* com o passo *forward* são necessários ajustes dos parâmetros das RNAs de contra-propagação e RBF, que é mostrado na seção a seguir.

3.3 ESTRUTURAS DAS RNAS DE CONTRA-PROPAGAÇÃO E RBF

Os programas de RNAs de contra-propagação e RBF foram escritos pelo Prof. Ricardo Carvalho de Almeida e ambos são compostos por dois códigos escritos em linguagem de programação Fortran, um para a operação da rede (utilizando somente 10% do total de pares entrada-saída), e outro para os conjuntos de treinamento, teste e validação (usando 90% do total de dados), sendo estes subdivididos em 80%, 10% e 10% para a rede CPN e 85%, 10% e 5% para a rede RBF, respectivamente.

A preparação dos dados consistiu em sua normalização dentro do intervalo $[-1,+1]$, para evitar que parâmetros de ordem de grandeza diferentes prejudiquem o aprendizado da rede, pois valores de ordem de grandeza maior tendem a mascarar a importância de valores de ordem de grandeza menor.

O número de neurônios na camada de entrada variou de um até o número de possíveis preditores, conforme o método *stepwise* ia avaliando as variáveis. Já o número de neurônios na camada oculta foi determinado experimentalmente. Na camada de saída, foi utilizado apenas um neurônio, pois trata-se de um problema do tipo aproximação de função, em que buscou-se prever apenas uma variável de cada vez (precipitação ou temperatura).

As redes neurais foram avaliadas para previsões mensais de precipitação e temperatura com os dados originais e, também, para as médias móveis de três meses. O objetivo de aplicar a média móvel nos dados foi filtrar variações de alta frequência (“ruído”) das séries temporais e avaliar o impacto do “ruído” no desempenho do modelo de RNA.

Para o treinamento da RNA de CPN utilizaram-se taxas de aprendizagem α e a de 0,001 e os critérios de parada dos ajustes dos pesos sinápticos foram que o erro entre a saída desejada e a resposta da rede fosse menor ou igual a 1×10^{-15} ou que atingisse 40.000 épocas. Já no treinamento da RNA RBF foi aplicado uma taxa de aprendizagem de 0,025 e os critérios de parada foram que o erro entre o valor desejado e a resposta da rede fosse menor ou igual a 1×10^{-12} ou que atingisse 15.000 épocas.

Como as RNAs definem aleatoriamente o valor do peso inicial no início do treinamento, um modelo de rede neural diferente é criado para cada processo do treinamento, gerando desempenhos diferentes. Portanto, os modelos ótimos de previsão foram selecionados com base no desempenho médio obtido pela repetição do processo de geração do modelo em cinco iterações no procedimento chamado validação cruzada.

3.4 PREVISÃO POR PERSISTÊNCIA

Fenômenos atmosféricos geralmente apontam dependência com seus próprios valores passados ou futuros. Na terminologia das ciências atmosféricas, essa dependência ao longo do tempo é geralmente conhecida como persistência. A persistência pode ser definida como a existência de dependência estatística (positiva) entre valores sucessivos da mesma variável ou entre ocorrências sucessivas de um determinado evento. Dependência positiva significa que o padrão da variável mantém-se ao longo do tempo, ou seja, valores grandes da variável tendem a ser seguidos por valores relativamente grande e valores pequenos da variável tendem a ser seguidos por valores relativamente pequenos (Wilks, 2011).

O método de previsão por persistência aplica uma técnica muito simples, que usa o fato da atmosfera possuir uma “memória”, ou seja, diversas variáveis meteorológicas medidas em instantes sucessivos, possuem uma significativa autocorrelação, em diferentes escalas temporais e espaciais.

Para realizar a previsão por persistência, simplesmente se considera que, no horizonte de previsão de interesse, a variável a ser prevista terá o mesmo valor que possui no instante inicial (ou de referência). Por exemplo, para a previsão de temperatura média de um determinado mês, considera-se que seu valor será igual ao da média registrada no mês imediatamente anterior, e para a previsão da temperatura média para os próximos três meses, considera-se que ela será igual à média das temperaturas do trimestre que se encerra imediatamente antes do próximo.

Esse método, por sua extrema simplicidade, é considerado como sem habilidade (*no skill*). Portanto, tendo seus resultados como referência, é necessário que qualquer metodologia tecnicamente mais sofisticada que ele seja capaz de produzir resultados de previsão com maior acurácia. Caso contrário, a técnica ou metodologia proposta não terá qualidade suficiente para ser considerada como útil em quaisquer aplicações.

3.5 AVALIAÇÃO DO DESEMPENHO DAS PREVISÕES

Para a avaliação do desempenho das redes neurais a serem empregadas para a previsão de temperatura e precipitação no estado do Paraná, os resultados obtidos pelo conjunto de dados identificados como operação, para realizar a validação do treinamento da rede neural,

foram submetidos a algumas métricas estatísticas para quantificar as semelhanças entre os dados observados (y_i) e os dados estimados (y_i^*). Essas métricas estatísticas são apresentadas a seguir.

- R - Correlação linear de Pearson: indica o grau de correlação linear entre duas variáveis. Varia entre os valores -1 e 1, sendo que $R = 0$ representa que não há correlação entre as variáveis analisadas e $R = +1/-1$ que há uma correlação positiva/negativa forte. O valor de R é obtido por:

$$R = \frac{\sum_{i=1}^n (y_i^* - \bar{y}^*)(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (y_i^* - \bar{y}^*)^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.1)$$

- CNS - Coeficiente de Nash-Sutcliffe: índice adimensional que varia de $-\infty$ a 1, sendo que $CNS = 1$ corresponde a uma combinação perfeita dos dados estimados com os dados observados, $CNS = 0$ indica que as previsões do modelo são tão precisas quanto a média dos dados observados e $CNS < 0$ ocorre quando a média observada é um melhor preditor que o modelo. O CNS é definido por:

$$CNS = 1 - \frac{\sum_{i=1}^n (y_i - y_i^*)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.2)$$

- REQM - Raiz do Erro Quadrático Médio: índice que mede a magnitude média dos erros estimados. Tem sempre valor positivo e quanto mais próximo de zero, maior a qualidade dos valores estimados. A REQM apresenta valor do erro na mesma dimensão da variável analisada e é definido por:

$$REQM = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^* - y_i)^2} \quad (3.3)$$

- EMA - Erro Médio Absoluto: representa o desvio absoluto médio entre os valores observados e preditos. Assim como a REQM, o EMA possui dimensão igual ao da variável analisada e é definido como:

$$EMA = \frac{1}{n} \sum_{i=1}^n |y_i^* - y_i| \quad (3.4)$$

Além das métricas estatísticas, os resultados das redes de aproximação de função também foram avaliados graficamente pelo diagrama de dispersão.

4 EXPERIMENTOS DE PREVISÃO DE TEMPERATURA E PRECIPITAÇÃO

4.1 PREVISÃO ATRAVÉS DE PERSISTÊNCIA

Na seção 3.4 foi realizada uma breve descrição a respeito de como se realiza a técnica de previsão por persistência. No presente estudo, a previsão climática por persistência foi realizada com toda a série histórica de temperatura e precipitação mensal e sazonal, ou seja, de 1979 a 2018. A técnica de persistência é considerada um método de previsão mais simples, visto que realiza a estimativa da série no instante t com base no valor no instante $t-1$. No entanto, com essa técnica é possível verificar se o fenômeno varia pouco ao longo do tempo (ou seja, persistente) ou se é muito variável (pouco persistente). Os resultados das previsões são avaliados com base das métricas estatísticas apresentadas da seção 3.5 e podem ser observados na Tabela 4.1, a seguir.

Tabela 4.1: Desempenho da previsão climática por persistência.

Previsão	R	CNS	REQM	EMA
Temperatura mensal	0,77	0,54	1,98 °C	1,59 °C
Temperatura sazonal	-0,01	-1,02	3,67 °C	3,22 °C
Precipitação mensal	0,35	-0,30	75 mm/mês	60 mm/mês
Precipitação sazonal	0,09	-0,82	66 mm/mês	54 mm/mês

Nota-se que as previsões mensais têm os melhores valores das métricas estatísticas. Este resultado já era esperado, pois a condição futura mais próxima da atual tende a se manter mais estável. Já os fenômenos sazonais sofrem as influências dos ciclos climáticos, o que dificulta a previsão através da persistência e geram resultados insuficientes.

4.2 PREVISÃO ATRAVÉS DA REDE NEURAL DE CONTRA-PROPAGAÇÃO

Nesta seção será apresentado o método de previsão através da rede neural artificial de contra-propagação para a temperatura e precipitação mensais e sazonais do estado do Paraná. Para isso, nas subseções seguintes são mostradas as seleções de variáveis de entrada dos modelos e na subseção 4.2.5 é apresentada a avaliação do desempenho das previsões pela rede neural CPN.

4.2.1 Seleção das variáveis de entrada do modelo de temperatura mensal

Para estabelecer o conjunto de dados de entrada da RNA, foi utilizado o método de regressão *stepwise*, conforme apresentado na seção 3.2. O desempenho do método de seleção de variáveis foi avaliado pela Raiz do Erro Quadrático Médio (REQM). O comportamento da REQM à medida que os preditores possíveis iam sendo adicionados apresentava um padrão de diminuição até uma certa quantidade de dados de entrada da rede, onde a partir desse ponto passou a elevar o valor do erro, como pode ser observado no exemplo apresentado na Figura 4.1, que mostra um ciclo do método de regressão *stepwise*. Note que há uma redução no valor da REQM até a inserção da terceira variável, seguido por um período de estabilidade e a partir da décima primeira variável, o erro eleva-se até 3,1 °C. Portanto, foram identificadas as variáveis em que o comportamento da REQM estava em decréscimo ou estável, que no exemplo mostrado

da Figura 4.1 foram 11 variáveis. Os resultados de todos os ciclos do método *stepwise* para identificar as variáveis predictoras dos modelos são apresentados no Apêndice B.

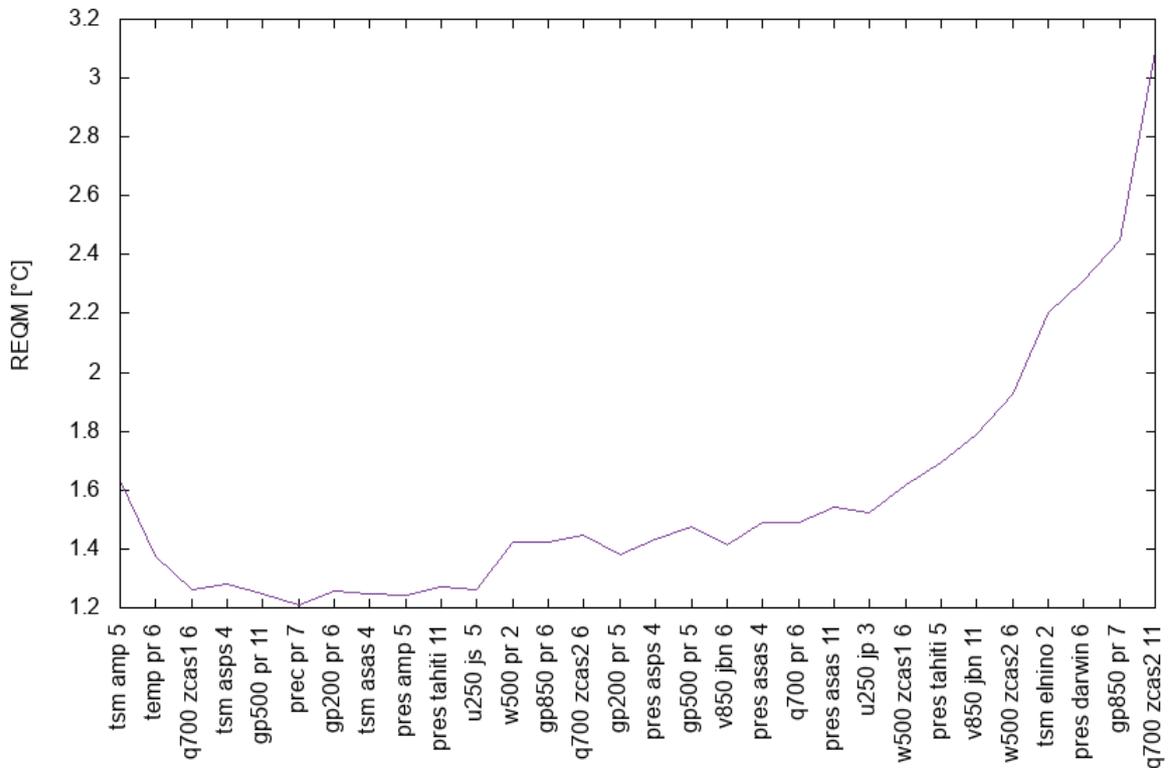


Figura 4.1: Exemplo do resultado de um ciclo do método *stepwise*.

O tempo de processamento de cada ciclo do método de regressão *stepwise* utilizando a rede de CPN foi de aproximadamente 7 horas. A configuração da camada oculta foi avaliada para 100, 150 e 200 neurônios, sendo que os melhores resultados foram obtidos com 150 neurônios, e cinco iterações no procedimento chamado validação cruzada. Como a rede neural inicializa os pesos sinápticos com valores aleatórios, em uma execução do método de seleção de variáveis pode ser indicado um conjunto de preditores relevantes para a previsão mensal de precipitação e em outra execução indicar outro conjunto. Para tentar minimizar essas diferenças, foram realizados cinco ciclos do método de regressão *stepwise*, em que o resultado pode ser visualizado na Figura 4.2.

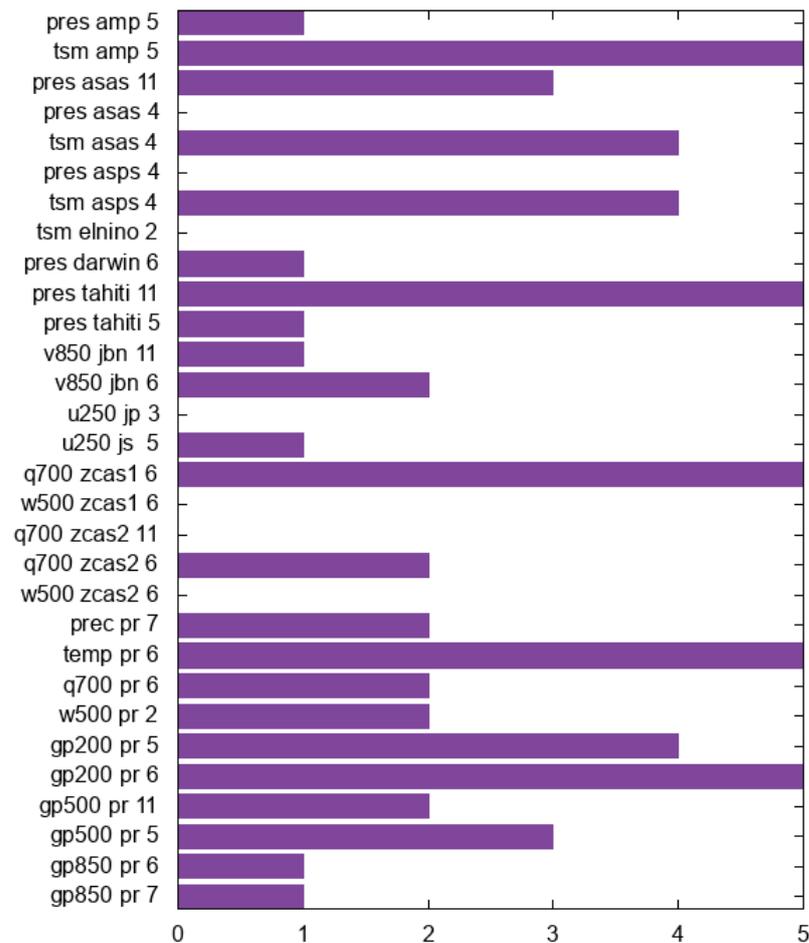


Figura 4.2: Resultado dos cinco ciclos do método de seleção de variáveis regressão *stepwise* para a temperatura mensal utilizando a RNA de CPN.

Das 30 variáveis avaliadas pelo método *stepwise*, 10 delas ocorreram três ou mais vezes, como pode-se observado na Figura 4.2. Nota-se também que cinco variáveis repetiram-se em todos os ciclos, indicando relevância para a previsão de temperatura mensal, são elas: TSM (5 meses anteriores) na região do Anticiclone Móvel Polar, pressão na superfície (11 meses) em Tahiti, umidade específica a 700 hPa (6 meses) na área 1 da Zona de Convergência do Atlântico Sul, temperatura (6 meses) e geopotencial a 200 hPa (6 meses) no Paraná.

Analisando-se a Tabela 3.3, que mostra as correlações lineares entre os preditores e a temperatura mensal, verifica-se que algumas variáveis com correlações maiores que as cinco citadas acima não se destacam como preditores importante nesses experimentos.

4.2.2 Seleção das variáveis de entrada do modelo de temperatura sazonal

Para selecionar os vetores de entrada da RNA de CPN na previsão de temperatura sazonal foram avaliados os 34 possíveis preditores, mostrados na Tabela 3.4, pelo método de seleção de variáveis regressão *stepwise*. O resultado dos cinco ciclos do método de seleção de variáveis para a temperatura sazonal é exibido na Figura 4.3.

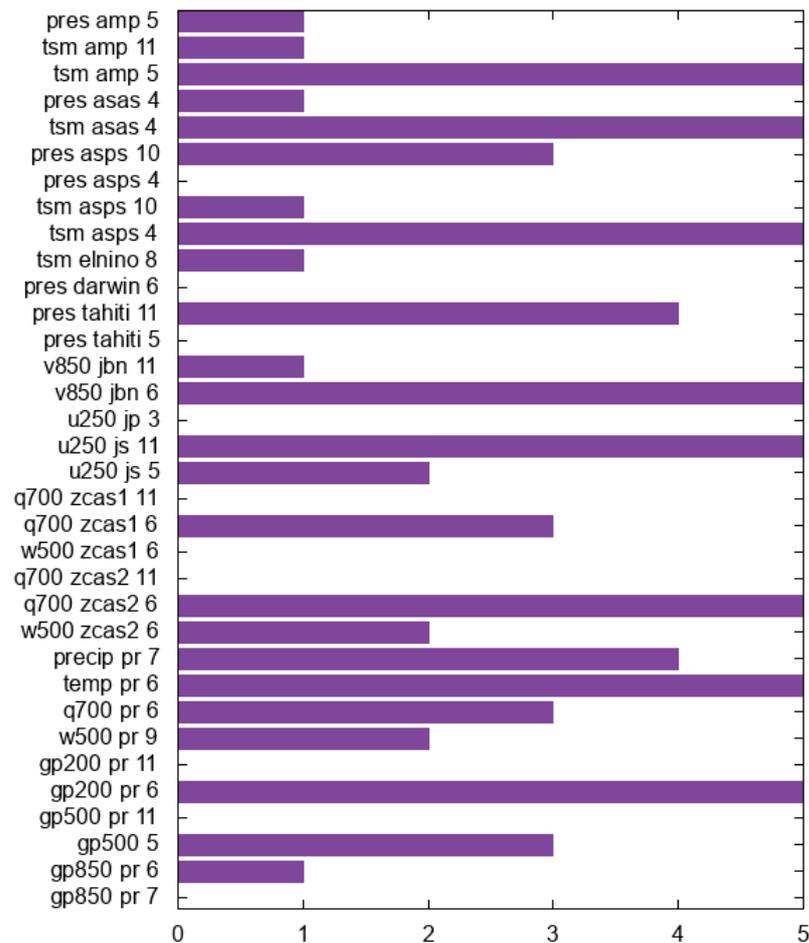


Figura 4.3: Resultado dos cinco ciclos do método de seleção de variáveis regressão *stepwise* para a temperatura sazonal utilizando a RNA de CPN.

Pode-se observar pela Figura 4.3 que 14 variáveis, das 34 analisadas, ocorreram em três ou mais ciclos do método *stepwise*. Um grande número de vetores foi selecionado como entrada da rede neural, no entanto foi esse conjunto de dados que apresentou o melhor desempenho do modelo. Também é possível notar que oito variáveis repetiram-se em todos os ciclos, indicando relevância para a temperatura sazonal, sendo eles os seguintes: TSM (5 meses de antecedência) na região do Anticiclone Móvel Polar, TSM (4 meses) nas regiões dos Anticiclones Subtropicais do Atlântico e Pacífico Sul, vento meridional a 850 hPa (6 meses) na região do Jato de Baixos Níveis, vento zonal a 250 hPa (11 meses) na região do Jato Subtropical, umidade específica a 700 hPa (6 meses) na região nomeada área 2 que representa a Zona de Convergência do Atlântico Sul e temperatura do ar (6 meses) e geopotencial a 500 hPa (6 meses) no Paraná.

4.2.3 Seleção das variáveis de entrada do modelo de precipitação mensal

Dos 34 preditores possíveis para a previsão de precipitação mensal, quase todas as variáveis candidatas foram selecionadas pelo menos uma vez durante os cinco ciclos do método de regressão *stepwise*, como pode ser observado pela Figura 4.4. Apenas a pressão na superfície e a TSM com atrasos de 8 e 3 meses, respectivamente, ambas pertencentes à região do Anticiclone Subtropical do Pacífico Sul, não foram selecionadas durante os ciclos do método *stepwise*.

Do total de preditores possíveis para a previsão de precipitação mensal, 19 deles ocorreram três ou mais vezes. Considerando que uma grande quantidade de vetores de entrada foi selecionada para o modelo, foi realizada uma segunda etapa da seleção de variáveis. A segunda etapa consistiu em avaliar apenas os 19 vetores escolhidos na primeira etapa, repetindo os cinco ciclos e selecionando aqueles que ocorreram três ou mais vezes (Figura 4.5).

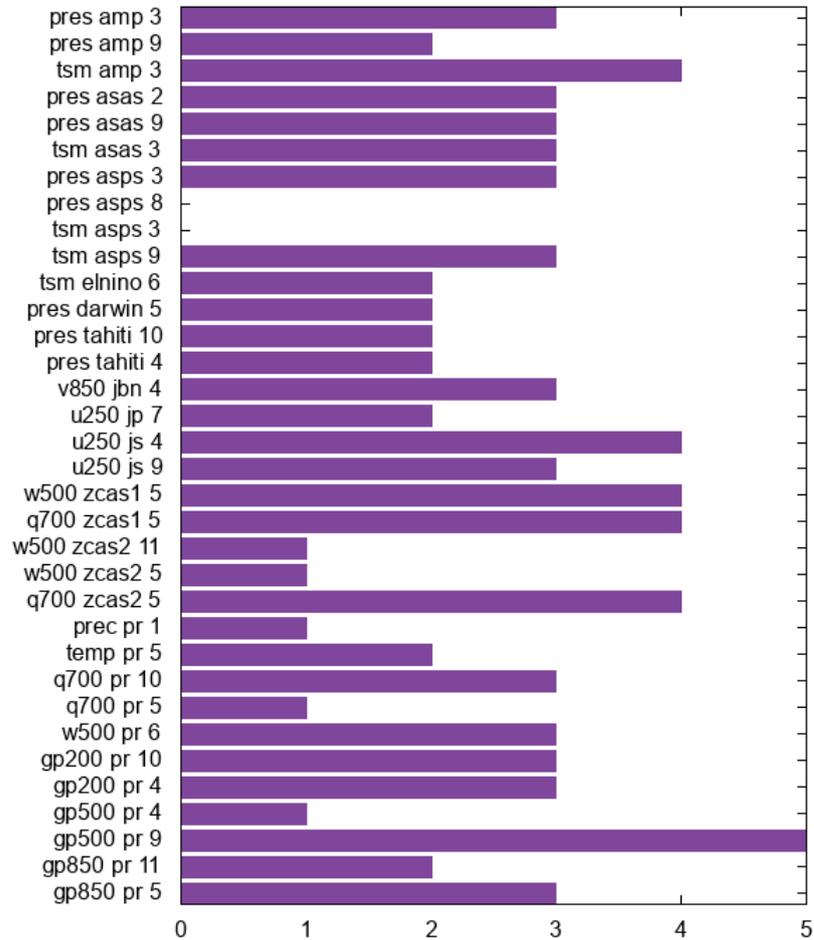


Figura 4.4: Resultado dos cinco ciclos do método de seleção de variáveis regressão *stepwise* para a precipitação mensal utilizando a RNA de CPN.

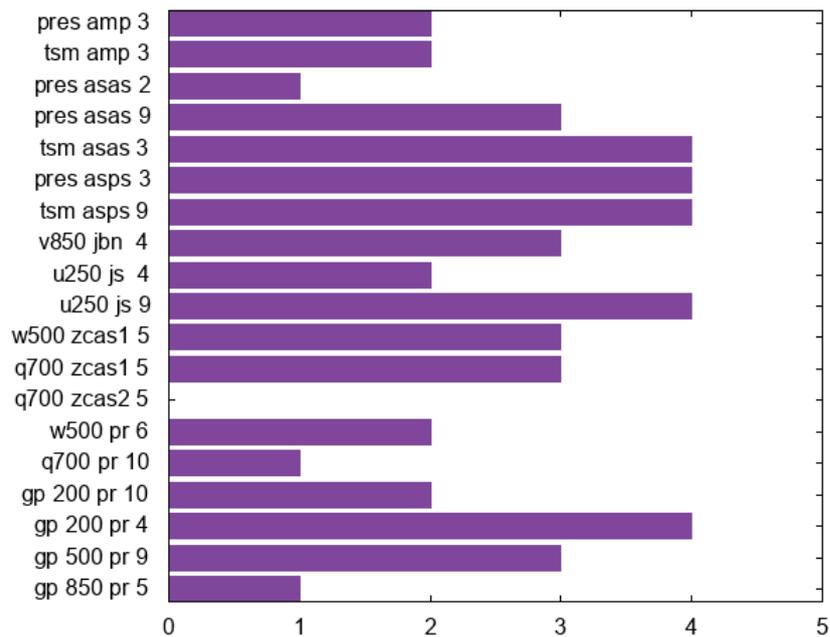


Figura 4.5: Resultado dos cinco ciclos da segunda etapa do método de seleção de variáveis regressão *stepwise* para a precipitação mensal utilizando a RNA de CPN.

Das 19 variáveis avaliadas na segunda etapa do método de regressão *stepwise*, 10 ocorreram três ou mais vezes, sendo essas selecionadas como entrada do modelo de RNA de CPN. Os 10 preditores selecionados foram: pressão na superfície (9 meses anteriores) e TSM (3 meses anteriores) na região do Anticiclone Subtropical do Atlântico Sul, pressão na superfície (3 meses anteriores) e TSM (9 meses anteriores) na região do Anticiclone Subtropical do Pacífico Sul, vento meridional (atraso de 4 meses) na região do Jato de Baixos Níveis, vento zonal (9 meses anteriores) na região do Jato Subtropical, umidade específica a 700 hPa (5 meses anteriores) e velocidade vertical a 500 hPa (5 meses anteriores) na região nomeada como área 1 que representa a Zona de Convergência do Atlântico Sul e os geopotenciais a 200 hPa (4 meses anteriores) e 500 hPa (9 meses anteriores) na região do Paraná.

4.2.4 Seleção das variáveis de entrada do modelo de precipitação sazonal

As variáveis de entrada do modelo de RNA de CPN para a previsão de precipitação sazonal foram obtidas por novas execuções do método *stepwise*. O resultado dos cinco ciclos de execuções do método de seleção de variáveis para a precipitação sazonal utilizando a RNA de CPN pode ser visualizado na Figura 4.6

Observa-se pela Figura 4.6 que das 35 variáveis avaliadas pelo método *stepwise*, 15 ocorreram três ou mais vezes. Os 15 vetores foram avaliados novamente pelo método de seleção de variáveis, pois considerou-se que ainda havia uma grande quantidade de entradas e possivelmente irrelevantes para o modelo de RNA. O resultado da segunda etapa do método de seleção de variáveis regressão *stepwise* para a precipitação sazonal é mostrado na Figura 4.7.

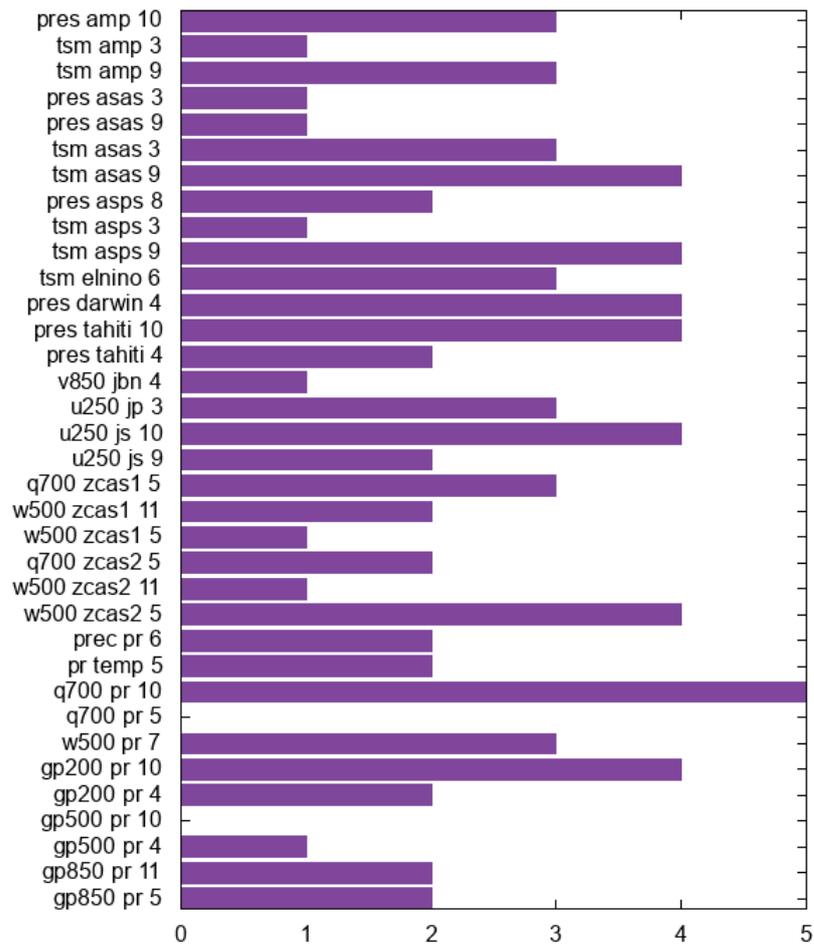


Figura 4.6: Resultado dos cinco ciclos do método de seleção de variáveis regressão *stepwise* para a precipitação sazonal utilizando a RNA de CPN.

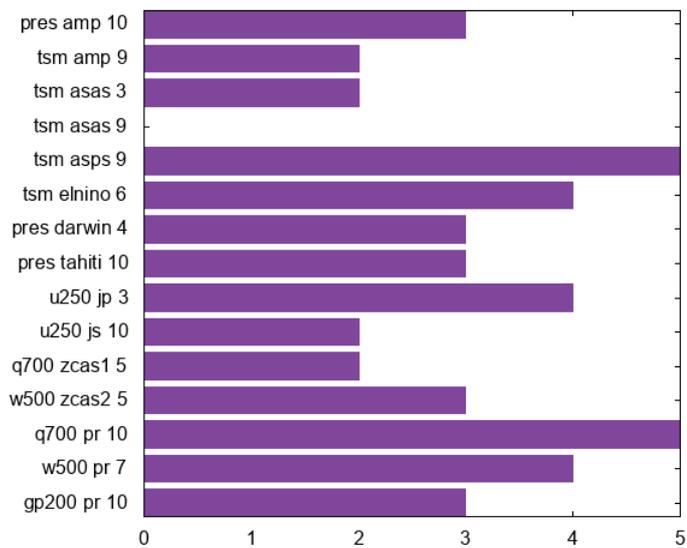


Figura 4.7: Resultado dos cinco ciclos da segunda etapa do método de seleção de variáveis regressão *stepwise* para a precipitação sazonal utilizando a RNA de CPN.

O resultado da segunda etapa do método *stepwise* indica que das 15 variáveis analisadas, 10 delas repetiram-se em três ou mais ciclos, sendo que a TSM em 9 meses anteriores na região do Anticiclone Subtropical do Pacífico Sul e a umidade específica a 700 hPa com 10 meses passados na região do Paraná mostraram-se relevantes para o modelo de previsão sazonal de precipitação utilizando a RNA de CPN, pois foram selecionadas como possíveis preditores nas cinco execuções do método.

4.2.5 Avaliação do desempenho das previsões pela rede neural de contra-propagação

Após selecionar os vetores de entrada, foi avaliado o desempenho da RNA de CPN na previsão climática do Paraná (Tabela 4.2). O número de neurônios na camada de entrada foram: 10, 14, 10 e 10 para as previsões de temperatura mensal, temperatura sazonal, precipitação mensal e precipitação sazonal, respectivamente. Em relação ao número de neurônios da camada oculta, manteve-se igual ao utilizado no método *stepwise*, ou seja, 150 neurônios. Os gráficos de dispersão entre os dados simulados pela rede neural e os dados observados no experimento são apresentados nas Figuras 4.8 a 4.11.

Tabela 4.2: Desempenho da previsão climática por redes neurais de contra-propagação.

Previsão	R	CNS	REQM	EMA
Temperatura mensal	0,87	0,72	1,40 °C	1,09 °C
Temperatura sazonal	0,94	0,88	0,94 °C	0,75 °C
Precipitação mensal	0,46	0,21	63 mm/mês	51 mm/mês
Precipitação sazonal	0,53	0,17	39 mm/mês	30 mm/mês

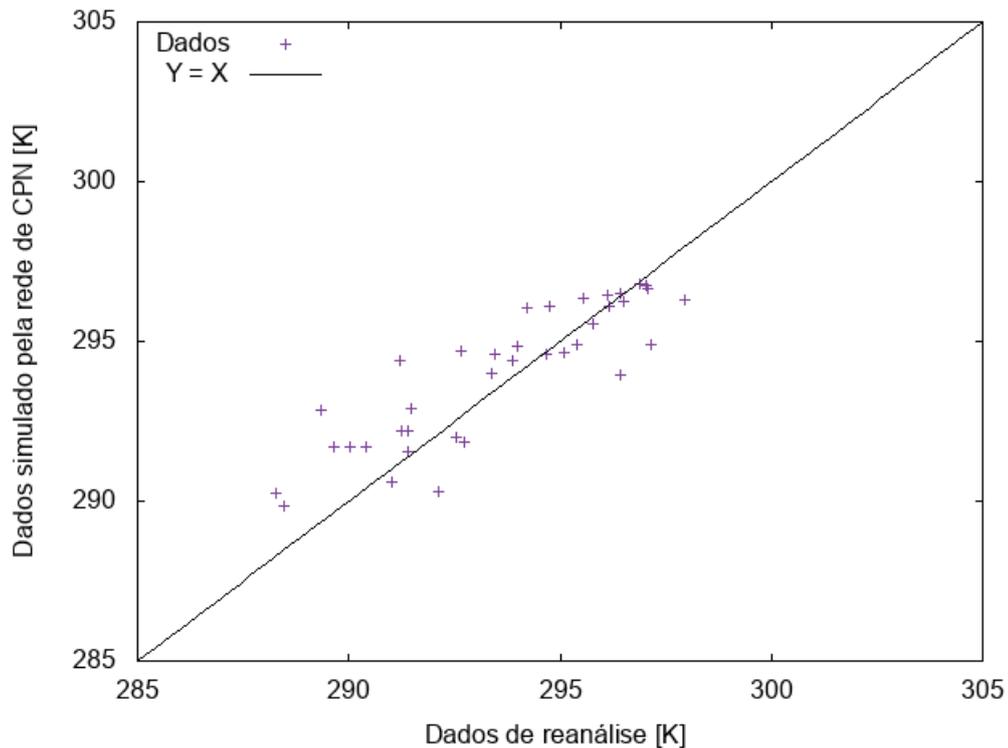


Figura 4.8: Gráfico de dispersão entre os dados de reanálise e o resultado da rede de contra-propagação para a temperatura mensal.

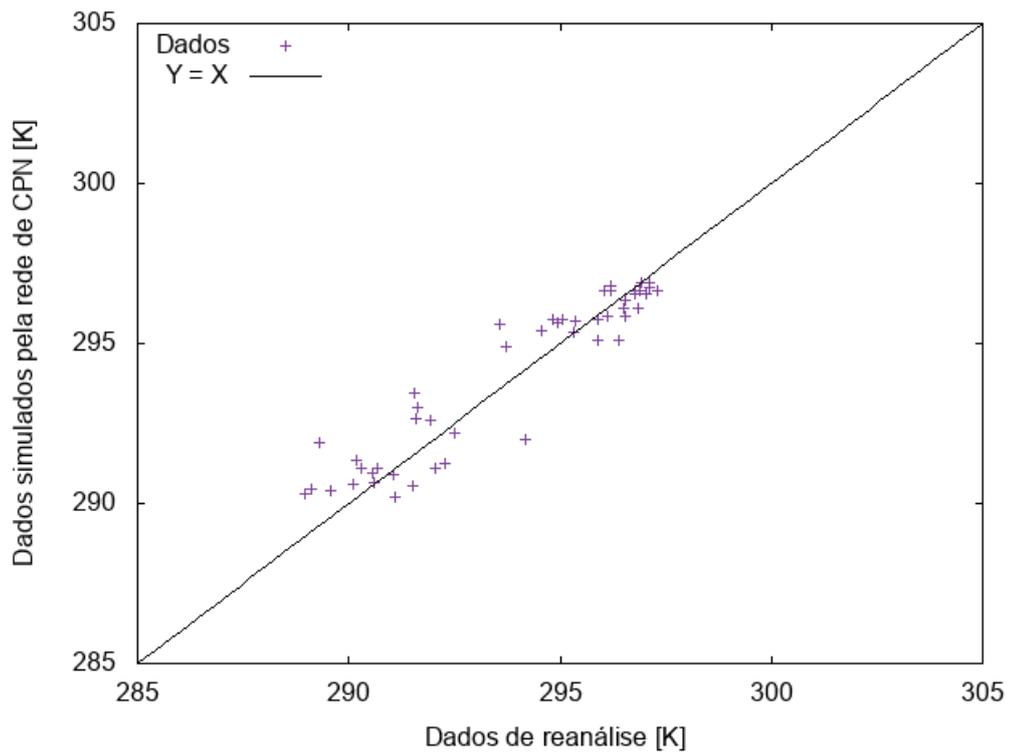


Figura 4.9: Gráfico de dispersão entre os dados de reanálise e o resultado da rede de contra-propagação para a temperatura sazonal.

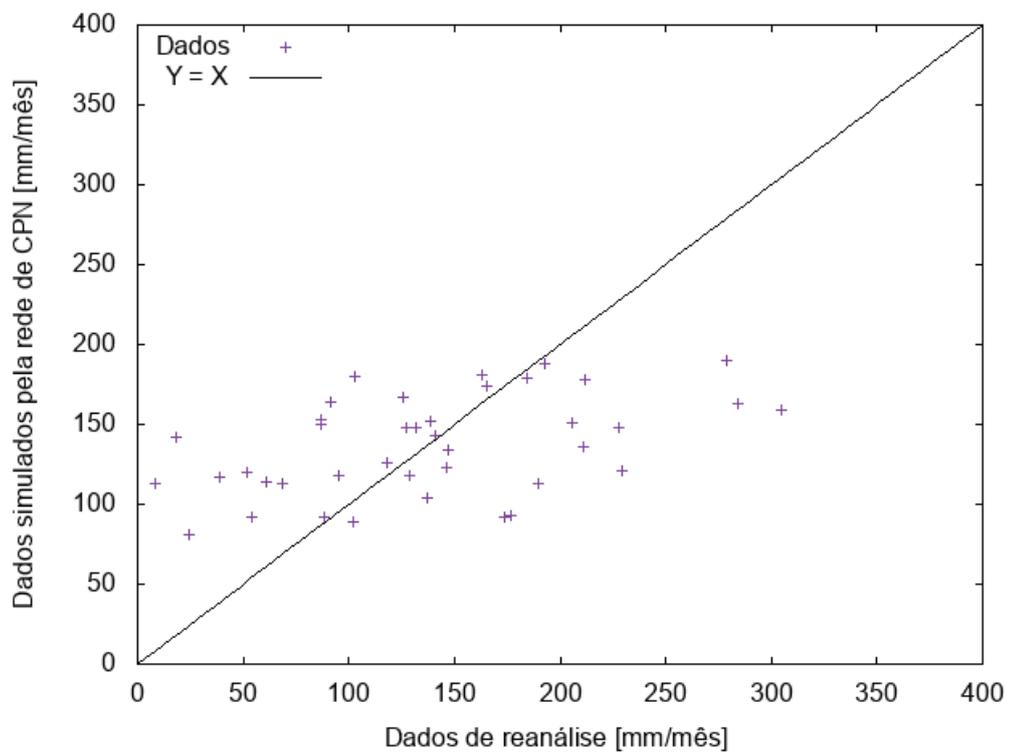


Figura 4.10: Gráfico de dispersão entre os dados de reanálise e o resultado da rede de contra-propagação para a precipitação mensal.

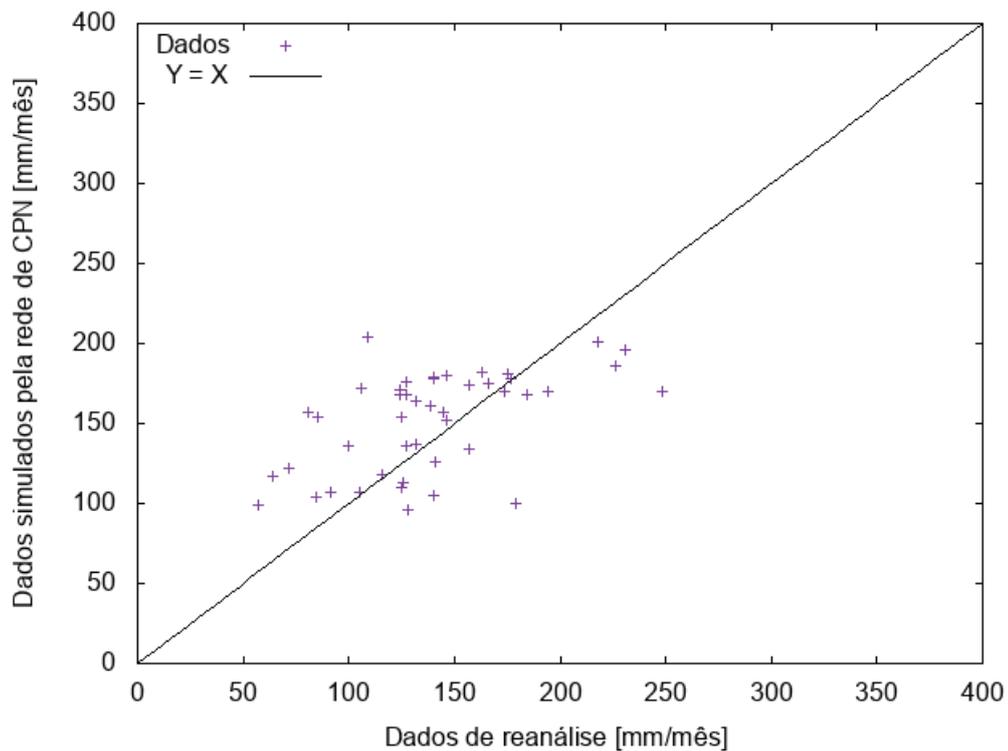


Figura 4.11: Gráfico de dispersão entre os dados de reanálise e o resultado da rede de contra-propagação para a precipitação sazonal.

4.3 PREVISÃO ATRAVÉS DA REDE NEURAL RBF

4.3.1 Seleção das variáveis de entrada do modelo de temperatura mensal

O procedimento da seleção de variáveis para a rede RBF foi o mesmo adotado para a rede neural de CPN. No entanto, o tempo de processamento do método de regressão *stepwise* foi muito superior. Enquanto o tempo de processamento de um ciclo do método de seleção de variáveis com a rede CPN levou aproximadamente 7 horas, com a rede RBF levou aproximadamente 22 horas. A duração de execução do método aumentava com o número de neurônios na camada oculta. Por isso, não foi possível avaliar o desempenho da rede RBF com um grande número de neurônios na camada oculta, limitando-se a 50 neurônios.

Para a temperatura mensal, foram avaliados 30 possíveis preditores, mostrados na Tabela 3.3, pelo método de seleção de variáveis regressão *stepwise*. O resultado do método para selecionar as entradas do modelo de RNA RBF é exibido na Figura 4.12.

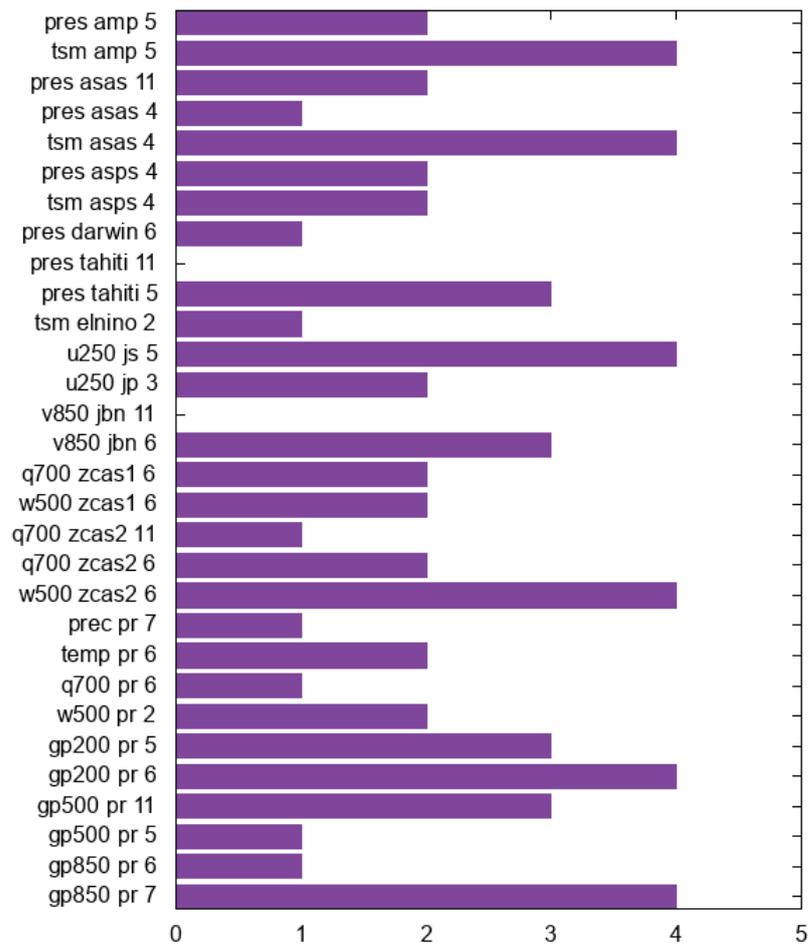


Figura 4.12: Resultado dos cinco ciclos do método de seleção de variáveis regressão *stepwise* para a temperatura mensal utilizando a RNA RBF.

Observa-se pela Figura 4.12 que a maioria das variáveis foi selecionada pelo menos uma vez durante os cinco ciclos do método *stepwise* e 10 delas repetiram-se por três ou mais vezes. Notou-se que poderia ser realizado mais um ciclo de execuções do método *stepwise* apenas com as 10 variáveis, pois o comportamento da REQM continuava sendo de declínio e ascensão à medida que os preditores iam sendo apresentados à rede neural, indicando que algumas variáveis sejam irrelevantes para o modelo. Portanto, foi realizada a segunda etapa do método de seleção de variáveis apenas com os 10 possíveis preditores e o resultado é mostrado na Figura 4.13.

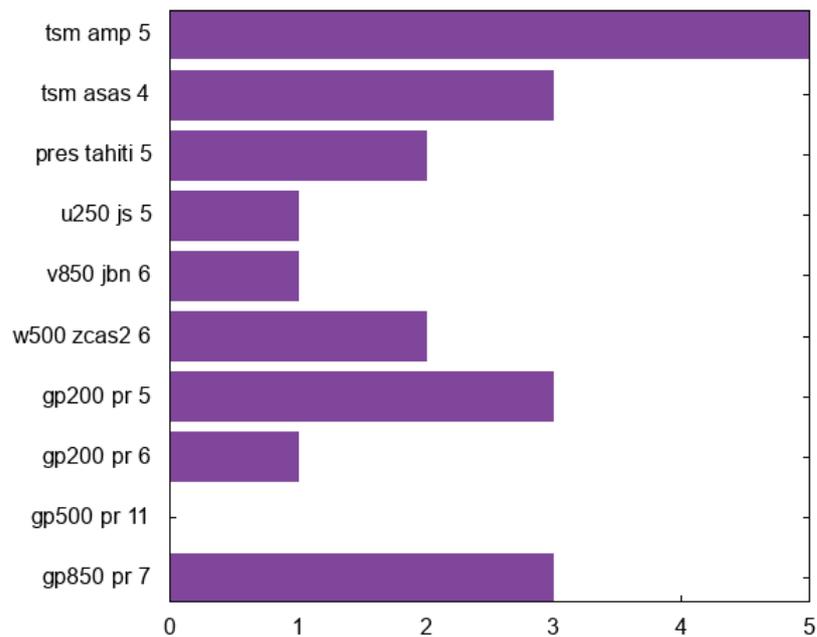


Figura 4.13: Resultado dos cinco ciclos da segunda etapa do método de seleção de variáveis regressão *stepwise* para a temperatura mensal utilizando a RNA RBF.

Após a segunda etapa de ciclos de execuções do método *stepwise*, quatro variáveis foram escolhidas como vetores de entrada da RNA. As variáveis selecionadas foram: TSM (5 meses anteriores) da região do Anticiclone Móvel Polar, TSM (4 meses) na região do Anticiclone Subtropical do Atlântico Sul e geopotenciais a 200 hPa (5 meses) e 850 hPa (7 meses) no Paraná. A quantidade de vetores de entrada da RNA RBF foi muito menor que a da rede CPN. No entanto, essa combinação de variáveis foi a que apresentou o melhor dos testes realizados.

4.3.2 Seleção das variáveis de entrada do modelo de temperatura sazonal

O resultado dos cinco ciclos do método de seleção de variáveis regressão *stepwise* para a temperatura sazonal utilizando a RNA RBF com 50 neurônios na camada oculta é mostrado na Figura 4.14. Das 34 variáveis analisadas, oito ocorreram em três ou mais execuções do método *stepwise*, sendo essas selecionadas como vetores de entrada da RNA RBF, que são: TSM (5 e 11 meses de antecedência) na região do Anticiclone Móvel Polar, TSM (4 meses) na região do Anticiclone Subtropical do Atlântico Sul, pressão na superfície (4 meses) na região do Anticiclone Subtropical do Pacífico Sul, TSM (8 meses) na região do El Niño, pressão na superfície (5 meses) em Tahiti, vento zonal a 250 hPa (5 meses) na região do Jato Subtropical e velocidade vertical a 500 hPa (9 meses) no Paraná.

Nota-se que a TSM na região do Anticiclone Móvel Polar é um preditor importante tanto na previsão da temperatura sazonal do Paraná quanto na mensal, pois foi selecionado como vetor de entrada em todas as redes neurais estudadas.

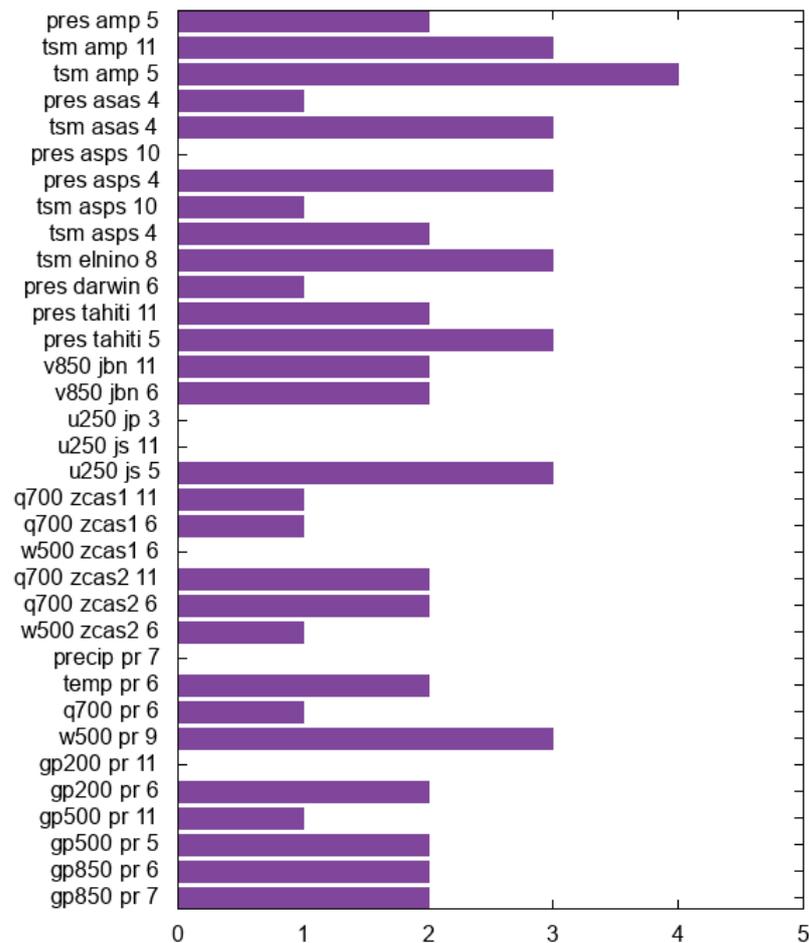


Figura 4.14: Resultado dos cinco ciclos do método de seleção de variáveis regressão *stepwise* para a temperatura sazonal utilizando a RNA RBF.

4.3.3 Seleção das variáveis de entrada do modelo de precipitação mensal

O resultado da sequência de execuções do método *stepwise* para a precipitação mensal utilizando a RNA RBF é apresentado na Figura 4.15, onde pode ser observado que 6 dos 34 possíveis preditores ocorreram três ou mais vezes durante a seleção dos vetores de entrada da RNA, são elas: TSM (3 meses anteriores) na região do Anticiclone Subtropical do Pacífico Sul; umidade específica (5 meses anteriores) e velocidade vertical (5 meses anteriores) na região nomeada como área 1 que representa a Zona de Convergência do Atlântico Sul; umidade específica (5 meses anteriores) na área 2 da Zona de Convergência do Atlântico Sul; precipitação (1 mês anterior) e velocidade vertical (5 meses anteriores) na região do Paraná.

Entre as variáveis selecionadas como entrada para os modelos de redes neurais, apenas a umidade específica a 700 hPa e a velocidade vertical a 500 hPa na região escolhida como área 1 que representa a Zona de Convergência do Atlântico Sul foram usadas nas duas RNA estudadas.

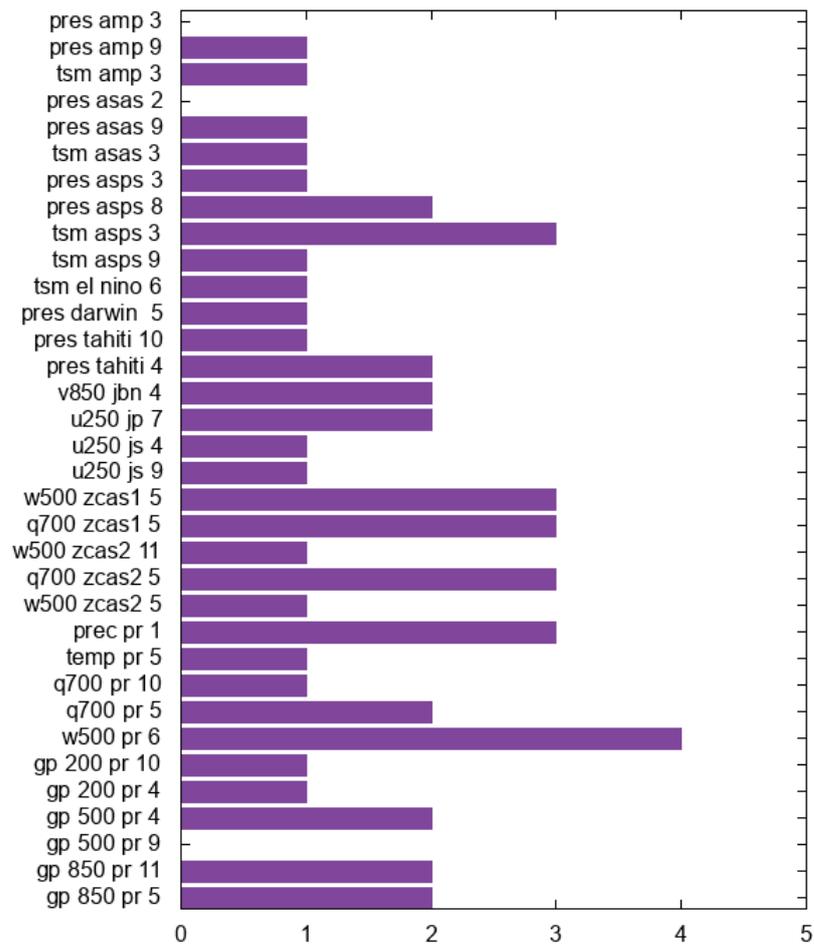


Figura 4.15: Resultado dos cinco ciclos do método de seleção de variáveis regressão *stepwise* para a precipitação mensal utilizando a RNA RBF.

4.3.4 Seleção das variáveis de entrada do modelo de precipitação sazonal

A Figura 4.16 apresenta o método de seleção de variáveis regressão *stepwise* para a precipitação sazonal utilizando a RNA RBF com 50 neurônios na camada oculta. Nota-se que os preditores do modelo de previsão da precipitação trimestral ficaram mais evidentes, visto que muitas variáveis não foram selecionadas nos cinco ciclos da execução do método *stepwise*.

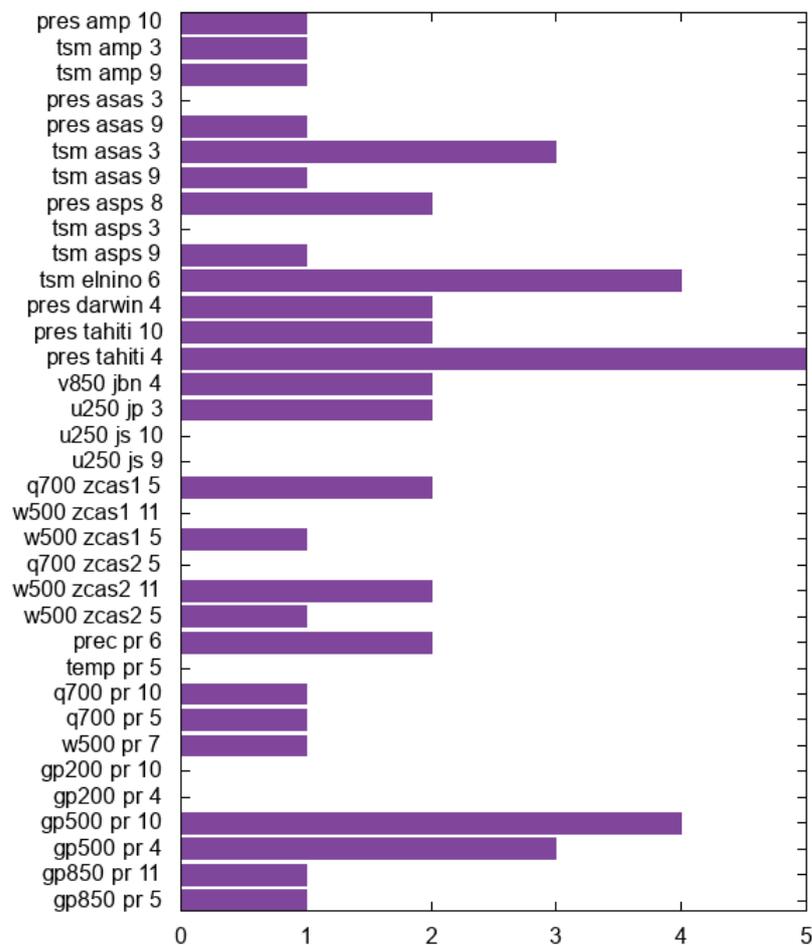


Figura 4.16: Resultado dos cinco ciclos do método de seleção de variáveis regressão *stepwise* para a precipitação sazonal utilizando a RNA RBF.

Das 35 variáveis avaliadas pelo método, cinco vetores foram utilizados como entrada da rede RBF para a previsão sazonal de precipitação, a saber: TSM (3 meses anteriores) na região do Anticiclone Subtropical do Atlântico Sul, TSM (6 meses) na região do El Niño, pressão na superfície (4 meses) no Tahiti e geopotenciais a 500 hPa (4 e 10 meses) no Paraná. Sendo a TSM (6 meses anteriores) na região do El Niño selecionada como entrada tanto na rede de CPN quanto na rede RBF para a previsão sazonal de precipitação.

4.3.5 Avaliação do desempenho das previsões pela rede neural RBF

Após selecionar os vetores de entrada pelo método *stepwise*, o desempenho da rede neural RBF foi avaliado na previsão climática do Paraná utilizando 50 neurônios na camada oculta, conforme mostrado na Tabela 4.3. Já o número de neurônios na camada de entrada foram: 4, 8, 6 e 5 para as previsões da temperatura mensal, temperatura sazonal, precipitação mensal e precipitação sazonal, respectivamente. Os gráficos de dispersão entre os dados simulados pela rede neural e os dados observados de temperatura e precipitação mensal e sazonal são apresentados da Figura 4.17 a 4.20.

Tabela 4.3: Desempenho da previsão climática por redes neurais RBF.

Previsão	R	CNS	REQM	EMA
Temperatura mensal	0,92	0,84	1,09 °C	0,84 °C
Temperatura sazonal	0,96	0,92	0,67 °C	0,53 °C
Precipitação mensal	0,63	0,39	45 mm/mês	36 mm/mês
Precipitação sazonal	0,76	0,54	30 mm/mês	24 mm/mês

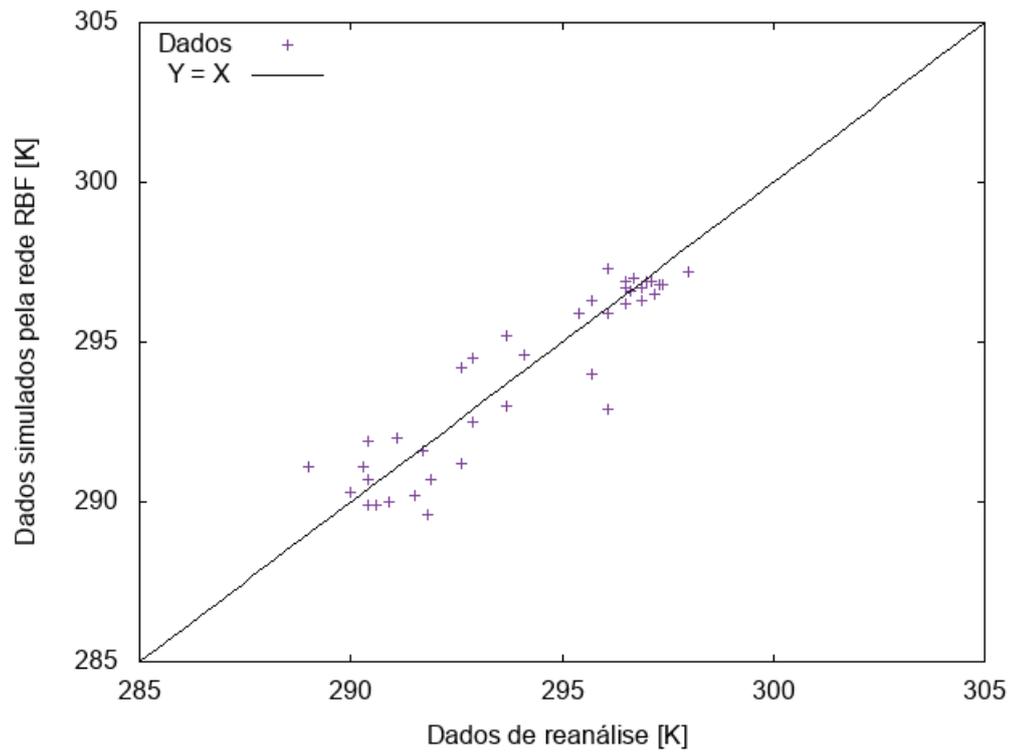


Figura 4.17: Gráfico de dispersão entre os dados de reanálise e o resultado da rede RBF para a temperatura mensal.

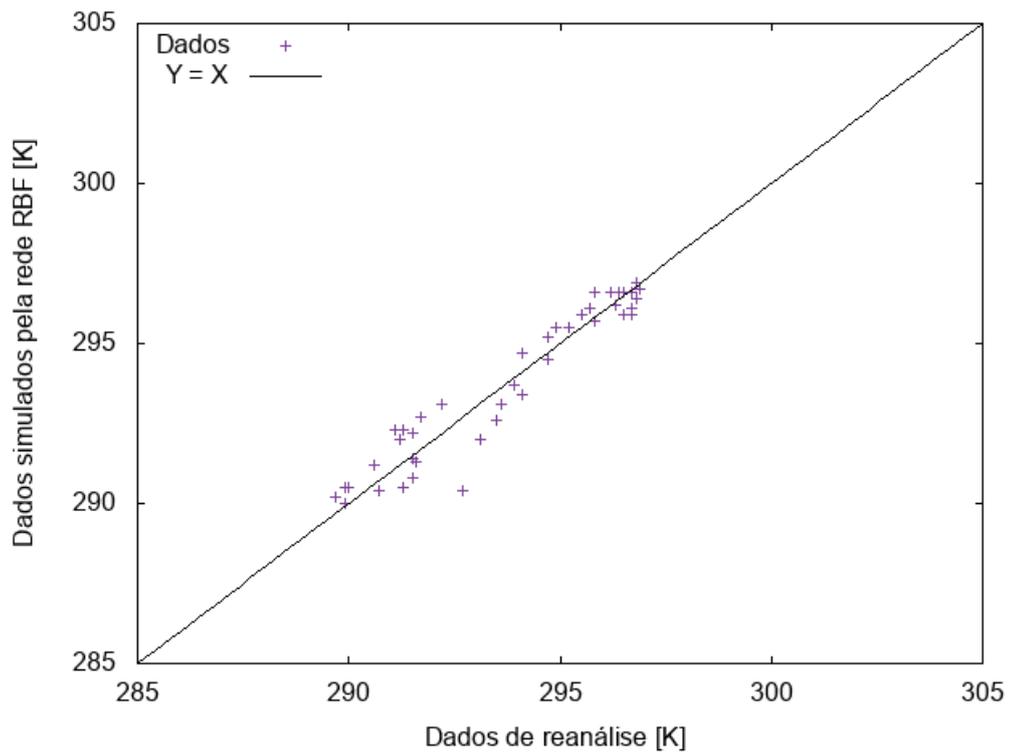


Figura 4.18: Gráfico de dispersão entre os dados de reanálise e o resultado da rede RBF para a temperatura sazonal.

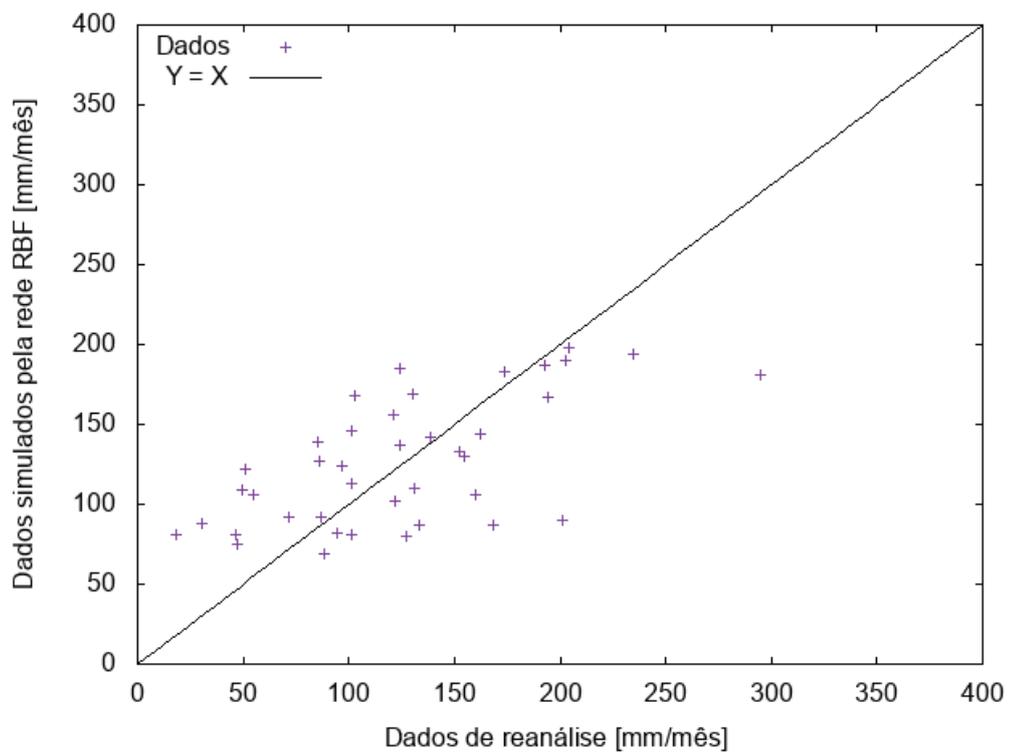


Figura 4.19: Gráfico de dispersão entre os dados de reanálise e o resultado da rede neural RBF para a precipitação mensal.

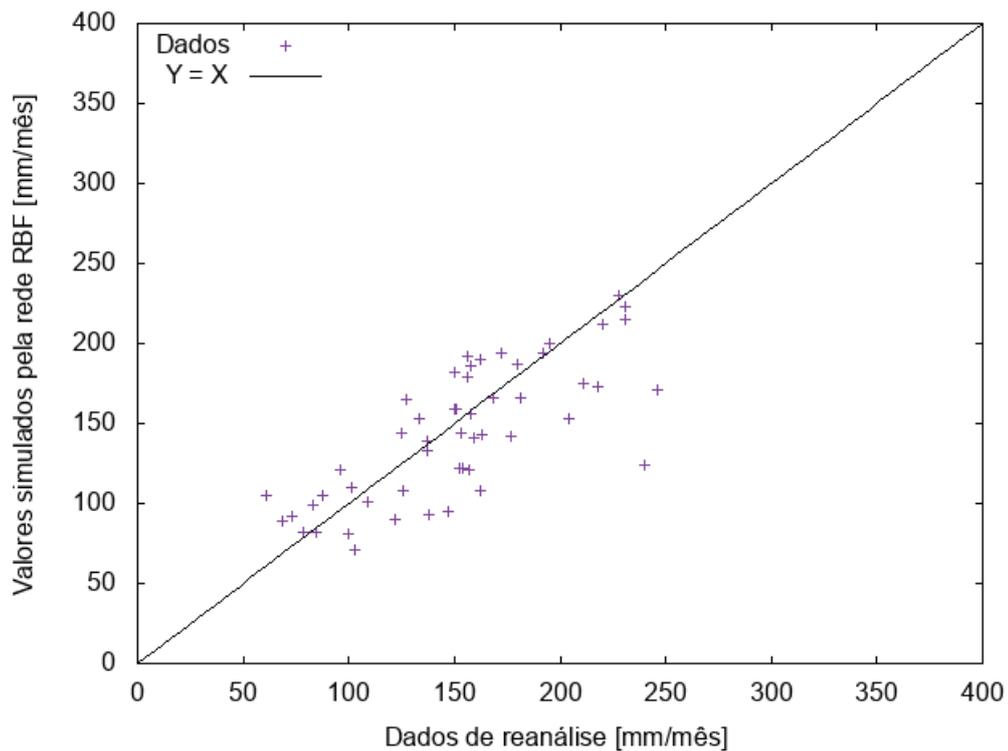


Figura 4.20: Gráfico de dispersão entre os dados de reanálise e o resultado da rede RBF para a precipitação sazonal.

4.4 DISCUSSÃO DOS RESULTADOS

Consolidando todos os resultados dos experimentos de previsão climática obtidos pelas técnicas de persistência, rede neural de CPN e a rede RBF em uma única tabela, como a Tabela 4.4, a seguir, é possível comparar o desempenhos das três metodologias empregadas para a previsão.

Tabela 4.4: Comparação dos resultados obtidos com diferentes técnicas para a previsão climática do Paraná.

Previsão	Técnica	R	CNS	REQM	EMA
Temperatura mensal	Persistência	0,77	0,54	1,98 °C	1,59 °C
	RNA de CPN	0,87	0,72	1,40 °C	1,09 °C
	RNA RBF	0,92	0,84	1,09 °C	0,84 °C
Temperatura sazonal	Persistência	-0,01	-1,02	3,67 °C	3,22 °C
	RNA de CPN	0,94	0,88	0,94 °C	0,75 °C
	RNA RBF	0,96	0,92	0,67 °C	0,53 °C
Precipitação mensal	Persistência	0,35	-0,30	75 mm/mês	60 mm/mês
	RNA de CPN	0,46	0,21	63 mm/mês	51 mm/mês
	RNA RBF	0,63	0,39	45 mm/mês	36 mm/mês
Precipitação sazonal	Persistência	0,09	-0,82	66 mm/mês	54 mm/mês
	RNA de CPN	0,53	0,17	39 mm/mês	30 mm/mês
	RNA RBF	0,76	0,54	30 mm/mês	24 mm/mês

É possível notar que o resultado obtido na previsão climática através da persistência é bastante inferior aos resultados obtidos pelas técnicas de redes neurais. Uma diferença

observada entre os dois métodos avaliados neste estudo é que a técnica de persistência obteve melhores desempenhos para as previsões mensais de temperatura e precipitação, já as redes neurais foram superiores nas previsões sazonais. Esses melhores desempenhos das redes neurais nas estimativas sazonais pode ser explicado pelo fato dos dados serem filtrados pela média móvel de três meses, o que diminui as variações de alta frequência ("ruídos") das séries temporais e facilita a previsão do fenômeno atmosférico. Em relação às redes neurais, nota-se que o desempenho da rede do tipo RBF em todos os experimentos de previsões foi superior ao da rede de CPN.

As estimativas de temperatura mensal e sazonal geradas pelas redes neurais foram satisfatórias, com valores de correlação de Pearson acima de 0,9, principalmente a rede RBF que produziu melhores resultados nas simulações. Os desempenhos das previsões de precipitação foram menores que da temperatura, devido à característica do fenômeno mais variável, que torna difícil sua estimativa.

Como referência para contextualização desses resultados, pode-se citar o trabalho de Abbot e Marohasy (2015) que teve como um de seus objetivos o mesmo ao qual nos propusemos nesse estudo, ou seja, prever as chuvas mensais utilizando redes neurais artificiais, no entanto, a área de estudo foi a bacia hidrográfica Murray-Darling, na Austrália. Os autores obtiveram valores de coeficiente de correlação de Pearson de 0,82 e 0,74 para a previsão da precipitação média de 1 e 3 meses adiante, respectivamente. Tal resultado foi superior ao obtido neste trabalho para a precipitação mensal, no entanto, é inferior ao desempenho obtido na precipitação sazonal com a rede RBF, que alcançou um valor de $R = 0,76$, como pode ser observado na Tabela 4.4.

5 CONCLUSÃO

Previsões climáticas são de grande importância para a sociedade, pois é possível antecipar a ocorrência de eventos extremos (tais como, secas drásticas, tempestades, temperaturas muito baixas ou muito altas, etc) e tomar decisões para diminuir seus impactos. No entanto, a previsão climática envolve períodos superiores a 15 dias o que também está suscetível ao efeito do caos.

A fim de contribuir para a obtenção de resultados satisfatórios na previsão climática, no presente trabalho buscou-se avaliar as técnicas de aprendizado de máquina, que foram as Redes Neurais Artificiais (RNAs), pois apresentam a vantagem de identificar relações não lineares altamente complexas entre variáveis de entrada e saída. As RNAs avaliadas neste estudo foram do tipo de contra-propagação (CPN) e a de função de base radial (RBF).

Para isso, primeiro, foram analisados os processos físicos considerados relevantes na variabilidade climática do estado do Paraná, a partir da literatura. A segunda etapa consistiu em definir as variáveis que representam a ação de cada um desses processos físicos considerados relevantes para a previsão de temperatura e precipitação no Paraná, por exemplo, as ZCAS, que causam as chuvas, foram representadas pela umidade e velocidade vertical. Após definidas as variáveis importantes do modelo de previsão, foi necessário identificar os meses de atrasos que apresentavam os mais significativos valores de correlação cruzada com os preditandos. Por fim, para finalizar a etapa de seleção de entradas da rede, foi aplicado o método de seleção *stepwise* para o conjunto de variáveis candidatas.

Os desempenhos das RNAs foram avaliados através das métricas estatísticas correlação linear de Pearson (R), coeficiente de Nash-Sutcliffe (CNS), raiz do erro quadrático médio (REQM) e erro médio absoluto (EMA). Também, foi aplicada uma técnica simples de previsão, que é a persistência, no intuito de verificar se as RNAs apresentavam resultados superiores.

Analisando-se os desempenhos observados nos experimentos de previsão através das técnicas de persistência e de redes neurais de CPN e RBF, podem-se destacar alguns resultados relevantes dos três métodos de previsão climática.

A persistência produziu resultados insatisfatórios para a previsão climática do Paraná, principalmente para a temperatura sazonal, onde o valor de R foi de -0,01, e para as precipitações mensais e sazonais, cujos valores do R foram 0,35 e 0,09, respectivamente. Já para a temperatura mensal, o desempenho do método de persistência foi um pouco mais satisfatório, no entanto, ainda inferior que o método de redes neurais.

Em relação aos desempenhos das redes neurais, notou-se que a rede RBF obteve maior acurácia que a rede de CPN e que a suavização das séries temporais com médias móveis de três meses permitiu a obtenção de melhores resultados nas previsões de precipitação e temperatura. Portanto, fenômenos atmosféricos que variam bastante são mais complexos para as RNAs identificarem e, conseqüentemente, prever um padrão.

O método de seleção de variáveis *stepwise* identificou diferentes conjuntos de vetores quando utilizou as redes CPN e RBF. Essa dificuldade em se estabelecer um conjunto de preditores relevantes para os modelos de previsão climática deve-se ao fato da complexidade do problema. No entanto, notou-se uma dominância na seleção de alguns fenômenos atmosféricos para a previsão climática, por exemplo, a Temperatura da Superfície do Mar (TSM) na região do Anticiclone Móvel Polar (AMP) com atraso de cinco meses indicou ser um preditor importante para a variação da temperatura mensal e sazonal do Paraná, pois foi a variável que mais vezes foi selecionada pelo método *stepwise*. Já para a previsão da precipitação mensal não foi possível

detectar qual variável tem a maior importância no modelo, visto que não houve destaque em nenhum dos possíveis preditores avaliados e os resultados das redes foram poucos satisfatórios. Já para a precipitação sazonal destaca-se o modelo de RNA RBF, que atingiu valores de R de 0,76, e as variáveis de entrada utilizadas pelo método. Nota-se que a pressão na superfície na região de Tahiti com atraso de 4 meses, a TSM na região do El Niño com atraso de 6 meses, ou seja, variáveis correspondentes ao El Niño Oscilação Sul (ENOS), e a altura geopotencial a 500 hPa no Paraná com atraso de 10 meses foram os principais vetores do modelo de previsão da precipitação sazonal.

Um ponto importante a ser destacado é que, a partir da seleção pelo método *stepwise*, pôde-se identificar as variáveis mais relevantes no controle de processos climáticos em uma região específica, no caso, o estado do Paraná. Essa seleção pode vir a ser útil na orientação de estudos que visem não apenas prever variações climáticas, mas também compreender as relações físicas que controlam suas variações.

Em conclusão, este estudo mostrou que as RNAs RBF são mais adequadas que as CPN nas estimativas de temperatura e precipitação do Paraná. Os resultados mostram que as redes neurais têm desempenhos melhores quando os dados são suavizados no tempo. Dados muito ruidosos, como precipitação mensal, são mais difíceis de se prever. Portanto, como as séries de temperatura são menos ruidosas que a precipitação, os modelos de previsão de temperatura, em geral, tiveram desempenhos melhores que os modelos de previsão de precipitação.

6 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Dada a importância do estudo de previsões climáticas e do crescente uso das RNAs no campo dos estudos atmosféricos cabem aqui algumas recomendações para trabalhos futuros nessa área.

Recomenda-se avaliar outras variações atmosféricas, como a Oscilação de Madden e Julian, Oscilação da Antártica, El Niño Central e a Baixa do Chaco, que podem estar relacionadas com a variabilidade climática no estado do Paraná.

No que se refere à seleção dos preditores relevantes do modelo de redes neurais, sugere-se utilizar um método não-linear para identificar as correlações entre as variáveis relevantes com suas defasagens e os preditandos, que são a temperatura e precipitação do Paraná. Também recomenda-se usar outros métodos para a seleção de variáveis, por exemplo, a Análise de Componentes Principais não linear.

Outras recomendações para trabalhos futuros são: refinar as previsões de temperatura e precipitação para regiões menores do Paraná e que possuam as mesmas características climáticas e; substituir os dados de reanálise por dados observados de estações meteorológicas.

REFERÊNCIAS

- Abbot, J. e Marohasy, J. (2015). Forecasting of monthly rainfall in the murray darling basin, australia: Miles as a case study. *WIT Transactions on Ecology and the Environment*, 197:149–159.
- Alves, J., CAMPOS, J., MARTINS, E., Muncunill, D., SOMBRA, S. e Souza Filho, F. (2007). Um estudo inter-comparativo de previsão sazonal estatística-dinâmica de precipitação no nordeste do brasil. *Revista Brasileira de Meteorologia*, 22(3):354–372.
- Alves, L. M., Chou, S.-C. e Marengo, J. (2004). Avaliação das previsões de chuvas sazonais do modelo eta climático sobre o brasil. Em *XIII Congresso Brasileiro de Meteorologia*, volume 29. SBMET Fortaleza e CE.
- Alves, M. F., Lotufo, A. D. P. e Lopes, M. L. M. (2013). Seleção de variáveis stepwise aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas. *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, 1(1).
- Anderson, J., Van Den Dool, H., Barnston, A., Chen, W., Stern, W. e Ploshay, J. (1999). Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bulletin of the American Meteorological Society*, 80(7):1349–1362.
- Anochi, J. A. e de Campos Velho, H. F. (2016). Previsão climática de precipitação para a região sul por rede neural autoconfigurada. *Ciência e Natura*, 38:98–104.
- Ayoade, J. e dos Santos, M. (2002). *Introdução à climatologia para os trópicos*. Bertrand Brasil.
- Bjerknes, J. (1969). Atmospheric teleconnections from the equatorial pacific. *Mon. Wea. Rev.*, 97(3):163–172.
- Bodri, L. e Čermák, V. (2000). Prediction of extreme precipitation using a neural network: application to summer flood occurrence in moravia. *Advances in Engineering Software*, 31(5):311–321.
- Branković, Č. e Palmer, T. (1997). Atmospheric seasonal predictability and estimates of ensemble size. *Monthly weather review*, 125(5):859–874.
- Cera, J. C. e Ferraz, S. E. T. (2015). Variações climáticas na precipitação no sul do brasil no clima presente e futuro. *Revista Brasileira de Meteorologia*, 30(1).
- Copernicus (2017). Copernicus climate change service (c3s): Era5: Fifth generation of ecmwf atmospheric reanalyses of the global climate. <https://cds.climate.copernicus.eu/cdsapp#!/home>. Acessado em 04/07/2019.
- ECMWF (2019). Who we are. <https://www.ecmwf.int/en/about/who-we-are>. Acessado em 03/07/2019.
- Fausett, L. (1994). *Fundamentals of neural networks: architectures, algorithms, and applications*. Prentice-Hall, Inc.

- Gholizadeh, M. H. e Darand, M. (2009). Forecasting precipitation with artificial neural networks (case study: Tehran). *Applied Sciences*, 9:1786–1790.
- Grimm, A. M. (2009). Clima da região sul do brasil. *Tempo e clima no Brasil*, 2:259–275.
- Grimm, A. M., Ferraz, S. E. e Gomes, J. (1998). Precipitation anomalies in southern brazil associated with el niño and la niña events. *Journal of climate*, 11(11):2863–2880.
- Hecht-Nielsen, R. (1987). Counterpropagation networks. *Applied optics*, 26(23):4979–4984.
- Hsieh, W. W. (2001). Nonlinear principal component analysis by neural networks. *Tellus A*, 53(5):599–615.
- INMET (2019). Previsão sazonal do inmet. http://www.inmet.gov.br/portal/index.php?r=clima/prev_estocastica. Acessado em 30/05/2019.
- Khalili, N., Khodashenas, S. R., Davary, K., Baygi, M. M. e Karimaldini, F. (2016). Prediction of rainfall using artificial neural networks for synoptic station of mashhad: a case study. *Arabian Journal of Geosciences*, 9(13):624.
- Lee, J., Kim, C.-G., Lee, J., Kim, N. e Kim, H. (2018). Application of artificial neural networks to rainfall forecasting in the geum river basin, korea. *Water*, 10(10):1448.
- Machado, L. H. R., Mendonça, A. M., Mendonça, R. W. B. e Bonatti, J. P. (2010). Impacto da utilização de previsões "defasadas" no sistema de previsão de tempo por conjunto do cptec/inpe. *Revista Brasileira de Meteorologia*, 25(1):54–69.
- Maheshwari, V. (2019). Dimensionality reduction techniques. <https://medium.com/datadriveninvestor/dimensionality-reduction-techniques-27049b5a4c55>. Acessado em 11/07/2019.
- Marzouq, M., Bounoua, Z., El Fadili, H., Mechaqrane, A., Zenkouar, K. e Lakhliai, Z. (2019). New daily global solar irradiation estimation model based on automatic selection of input parameters using evolutionary artificial neural networks. *Journal of cleaner production*, 209:1105–1118.
- Neto, A. S. R. (2010). Agrometeorologia. http://www.agricultura.pr.gov.br/arquivos/File/deral/Prognosticos/agrometeorologia_2010_11.pdf. Acessado em 20/06/2019.
- Nezhad, E. F., Ghalhari, G. F. e Bayatani, F. (2019). Forecasting maximum seasonal temperature using artificial neural networks "tehran case study". *Asia-Pacific Journal of Atmospheric Sciences*, 55(2):145–153.
- Nimer, E. (1989). Climatologia do brasil. *IBGE*.
- Reboita, M. S., Gan, M. A., da Rocha, R. P. e Ambrizzi, T. (2010). Regimes de precipitação na américa do sul: uma revisão bibliográfica. *Revista Brasileira de Meteorologia*, 25(2).
- Sampaio, G. e da Silva Dias, P. L. (2014). Evolução dos modelos climáticos e de previsão de tempo e clima. *Revista USP*, (103):41–54.
- Satyamurty, P., Nobre, C. A. e Dias, P. L. S. (1998). *South America*. Springer.

- Scholz, M., Fraunholz, M. e Selbig, J. (2008). Nonlinear principal component analysis: neural network models and applications. Em *Principal manifolds for data visualization and dimension reduction*, páginas 44–67. Springer.
- Seluchi, M. E. e Garreaud, R. D. (2012). Campos médios e processos físicos associados ao ciclo de vida da baixa do chaco. *Revista Brasileira de Meteorologia*, 27(4).
- Shukla, J., Anderson, J., Baumhefner, D., Brankovic, C., Chang, Y., Kalnay, E., Marx, L., Palmer, T., Paolino, D., Ploshay, J. et al. (2000). Dynamical seasonal prediction. *Bulletin of the American Meteorological Society*, 81(11):2593–2606.
- Silva, I. d., Spatti, D. H. e Flauzino, R. A. (2010). *Redes neurais artificiais para engenharia e ciências aplicadas*, volume 23.
- Tracton, M. S. e Kalnay, E. (1993). Operational ensemble prediction at the national meteorological center: Practical aspects. *Weather and Forecasting*, 8(3):379–398.
- Velasco, I. e Fritsch, J. M. (1987). Mesoscale convective complexes in the americas. *Journal of Geophysical Research: Atmospheres*, 92(D8):9591–9613.
- Wikipedia (2019). Robinson projection. https://en.wikipedia.org/wiki/Robinson_projection. Acessado em 02/07/2019.
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*, volume 100. Academic press.
- World Meteorological Organization, W. M. O. (2014). *El Niño / Southern Oscillation*. Chairperson, Publications Board.

APÊNDICE A – MÉDIAS CLIMATOLÓGICAS E ANOMALIAS DAS VARIÁVEIS PREDITORAS DA RNA

A seguir são apresentadas as médias climatológicas e as anomalias das variáveis predictoras possíveis da RNA, que foram obtidas pelo banco de dados do ECMWF. As variáveis selecionadas são médias espaciais das regiões mostradas na Tabela 3.1 e que podem ocasionar a variação climática no estado do Paraná.

A.1 ANTICLONE SUBTROPICAL DO ATLÂNTICO SUL

O fenômeno Anticlone Subtropical do Atlântico Sul (ASAS) foi representado pela variação temporal da temperatura da superfície do mar (TSM) e da pressão na superfície.

As médias climatológicas e as anomalias da TSM mensais são apresentadas nas Figuras A.1 e A.2.

Pela Figura A.1 pode-se notar que as TSM são mais elevadas nos meses de janeiro, fevereiro, março e abril, quando as médias ficam acima de 22 °C, e o menor valor da temperatura do mar é no mês de setembro. Observa-se que a temperatura do mar tem o mesmo comportamento da temperatura do ar, sendo a resposta defasada em dois meses devido às propriedades físicas da água que faz com que leve mais tempo para resfriar e aquecer.

As anomalias da TSM na região do ASAS são mostradas na Figura A.2. Com base no histograma, observa-se que os valores de anomalias de temperatura do mar estão entre -2 °C e +2 °C.

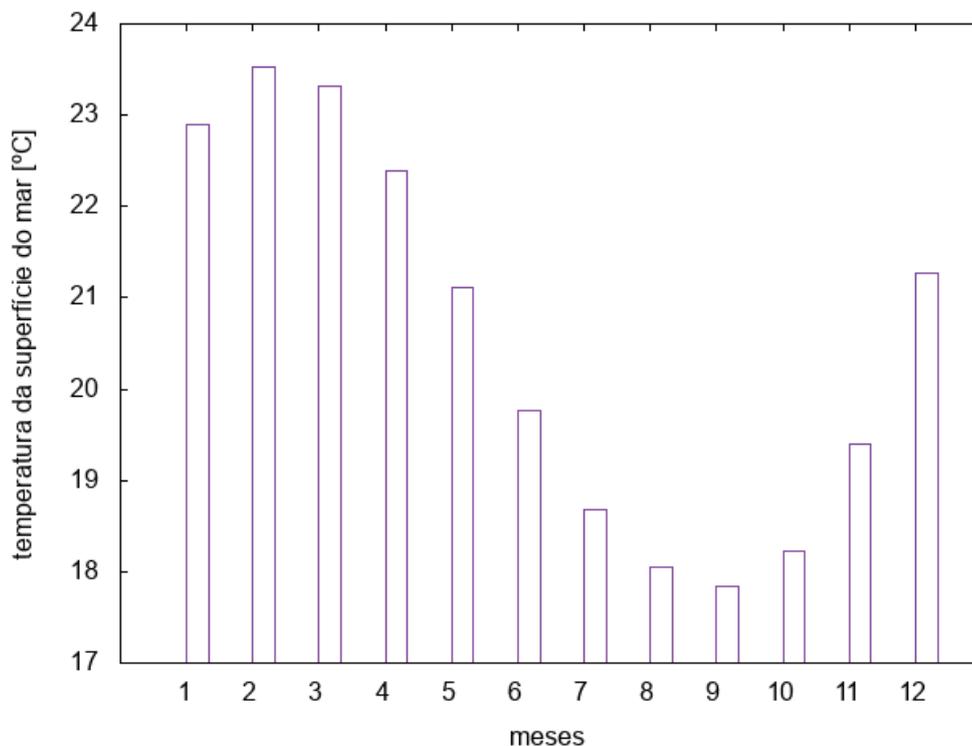


Figura A.1: Médias climatológicas da temperatura da superfície do mar na região do ASAS.

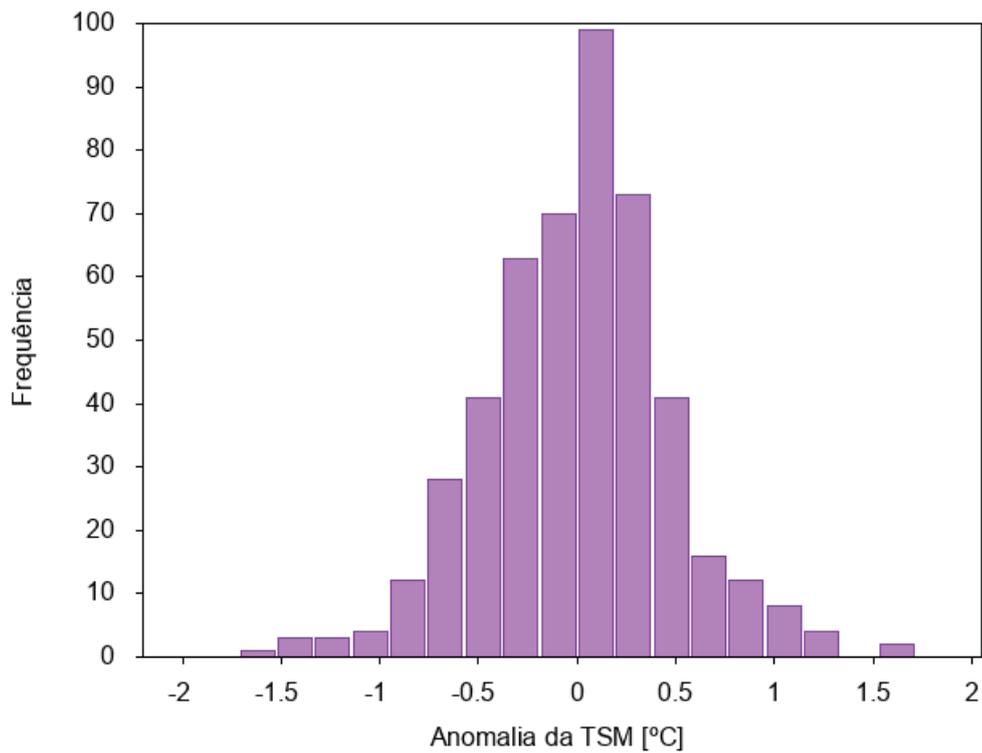


Figura A.2: Anomalias da temperatura da superfície do mar na região do ASAS.

As médias climatológicas e as anomalias das pressões na superfície são mostradas nas Figuras A.3 e A.4, respectivamente.

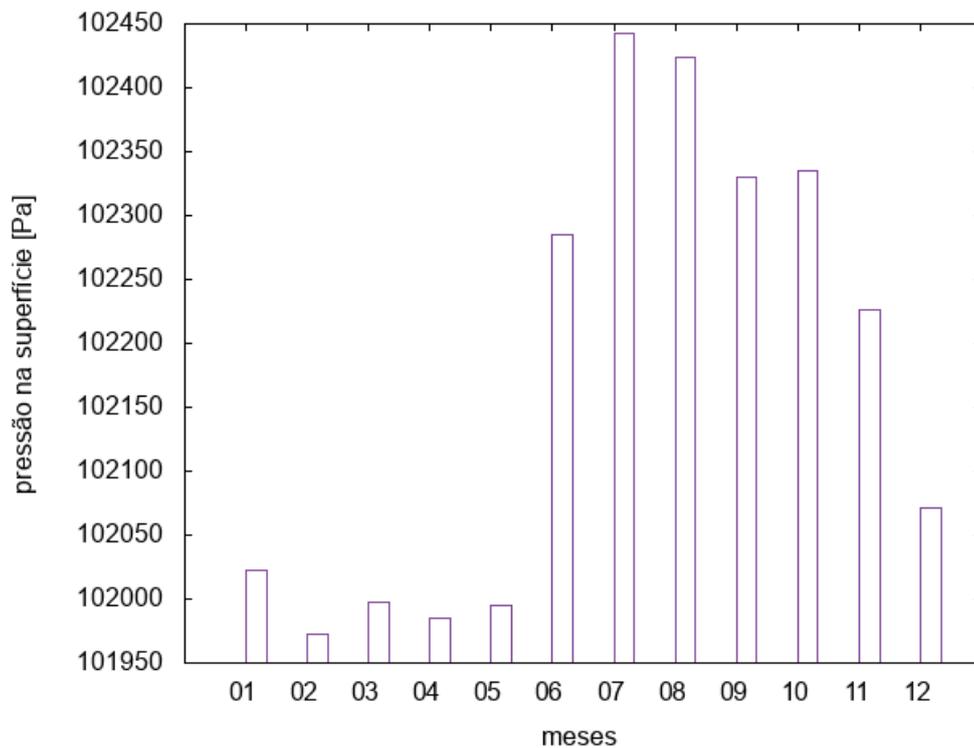


Figura A.3: Médias climatológicas da pressão na superfície na região do ASAS.

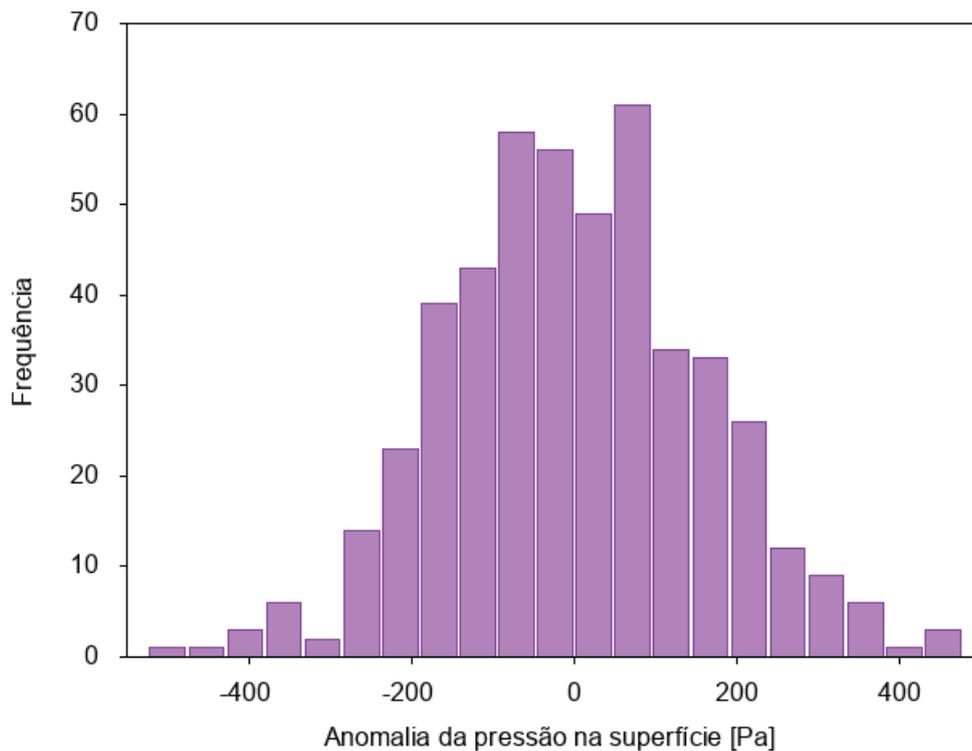


Figura A.4: Anomalias da pressão na superfície na região do ASAS.

O padrão das médias climatológicas da pressão na superfície apresenta uma relação inversamente proporcional ao padrão da TSM, ou seja, nos meses mais quentes os valores da pressão são menores, enquanto nos meses mais frios têm-se os maiores valores de pressão, como pode ser observado no Figura A.3. Esses resultados corroboram o estudo mostrado em Satyamurty et al. (1998) que indica que o centro de alta pressão se aproxima do continente no inverno e afasta no verão, quando os trópicos e subtropicais da América do Sul tem sua estação chuvosa.

As anomalias da pressão na superfície na região do ASAS são apresentadas na Figura A.4. Com base no histograma, nota-se que as anomalias de pressão estão entre +500 Pa e -500 Pa.

A.2 ANTICLONE SUBTROPICAL DO PACÍFICO SUL

Assim como o ASAS, o Anticiclone Subtropical do Pacífico Sul (ASPS) também foi representado pela variação temporal da TSM e da pressão na superfície.

As médias climatológicas e as anomalias das TSM mensais na região do ASPS são apresentadas nas Figuras A.5 e A.6.

O comportamento climático da TSM na região do ASPS é similar ao padrão da TSM na região do ASAS, com valores mais elevados nos meses de janeiro a abril e o menor valor em setembro, como mostrado na Figura A.5.

As anomalias da TSM na região do ASPS são apresentadas na Figura A.6, observa-se que o histograma é assimétrico e os valores estão entre $-1,5\text{ }^{\circ}\text{C}$ e $2\text{ }^{\circ}\text{C}$. Embora as anomalias positivas da TSM da região do ASPS sejam maiores que as negativas, é detectada uma maior frequência de TSM abaixo da média climatológica.

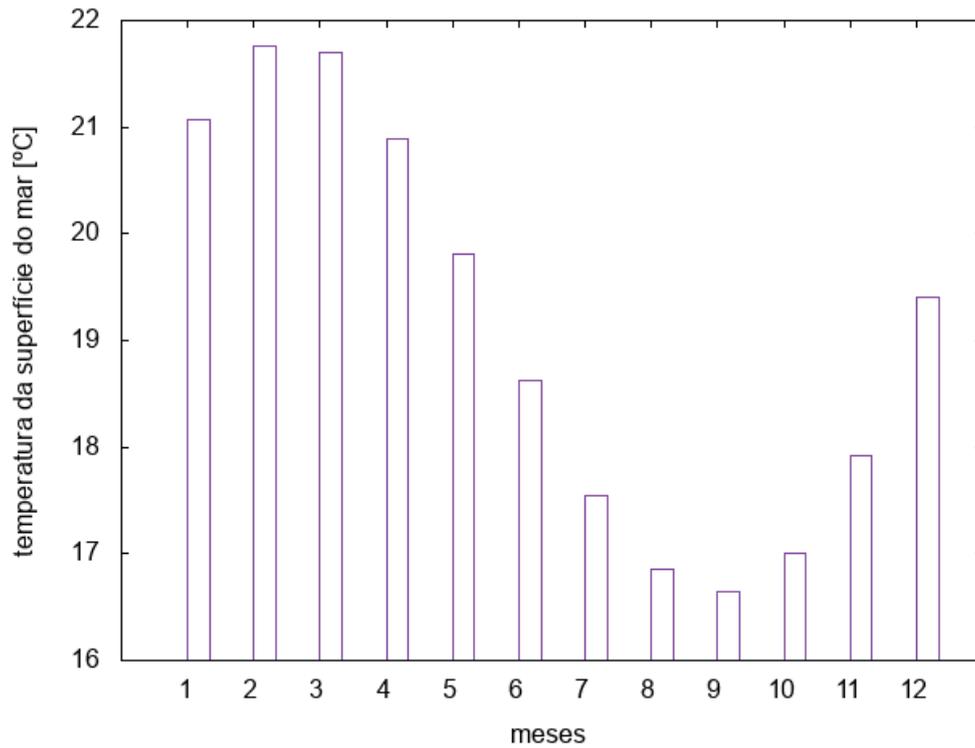


Figura A.5: Médias climatológicas da temperatura da superfície do mar na região do ASPS.

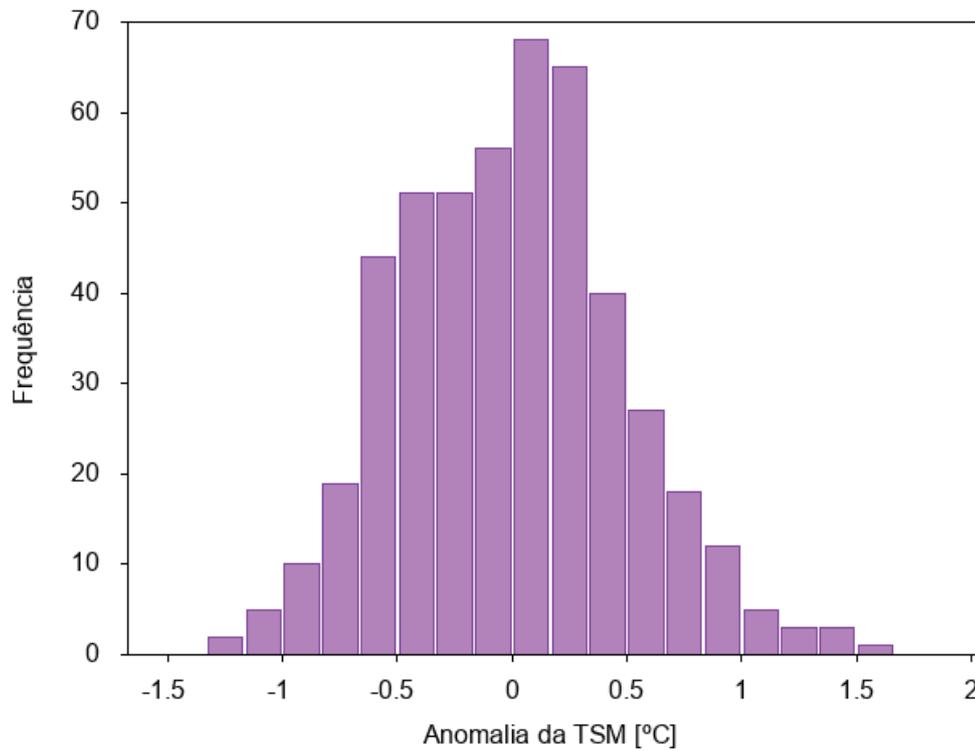


Figura A.6: Anomalias da temperatura da superfície do mar na região do ASPS.

As médias climatológicas e anomalias de pressões na superfície são apresentadas nas Figuras A.7 e A.8.

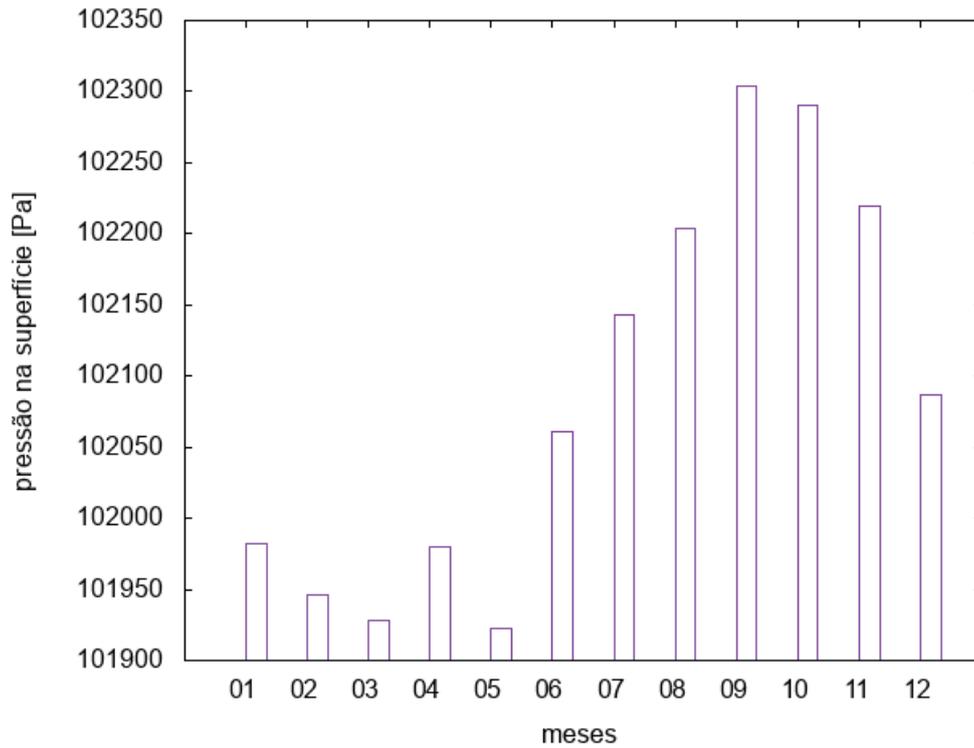


Figura A.7: Médias climatológicas da pressão na superfície na região do ASPS.

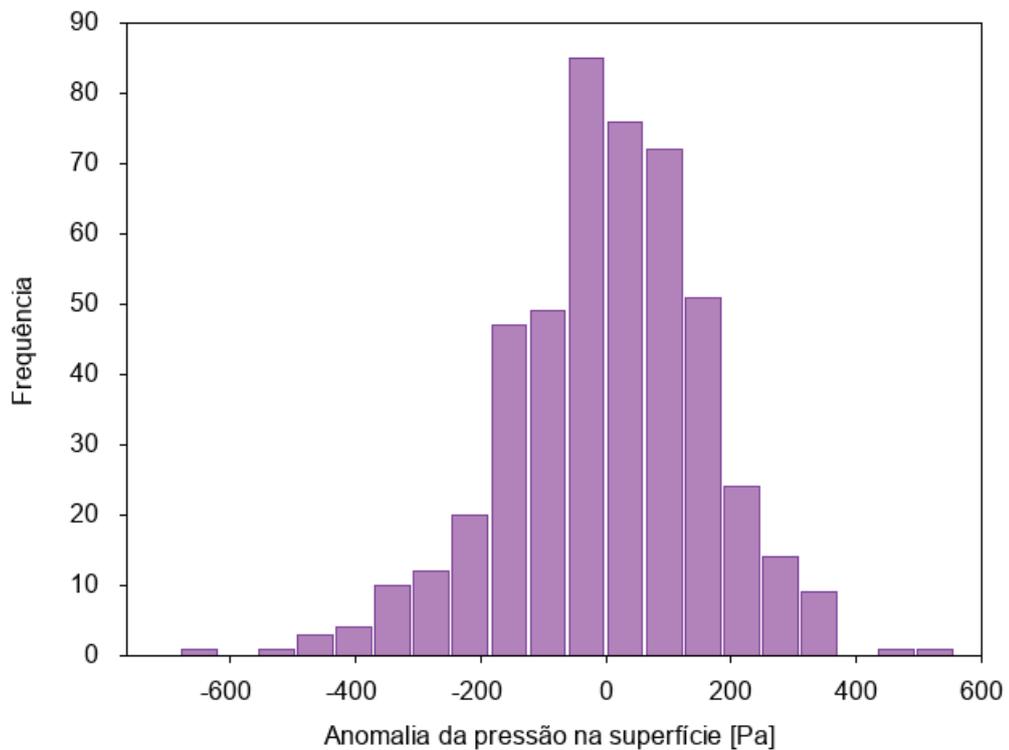


Figura A.8: Anomalias da pressão na superfície na região do ASPS.

Na região do ASPS, os valores de pressões na superfície e de TSM também são inversamente proporcionais, ou seja, no mês de setembro quando a TSM é baixa, tem-se o maior valor de pressão na superfície, como pode ser observado nas Figuras A.5 e A.7. Esse

comportamento é similar ao da pressão na superfície da região do ASAS, indicando que o centro de alta pressão se afasta do continente no verão e aproxima no inverno.

Na Figura A.8 os valores de anomalias de pressão na superfície na região do ASPS estão entre -700 Pa e +600 Pa. Observa-se uma maior frequência em valores de anomalias positivas nessa região.

A.3 ANTICICLONE MÓVEL POLAR

Outro sistema de pressão importante para a variação climática do sul do Brasil é o Anticiclone Móvel Polar (AMP), conforme apresentado na seção 2.2. Esse sistema também foi representado pela variação temporal da TSM e da pressão na superfície.

Nas Figuras A.9 e A.10 são apresentadas as médias climatológicas e as anomalias das TSMs mensais na região do AMP.

Assim como as médias climatológicas da TSM das regiões do ASAS e ASPS, as TSMs na região do AMP são mais elevadas nos meses correspondentes ao verão e início de outono e os menores valores são encontrados no final do inverno e início de primavera. Dentre os sistemas de pressão avaliados neste estudo, o AMP é que apresenta menores valores da TSM, sendo 14 °C o maior valor da média climatológica, o que é justificado pela sua posição mais ao sul do continente.

As anomalias da TSM na região do AMP são mostradas na Figura A.10 e observa-se que os valores estão entre +1,5 °C e -1,5 °C.

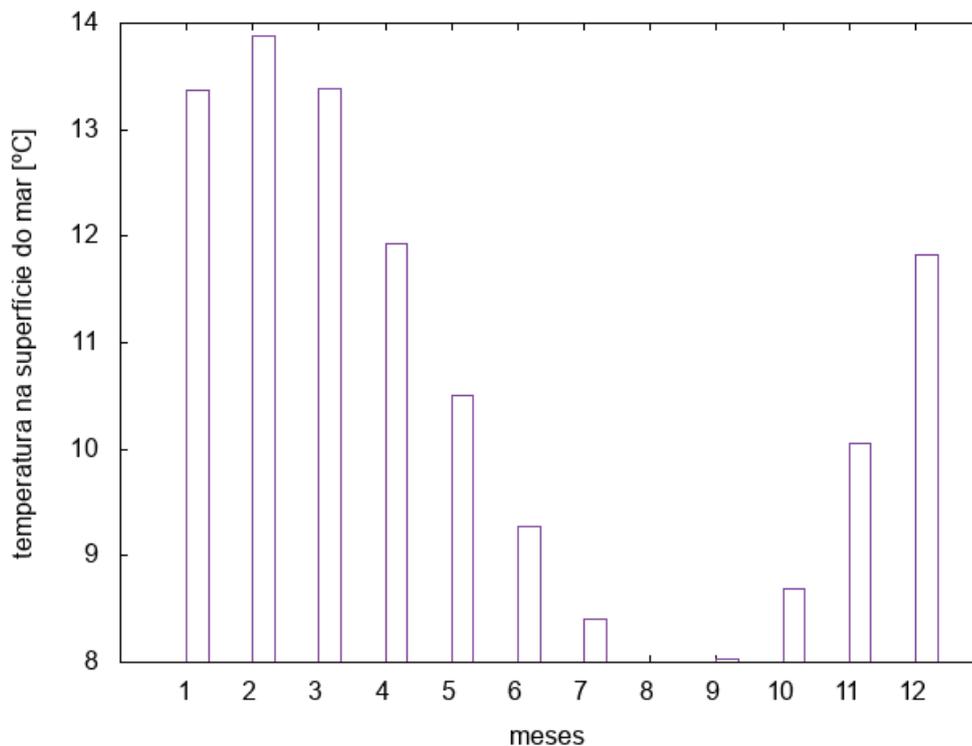


Figura A.9: Médias climatológicas da temperatura da superfície do mar na região do AMP.

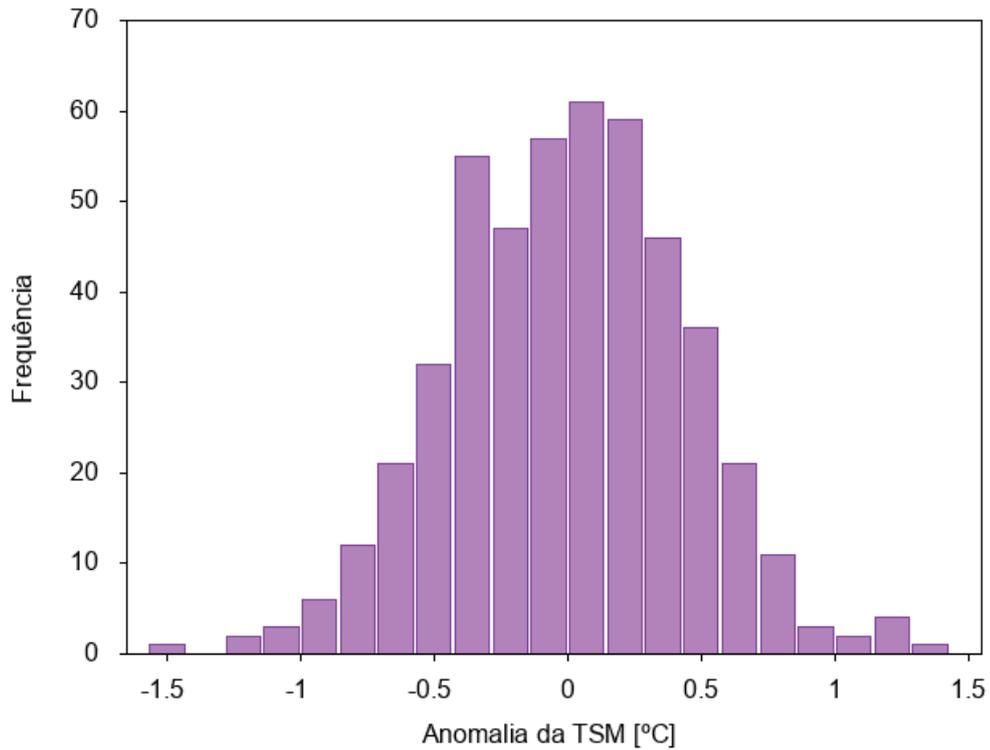


Figura A.10: Anomalias da temperatura da superfície do mar na região do AMP.

Nas Figuras A.11 e A.12 são apresentadas as médias climatológicas e anomalias da pressão na superfície na região do AMP.

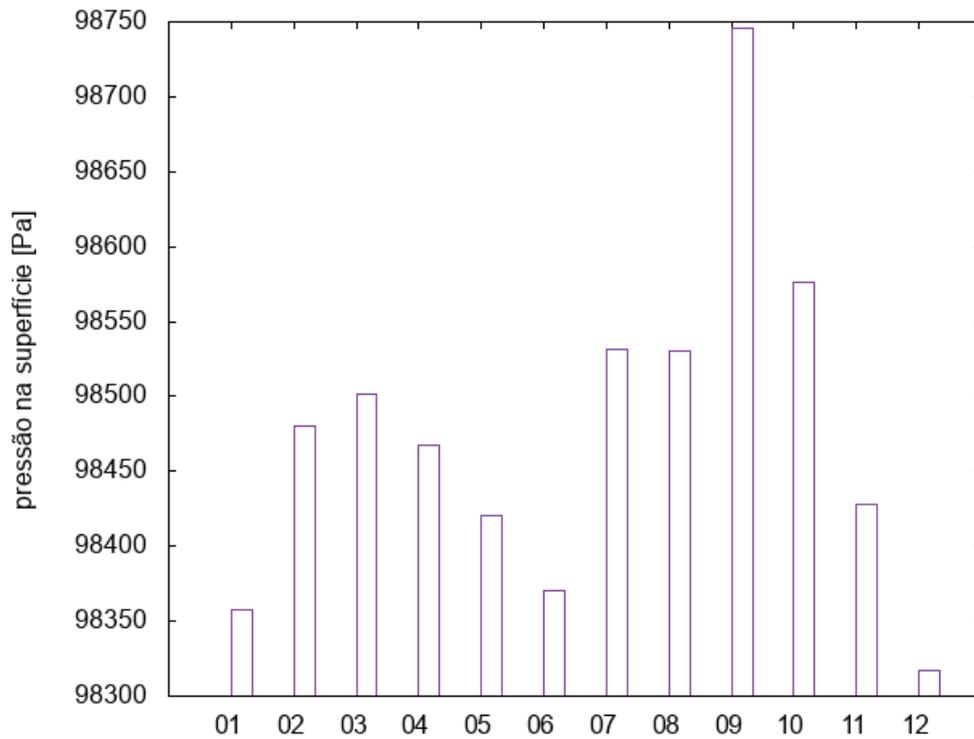


Figura A.11: Médias climatológicas da pressão na superfície na região do AMP.

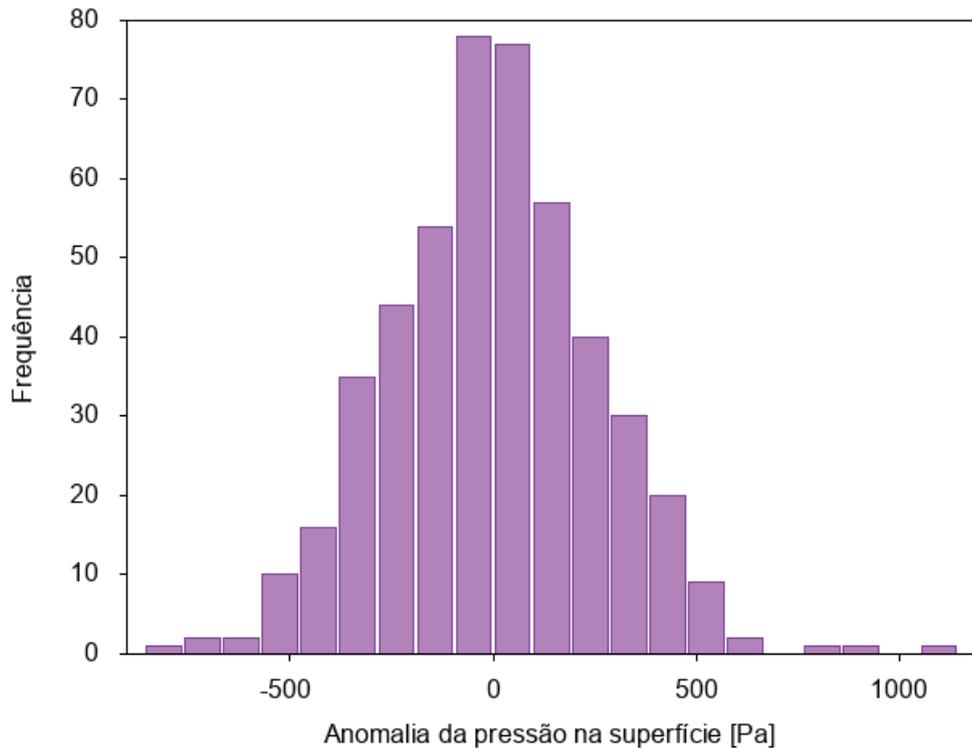


Figura A.12: Anomalias da pressão na superfície na região do AMP.

Na Figura A.11, observa-se que a média climatológica da pressão na superfície é maior no mês de setembro, o que corresponde ao mês que apresenta o menor valor da TSM (Figura A.9). No entanto, nota-se que o comportamento da pressão na superfície na região do AMP não é inversamente proporcional à TSM, pois há um máximo secundário no mês de março, quando a temperatura do mar está alta. Esse padrão da pressão na região do AMP pode estar associado à interferência do continente na variável, pois uma parte da América do Sul foi selecionada para a região do AMP.

As anomalias da pressão na superfície na região do AMP, apresentadas na Figura A.12, estão entre +1100 Pa e -1000 Pa, sendo valores de anomalias maiores que nas regiões do ASAS e ASPS.

A.4 EL NIÑO OSCILAÇÃO SUL

Na Figura A.13 são apresentados as anomalias da TSM na região do El Niño 3.4 e o Índice de Oscilação Sul (IOS) no período de 1979 a 2018.

O IOS é calculado usando as diferenças de pressão padronizadas pelo desvio padrão entre o Tahiti e Darwin. Valores negativos do IOS indicam que a pressão na superfície em Darwin está mais alta que em Tahiti, o que sugere episódios de El Niño. Enquanto valores positivos do IOS são típicos em períodos de La Niña. Esse padrão climático pode ser observado na Figura A.13.

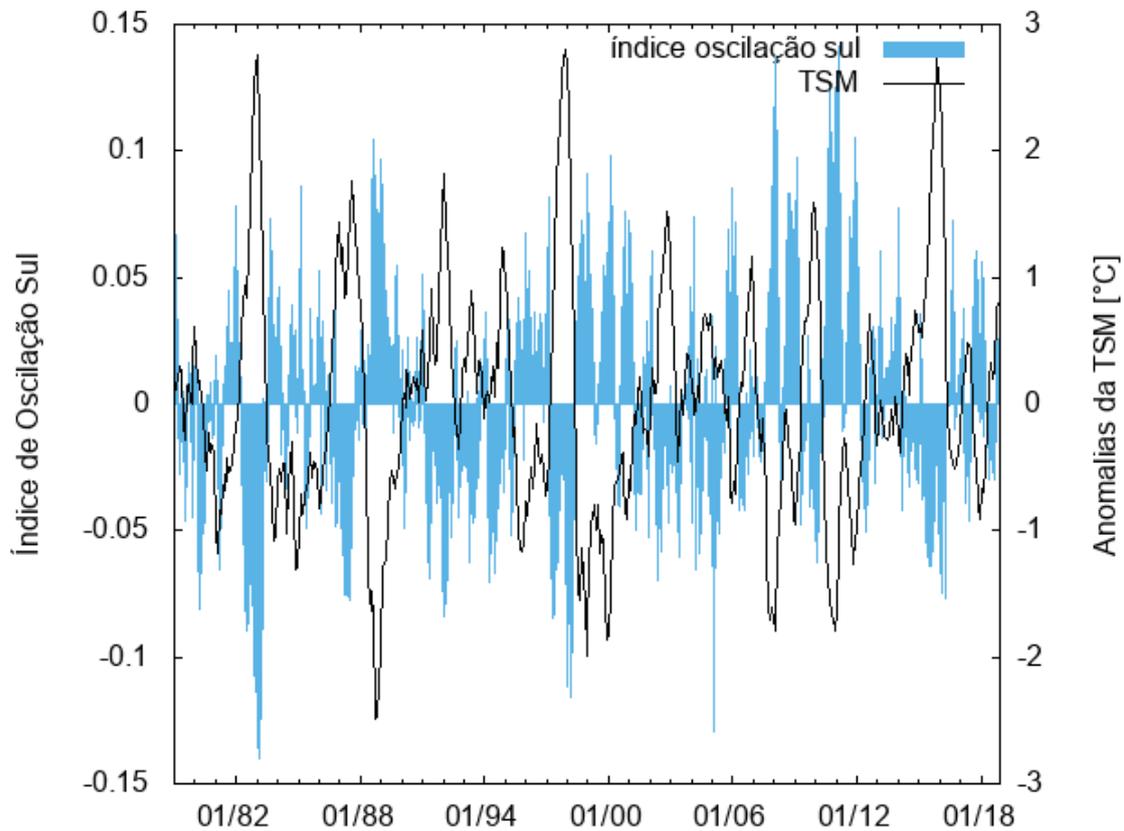


Figura A.13: Anomalias da TSM na região do El Niño 3.4 e Índice de Oscilação Sul de 1979 a 2018.

Na Figura A.13 pode-se notar que as anomalias da TSM na região do El Niño estão entre $+3$ e -3 °C, os valores do IOS estão no intervalo de $+0.15$ e -0.15 e que as anomalias da TSM e o IOS são inversamente proporcionais.

A.5 JATO DE BAIXOS NÍVEIS

As médias climatológicas e anomalias da velocidade dos ventos meridionais a 850 hPa (variável utilizada para representar o Jato de Baixos Níveis - JBN) são apresentadas nas Figuras A.14 e A.15, respectivamente.

Pela Figura A.14 é possível observar que as médias climatológicas da velocidade do vento meridional a 850 hPa são maiores nos meses de junho a agosto, o que corresponde ao inverno austral. Enquanto os menores valores da velocidade do vento meridional são correspondentes aos meses de verão.

Esse comportamento apresentado na Figura A.14 é coerente com o estudo de Grimm (2009) que afirma que os eventos de JBN na porção sul são mais presentes nas estações de inverno e primavera, quando associados a outra fonte de umidade, a ASAS, mais intensa e deslocada para o continente, contribuindo para as chuvas desta região.

O histograma apresentado na Figura A.15 mostra que os valores de anomalias do vento meridional a 850 hPa na região do JBN estão entre $+6$ m/s e -7 m/s.

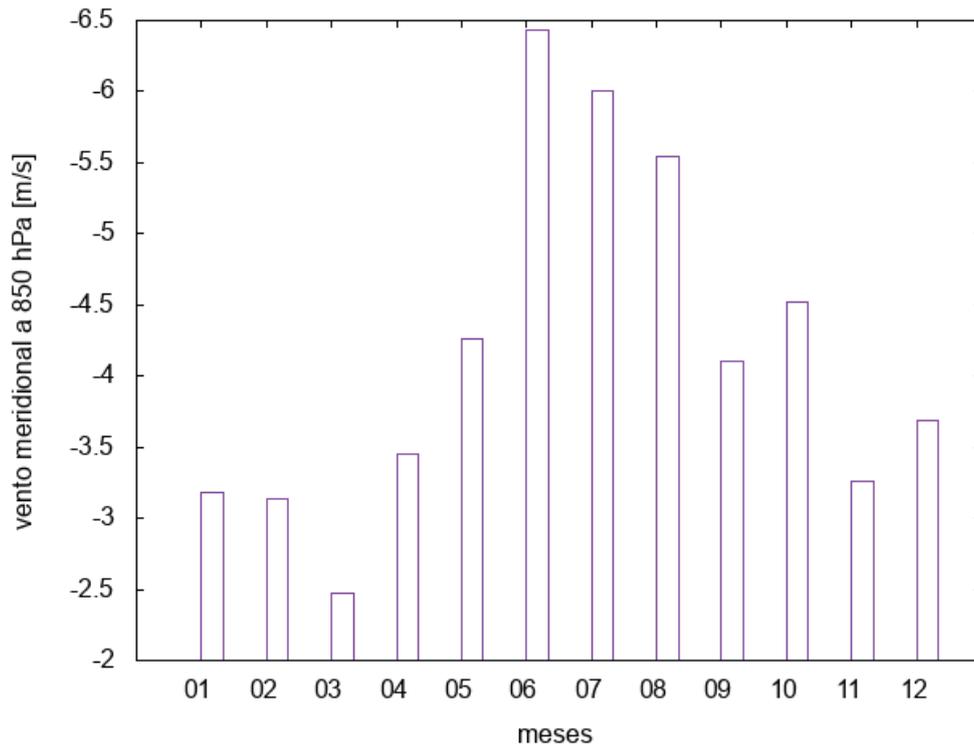


Figura A.14: Médias climatológicas do vento meridional a 850 hPa na região do JBN.

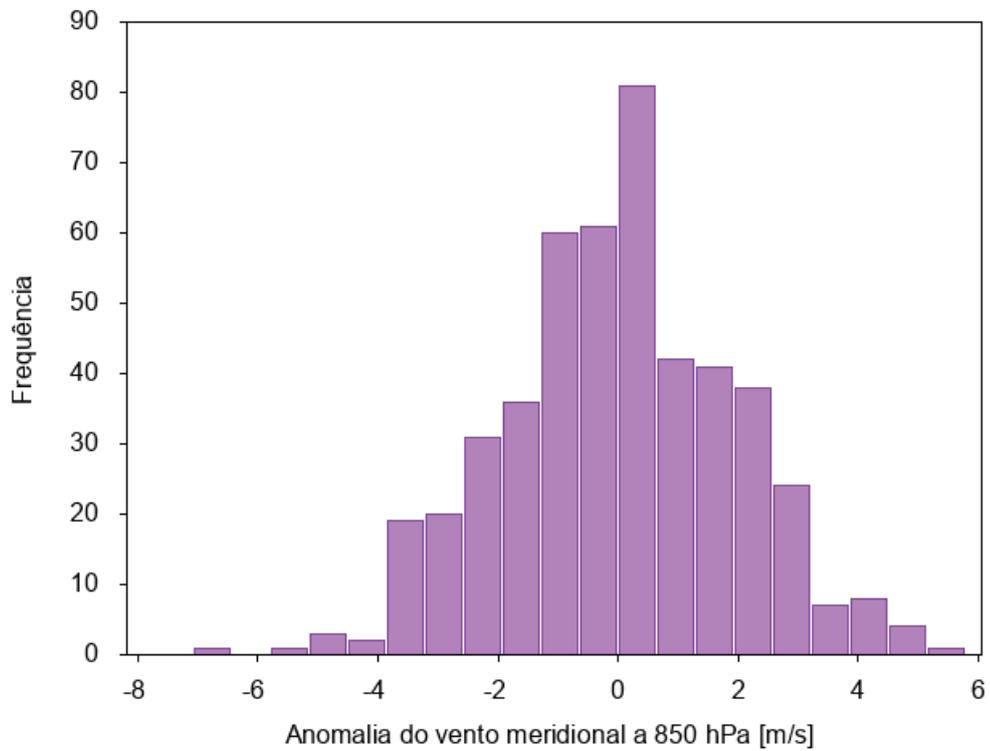


Figura A.15: Anomalias do vento meridional a 850 hPa na região do JBN.

A.6 JATO POLAR

O Jato Polar (JP) foi outro fenômeno atmosférico considerado como relevante no controle das variações no quadro médio climático do Paraná, e para caracterizá-lo foi utilizado a velocidade dos ventos zonais a 250 hPa.

Pela Figura A.16 observa-se que não é possível identificar um padrão para as médias climatológicas do vento zonal a 250 hPa na região do JP, pois há uma sobreposição de informações. No verão o JP é deslocado para o sul, onde foi selecionada a região de atuação do fenômeno, assim verifica-se um pico da velocidade dos ventos zonais nos meses de dezembro e janeiro. Durante o inverno, o JP é mais forte que no verão, no enquanto é deslocado no sentido do equador, assim também é observado um pico durante os meses mais frios do ano. Portanto, são observados dois picos da velocidade dos ventos zonais, um no verão devido à presença do JP na região avaliada, e outro no inverno devido à intensificação do JP durante essa época do ano.

As anomalias do vento zonal a 250 hPa na região do JP são apresentadas na Figura A.17 e pode-se notar as anomalias entre +17 m/s e -17 m/s.

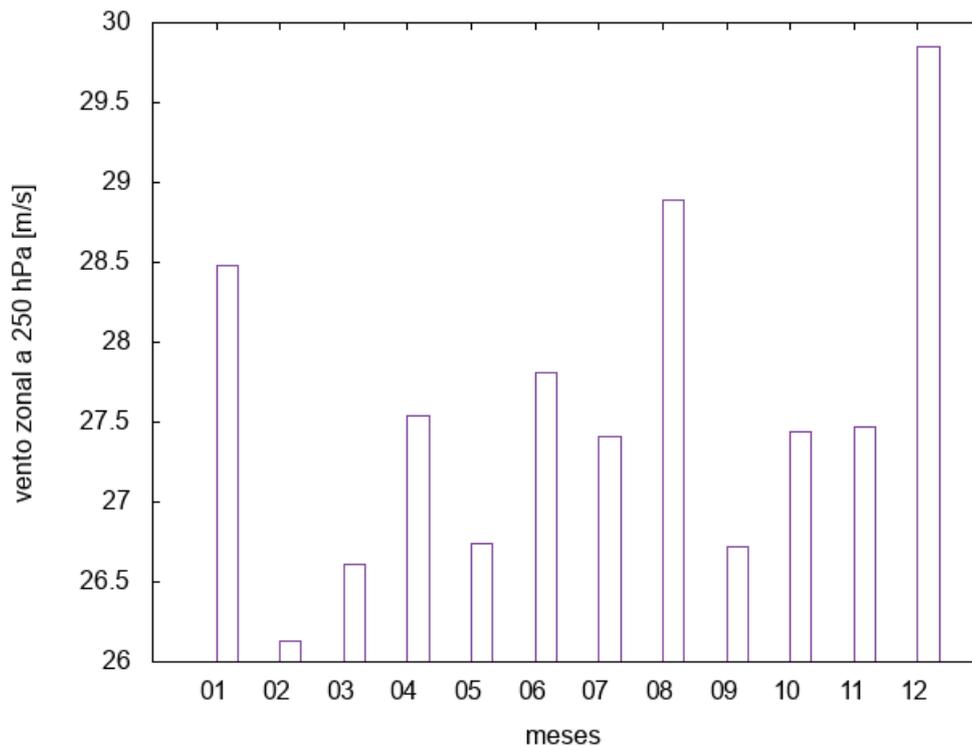


Figura A.16: Médias climatológicas do vento zonal a 250 hPa na região do JP.

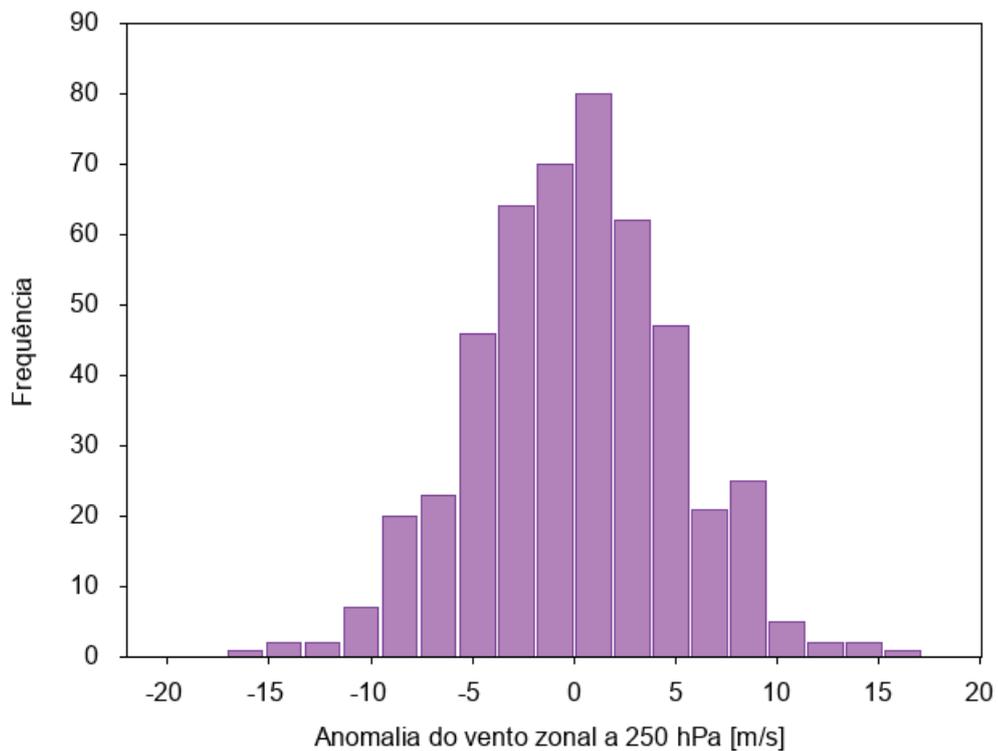


Figura A.17: Anomalias do vento zonal a 250 hPa na região do JP.

A.7 JATO SUBTROPICAL

O Jato Subtropical (JS) foi representado pelas variações temporais dos ventos zonais a 250 hPa na região de 28 a 33° de latitude e 53 a 60° de longitude. As médias climatológicas e as anomalias dos ventos zonais a 250 hPa são apresentados nas Figuras A.18 e A.19, respectivamente.

Pela Figura A.18 pode-se observar que os ventos zonais intensificam-se durante o inverno e primavera (ou seja, nos meses de junho a outubro) e enfraquecem durante o verão (nos meses de dezembro a março). Esse comportamento é justificado pelo maior contraste de temperatura da superfície durante o inverno que produzem ventos mais rápidos em ar superior.

O histograma apresentado na Figura A.19 mostra que as anomalias dos ventos zonais a 250 hPa estão entre +17 e -15 m/s.

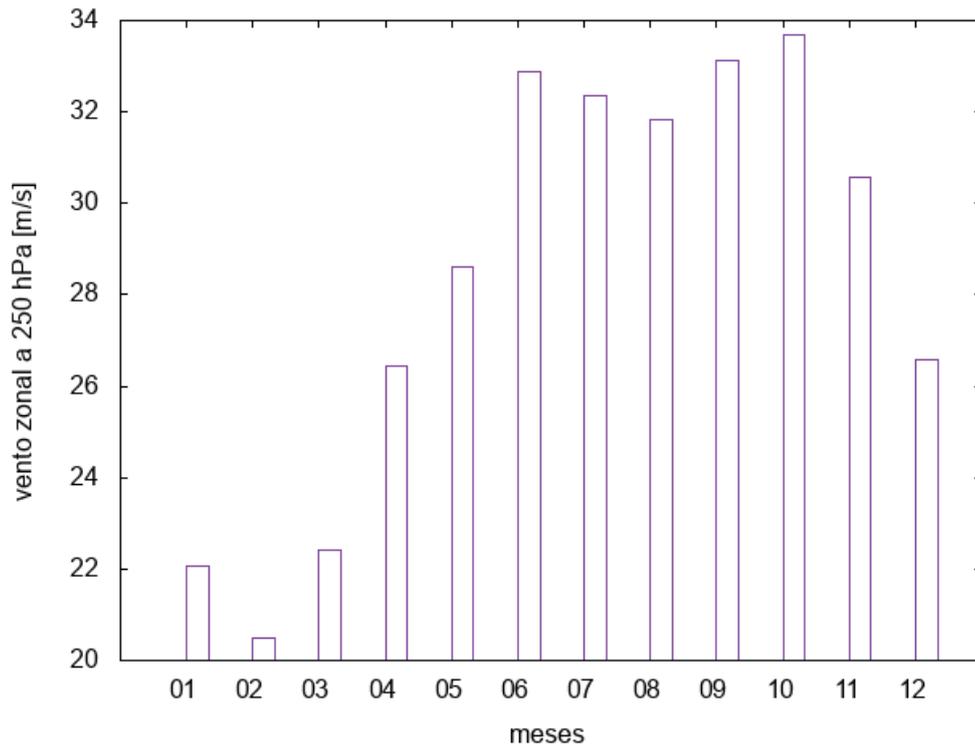


Figura A.18: Médias climatológicas do vento zonal a 250 hPa na região do JS.

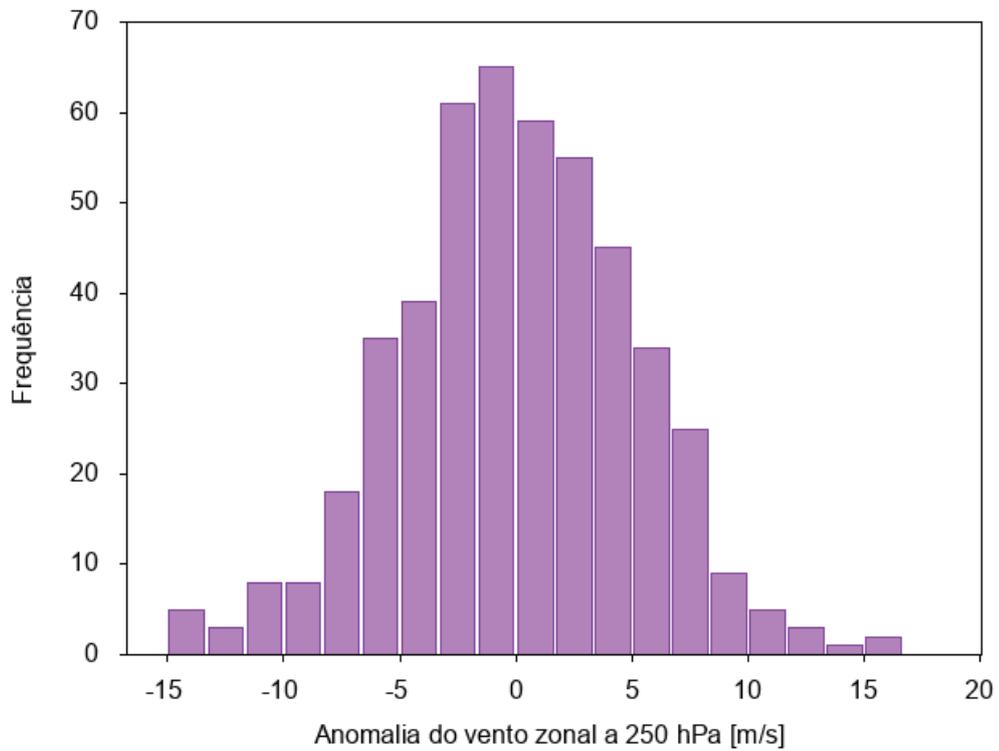


Figura A.19: Anomalias do vento zonal a 250 hPa na região do JS.

A.8 ZONA DE CONVERGÊNCIA DO ATLÂNTICO SUL

Para caracterizar a Zona de Convergência do Atlântico Sul (ZCAS) foram avaliados os comportamentos climáticos da umidade específica a 700 hPa e da velocidade vertical a 500 hPa em duas regiões do Brasil, sendo uma no sudeste brasileiro (área 1) e outra na região norte/centro oeste (área 2).

As médias climatológicas e as anomalias da umidade específica a 700 hPa nas duas áreas analisadas da ZCAS são mostradas nas Figuras A.20 e A.21, respectivamente.

Pela Figura A.20 observa-se que umidade específica a 700 hPa é maior nos meses mais quentes do ano (ou seja, de novembro a março). A partir de abril a umidade diminui até chegar em julho que apresenta o menor valor. Os meses de agosto a outubro nota-se um aumento na umidade a 700 hPa. Esse comportamento é observado nas duas áreas selecionadas para identificar a ZCAS. No verão, os fluxos ascendentes de ar provocados pelo aquecimento terrestre transportam a umidade para a atmosfera. Já no inverno, o resfriamento do continente e a redução da evaporação favorecem os fluxos descendentes de ar. Também na Figura A.20 nota-se que a área 2 tem valores de umidade específica a 700 hPa maiores que na área 1, e isso se deve à proximidade com a região amazônica.

As anomalias da umidade a 700 hPa das duas regiões selecionadas para avaliar a ZCAS são apresentadas na Figura A.21. Observa-se que a área 1 apresenta valores de anomalias maiores que a área 2.

Em suma, a Figuras A.20 indica que a umidade específica a 700 hPa é superior na área 2, no entanto, a Figura A.21 mostra que os valores de anomalias são menores.

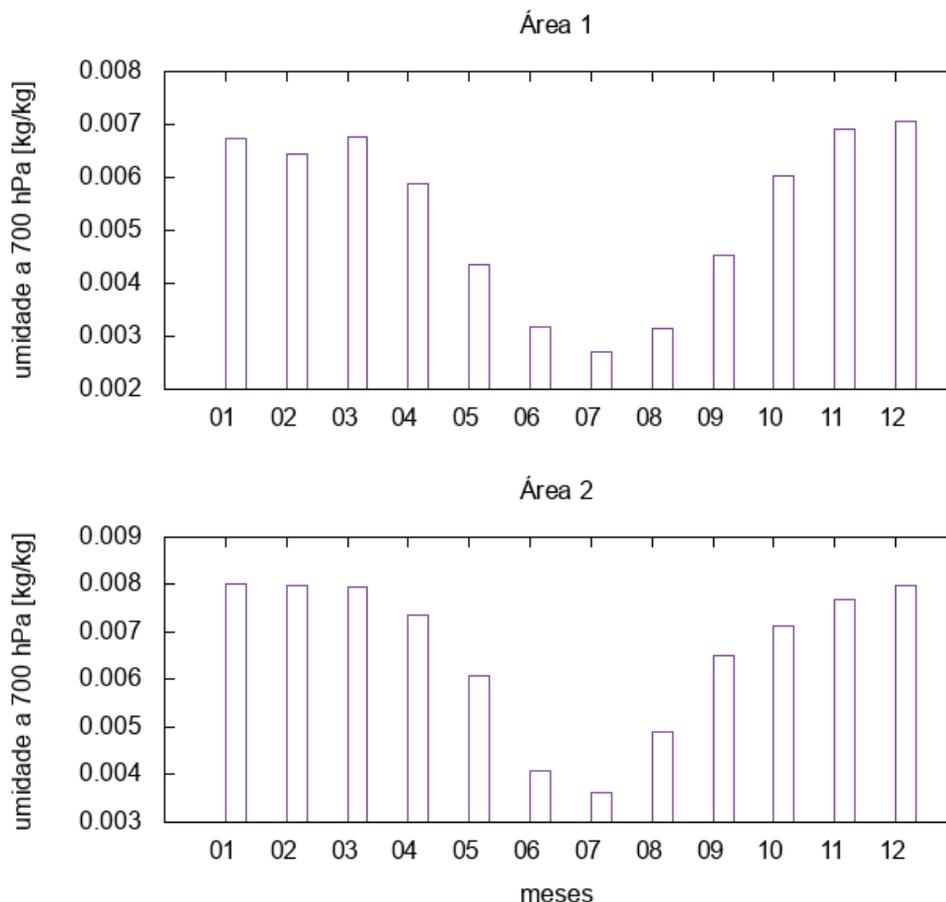


Figura A.20: Médias climatológicas da umidade específica a 700 hPa em duas áreas analisadas da ZCAS.

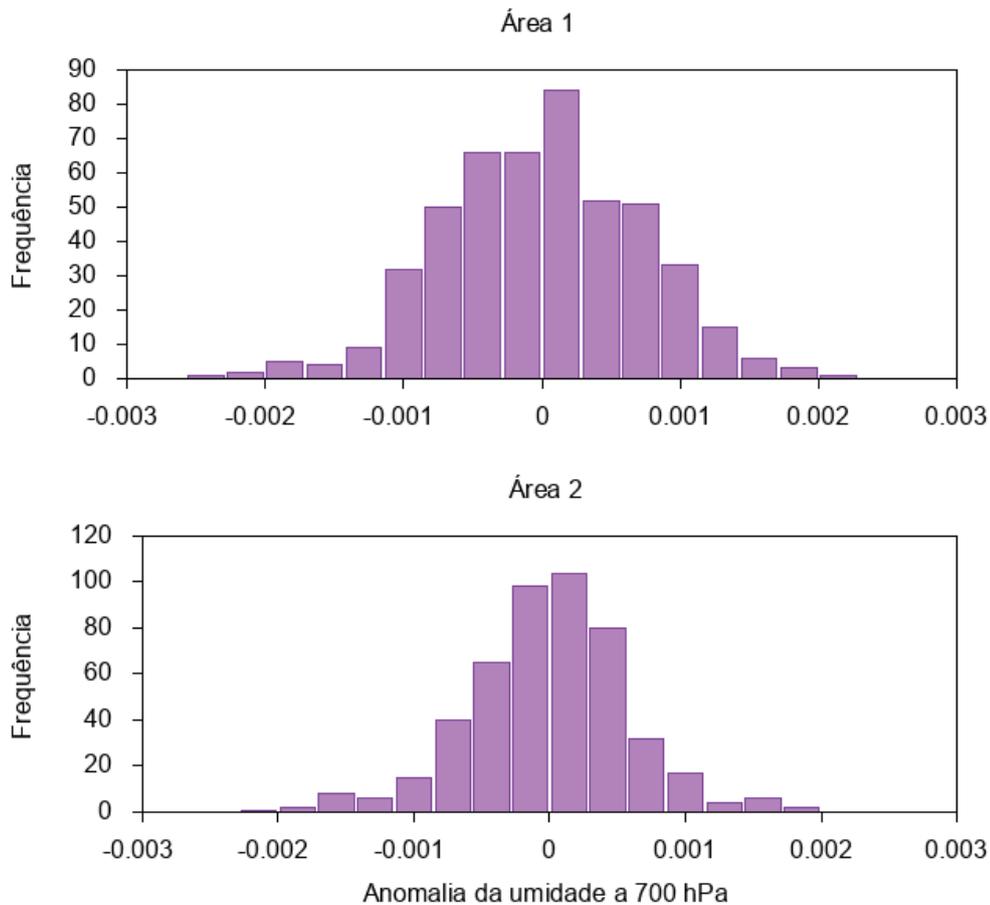


Figura A.21: Anomalias da umidade específica a 700 hPa em duas áreas analisadas da ZCAS.

As médias climatológica e anomalias das velocidades verticais em 500 hPa, expressas em Pa/s, nas duas regiões selecionadas para representar a ZCAS são apresentados nas Figuras A.22 e A.23, respectivamente.

A Figura A.22 mostra que as velocidades verticais a 500 hPa são negativas nos meses mais quentes do ano e positivas nos meses frios. O comportamento das médias climatológicas das velocidades verticais reforçam o padrão indicado na umidade específica a 700 hPa. Assim, no verão, quando há favorecimento dos fluxos ascendentes de ar, a velocidade vertical é negativa, devido ao sistema de coordenadas da variável que está orientado para baixo. Já no inverno, quando há favorecimento dos fluxos descendentes de ar, a velocidade vertical é positiva. Em relação à diferença da área 1 e área 2 apresentadas na Figura A.22, nota-se que os fluxos ascendentes são maiores na área 2, enquanto os fluxos descendentes são maiores na área 2.

Embora a ZCAS seja observada em todos os verões, sendo notada por uma banda de nebulosidade e chuvas com orientação noroeste-sudeste que se estende desde a Amazônia até o Sudeste do Brasil, importantes variações ocorrem na organização espacial e intensidade das anomalias. São essas variações muitas vezes responsáveis pela ocorrência de eventos severos, alagamentos, deslizamentos de terra ou até mesmo pela ausência de chuvas.

As anomalias da velocidade vertical a 500 hPa para as duas regiões selecionadas que representam a ZCAS são apresentadas na Figura A.23. Assim como mostrado para as anomalias da umidade específica a 700 hPa (Figura A.21), as anomalias da velocidade vertical a 500 hPa também são maiores na área 1 (Figura A.23).

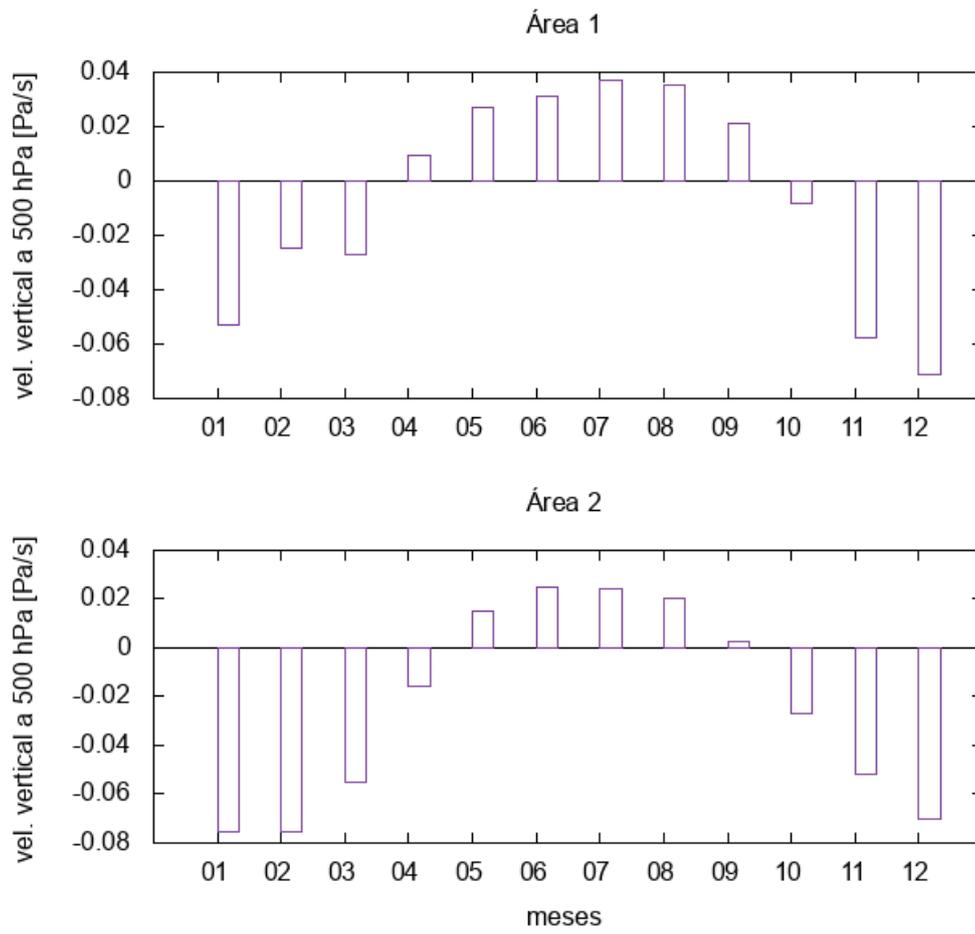


Figura A.22: Médias climatológicas da velocidade vertical a 500 hPa em duas áreas analisadas da ZCAS.

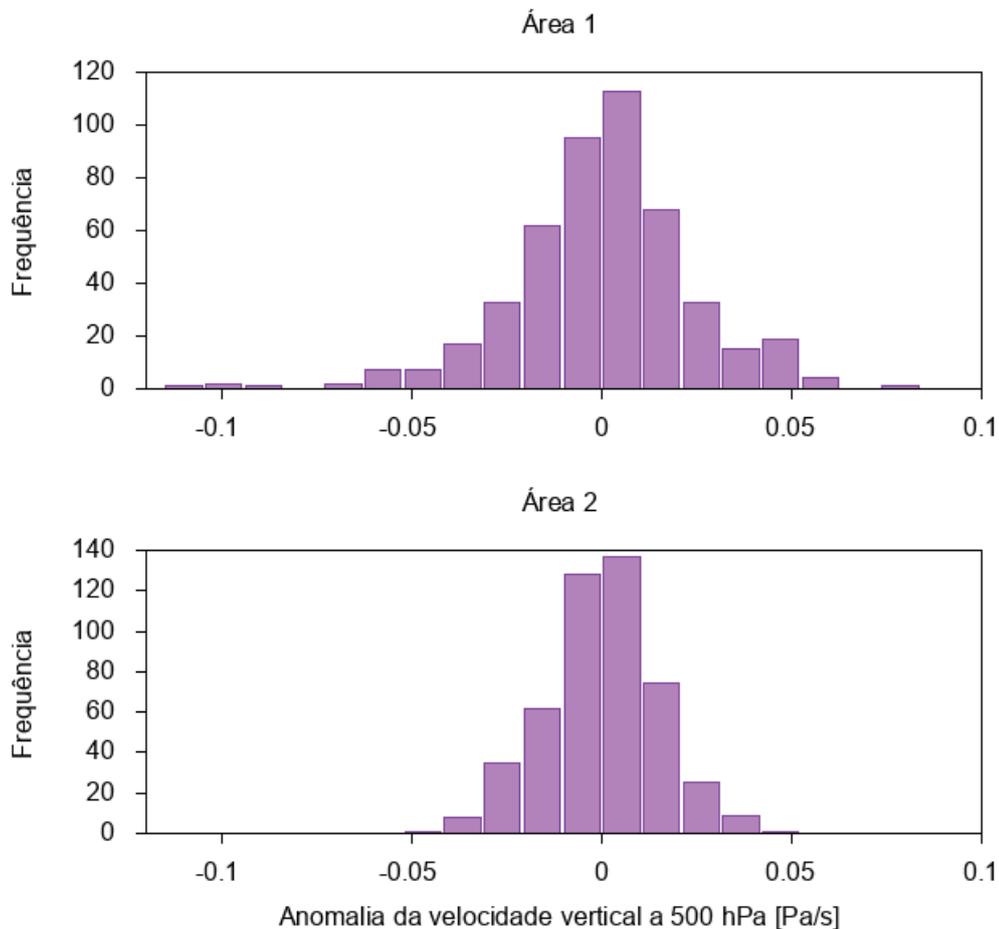


Figura A.23: Anomalias da velocidade vertical a 500 hPa em duas áreas analisadas da ZCAS.

A.9 PARANÁ

Para a região do Paraná (PR), além das variáveis de interesse que são a precipitação e a temperatura, foram selecionadas a velocidade vertical em 500 hPa (em Pa/s), umidade específica a 700 hPa e geopotenciais a 200, 500 e 850 hPa.

As médias climatológicas da umidade específica a 700 hPa no PR são apresentadas na Figura A.24. O comportamento da umidade específica no PR é similar à apresentada na região da ZCAS, com valores maiores nos meses mais quentes do ano (de dezembro a março) e o menor valor no mês de julho, quando a umidade média é de aproximadamente 0,003 kg/kg. No entanto, a região do PR difere-se da região da ZCAS na média máxima de umidade nos meses mais quentes, enquanto na ZCAS os valores máximos são 0,007 e 0,008 kg/kg nas áreas 1 e 2, respectivamente, no PR a umidade média máxima foi 0,0065 kg/kg.

As anomalias da umidade específica a 700 hPa no PR são apresentadas na Figura A.25, em que observa-se pelo histograma que as classes estão entre -0,002 e +0,002 kg/kg.

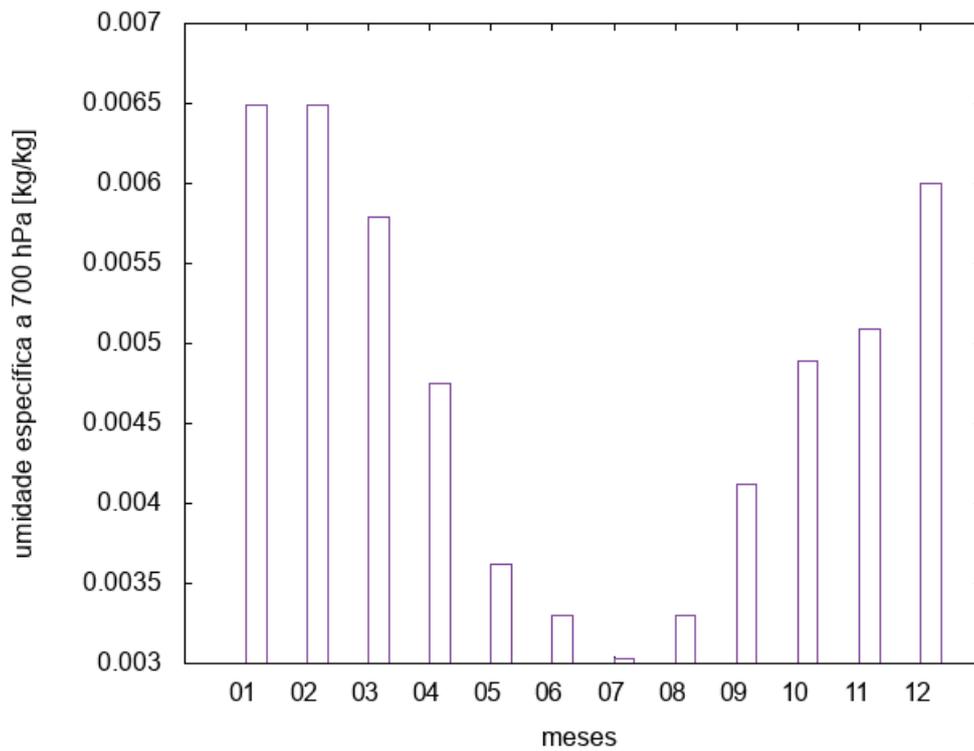


Figura A.24: Médias climatológicas da umidade específica a 700 hPa na região do PR.

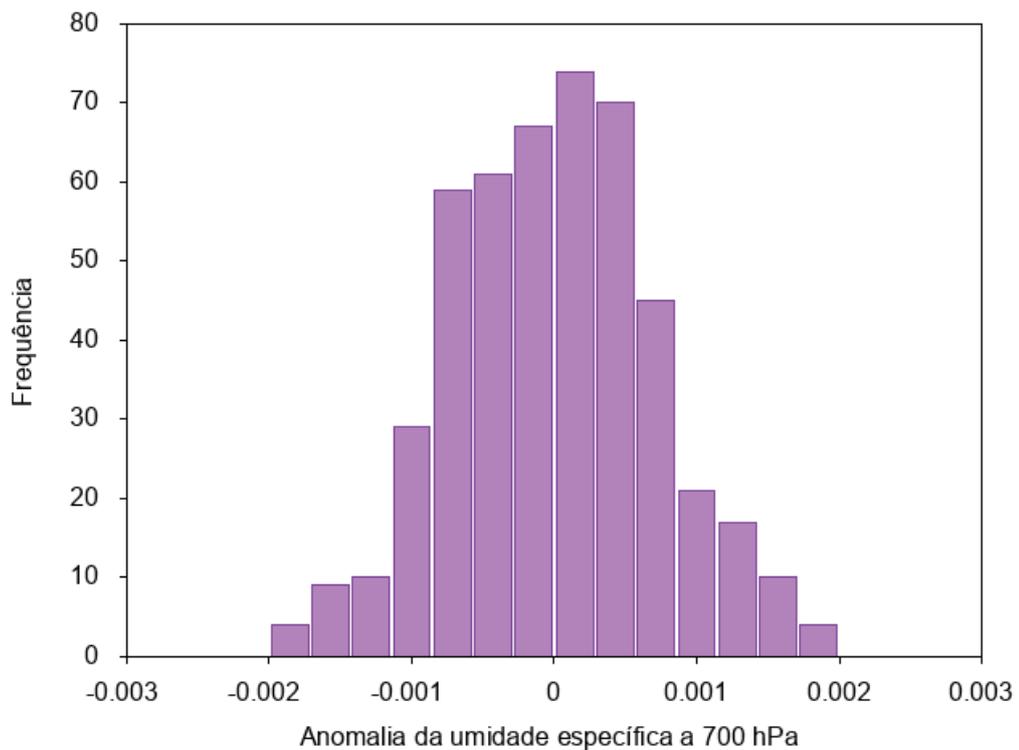


Figura A.25: Anomalias da umidade específica a 700 hPa na região do PR.

Outra variável avaliada no Paraná foi a velocidade vertical a 500 hPa e suas médias climatológicas são mostradas na Figura A.26. Nota-se que no Paraná prevalecem as velocidades

verticais médias mensais negativas durante todos os meses do ano. Portanto, os fluxos ascendentes de ar são predominantes nessa região. O mês de setembro é que apresenta a menor média mensal (com velocidade vertical de aproximadamente $-0,045$ Pa/s), enquanto o mês de abril é o período que possui maior valor da velocidade vertical média (aproximadamente $-0,005$ Pa/s).

As anomalias da velocidade vertical a 500 hPa são apresentadas na Figura A.27. Pode-se observar que as anomalias negativas são maiores que as positivas, pois as classes do histograma estão compreendidas entre $-0,12$ e $+0,8$ Pa/s.

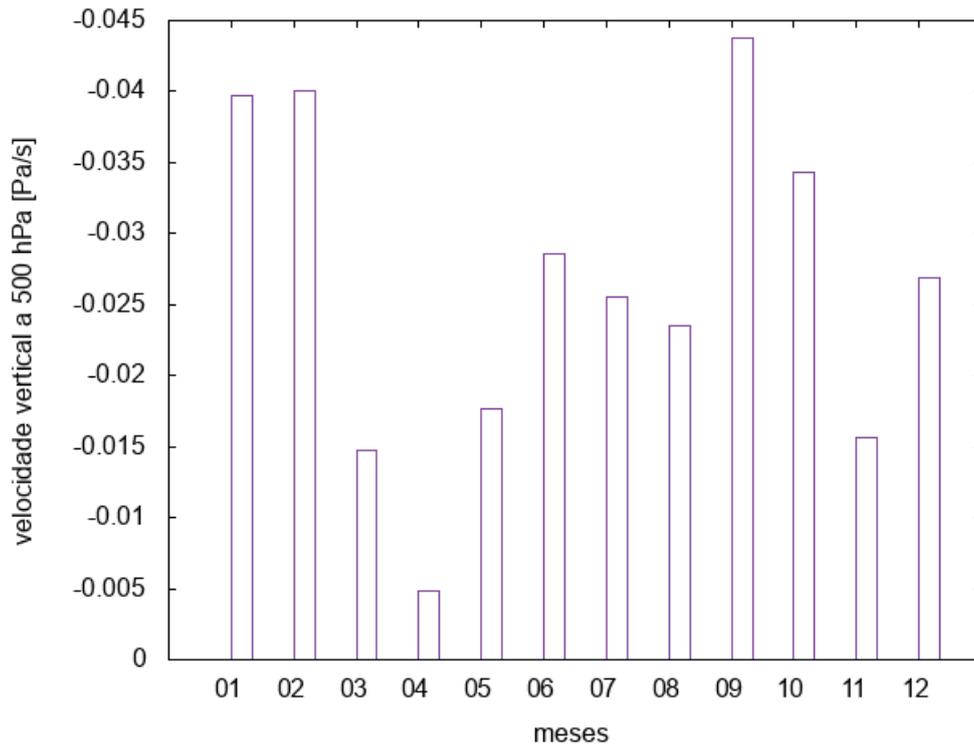


Figura A.26: Médias climatológicas da velocidade vertical a 500 hPa na região do PR.

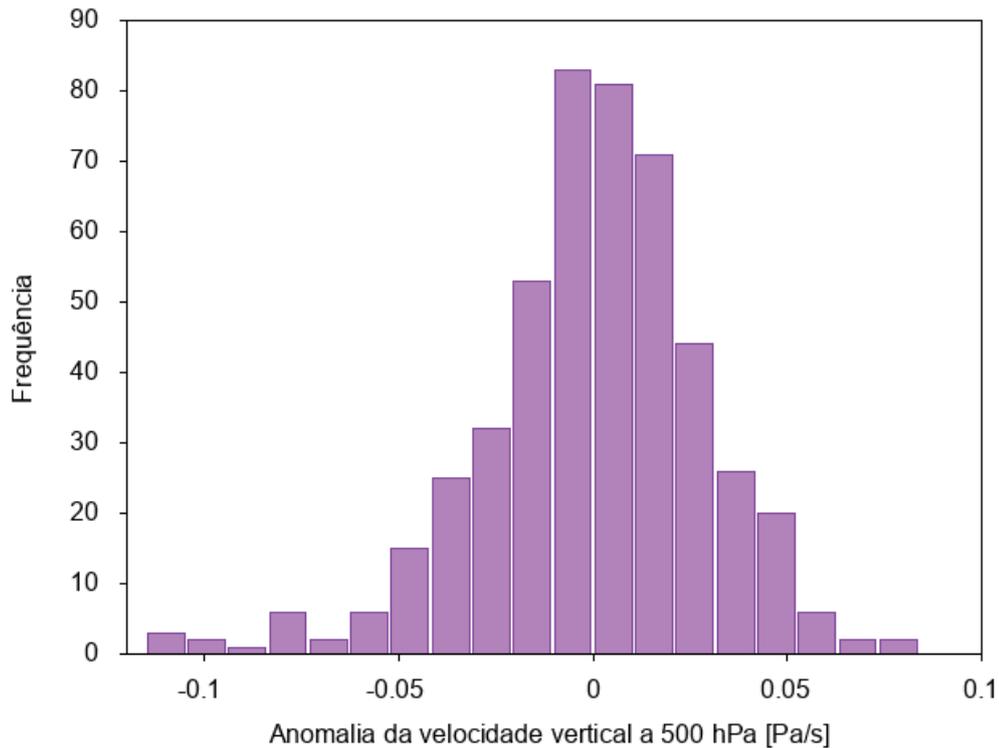


Figura A.27: Anomalias da velocidade vertical a 500 hPa na região do PR.

Também foram avaliados os geopotenciais a 200, 500 e 850 hPa no Paraná. Os gráficos das médias climatológicas e das anomalias são apresentados nas Figuras A.28 e A.29.

Pela Figura A.28 observa-se que as médias climatológicas do geopotencial a 850 hPa (camada baixa da troposfera) são maiores nos meses mais frios, enquanto os geopotenciais a 500 e 200 hPa (camadas médias e altas da troposfera) são menores nos meses correspondentes ao inverno. Os padrões mensais dos geopotenciais a 850, 500 e 200 hPa mostram que durante o inverno as camadas são comprimidas devido às temperaturas menores. Já no verão, as temperaturas mais elevadas fazem com que as camadas da atmosfera sejam expandidas.

Em relação às anomalias apresentadas pela Figura A.29, percebe-se que no nível de 200 hPa o geopotencial varia $-1.100 \text{ m}^2/\text{s}^2$ a $+1.100 \text{ m}^2/\text{s}^2$, nos níveis de 500 hPa e 850 hPa as variações dos geopotenciais estão entre -450 e $450 \text{ m}^2/\text{s}^2$. Embora os valores absolutos das anomalias de geopotencial a 200 hPa sejam maiores que nos níveis de 500 e 850 hPa, a relação entre a variação e a média é similar nas três camadas analisadas.

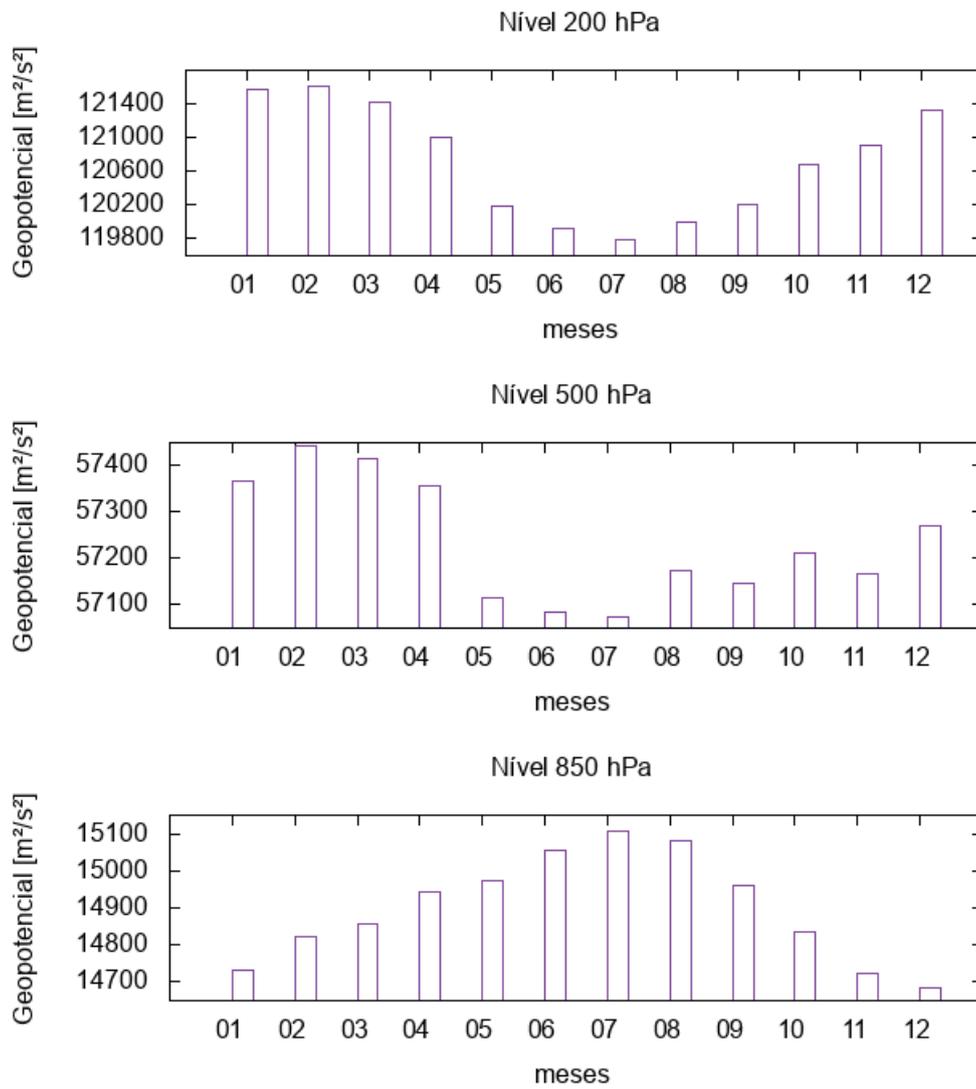


Figura A.28: Médias climatológicas de geopotencial a 200, 500 e 850 hPa na região do PR.

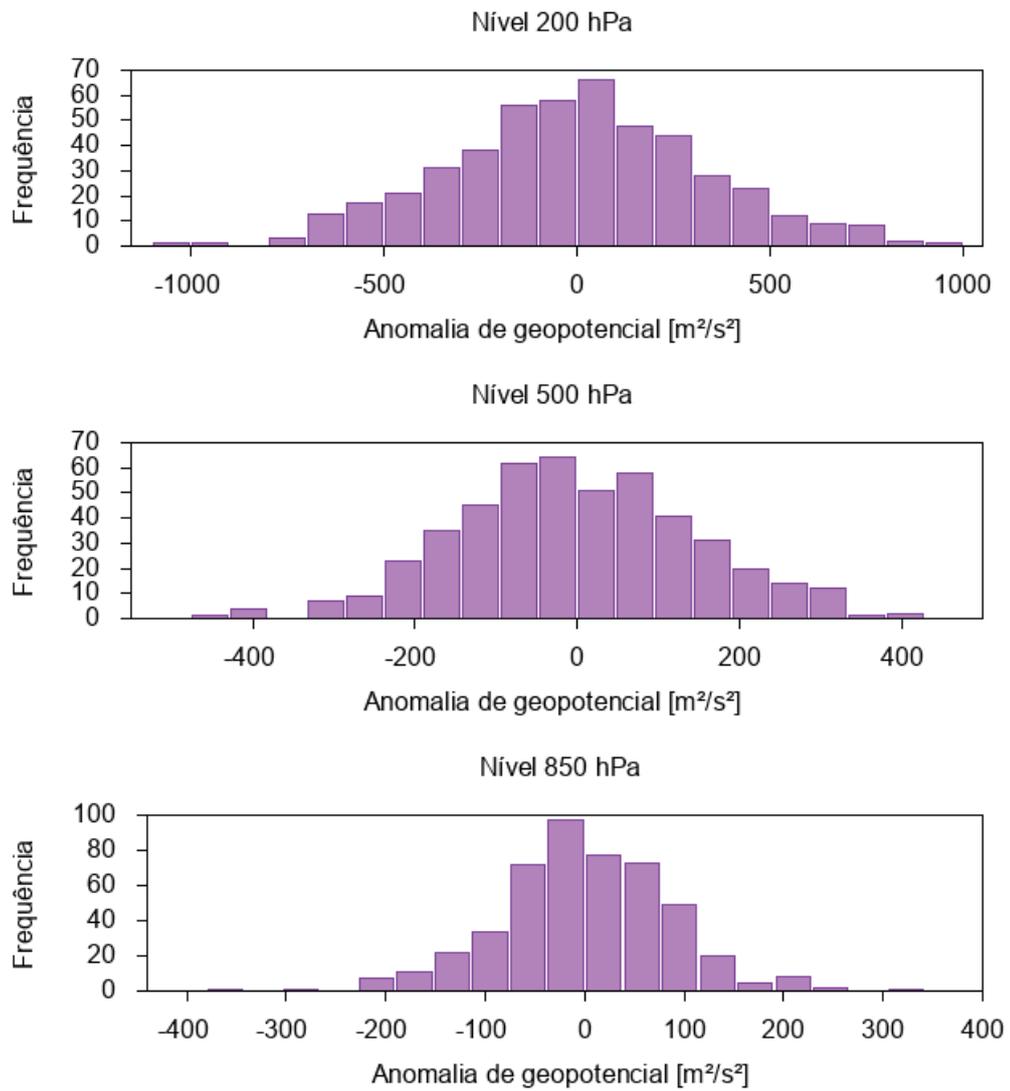


Figura A.29: Histograma das anomalias de geopotenciais a 200, 500 e 850 hPa na região do PR.

APÊNDICE B – RESULTADOS DO MÉTODO *STEPWISE*

B.1 REDE NEURAL DE CONTRA-PROPAGAÇÃO

Nas Figuras B.1, B.2, B.3, B.4 e B.5 são apresentados os resultados dos cinco ciclos do método *stepwise* para a temperatura mensal.

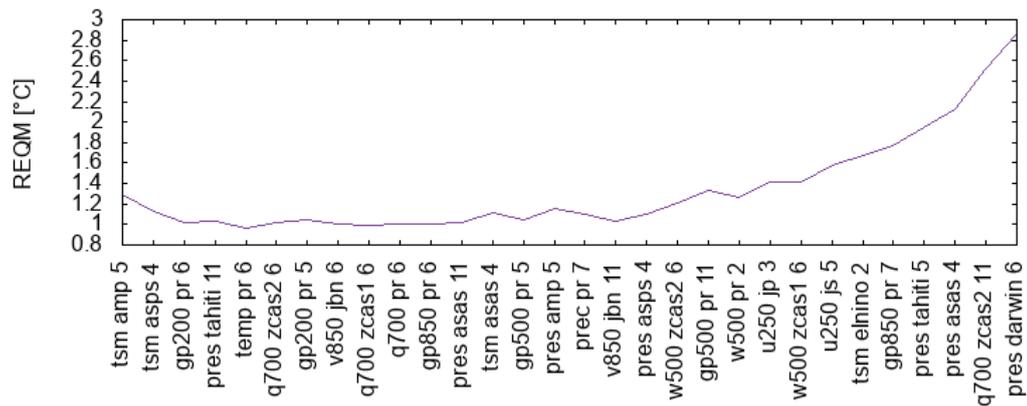


Figura B.1: Resultado do primeiro ciclo do método *stepwise* para a temperatura mensal: 12 variáveis selecionadas.

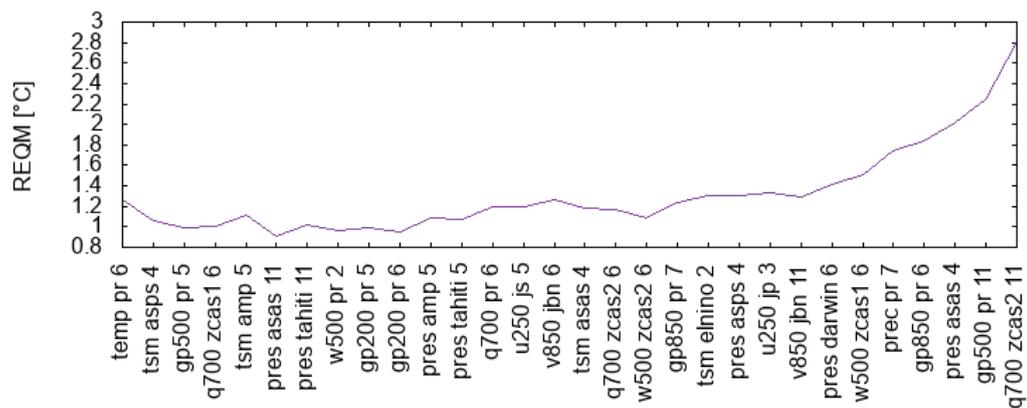


Figura B.2: Resultado do segundo ciclo do método *stepwise* para a temperatura mensal: 10 variáveis selecionadas.

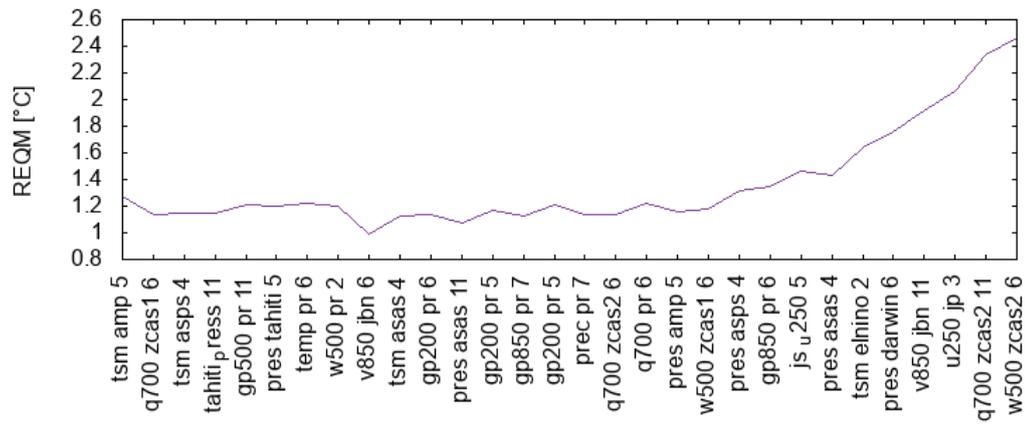


Figura B.3: Resultado do terceiro ciclo do método *stepwise* para a temperatura mensal: 15 variáveis selecionadas.

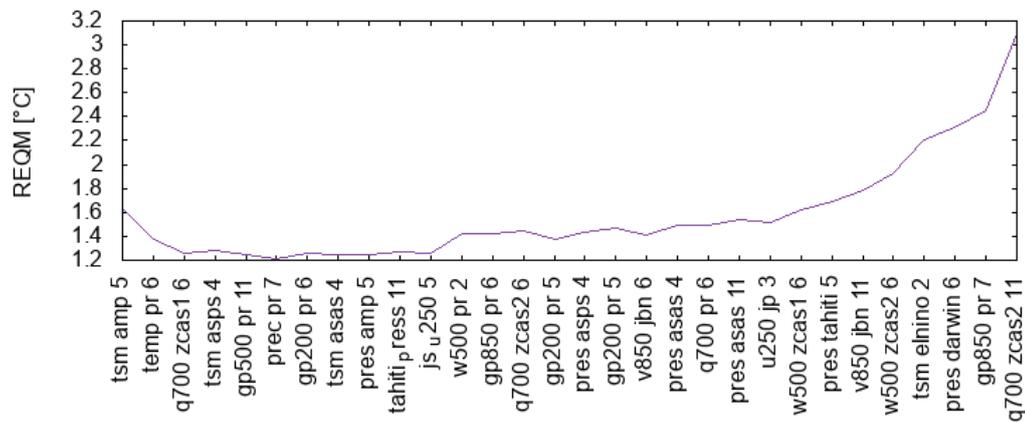


Figura B.4: Resultado do quarto ciclo do método *stepwise* para a temperatura mensal: 11 variáveis selecionadas.

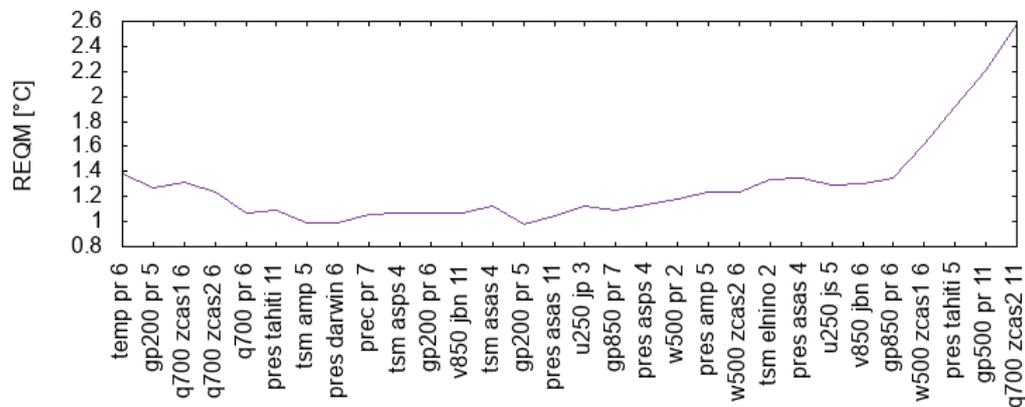


Figura B.5: Resultado do quinto ciclo do método *stepwise* para a temperatura mensal: 14 variáveis selecionadas.

Nas Figuras B.6, B.7, B.8, B.9 e B.10 são apresentados os resultados dos cinco ciclos do método *stepwise* para a temperatura sazonal.

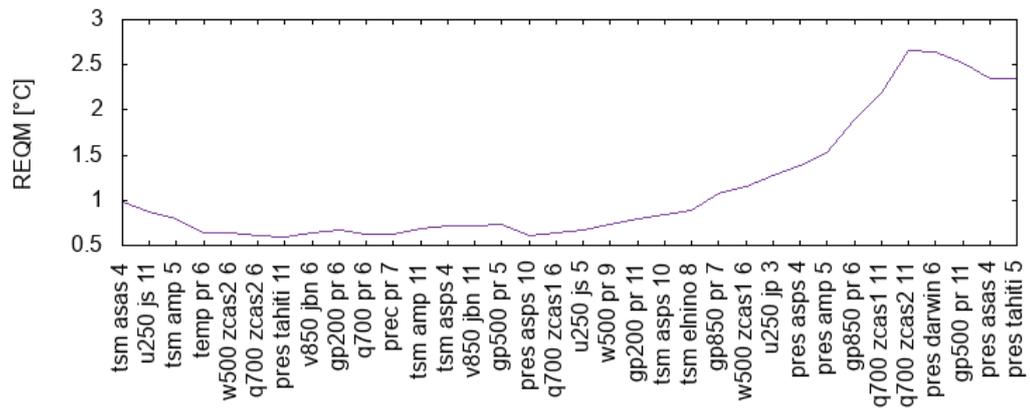


Figura B.6: Resultado do primeiro ciclo do método *stepwise* para a temperatura sazonal: 16 variáveis selecionadas.

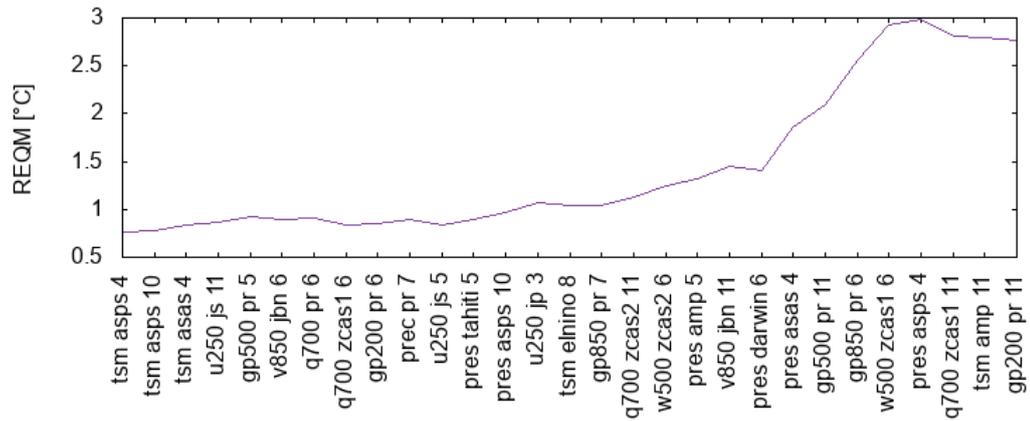


Figura B.7: Resultado do segundo ciclo do método *stepwise* para a temperatura sazonal: 11 variáveis selecionadas.

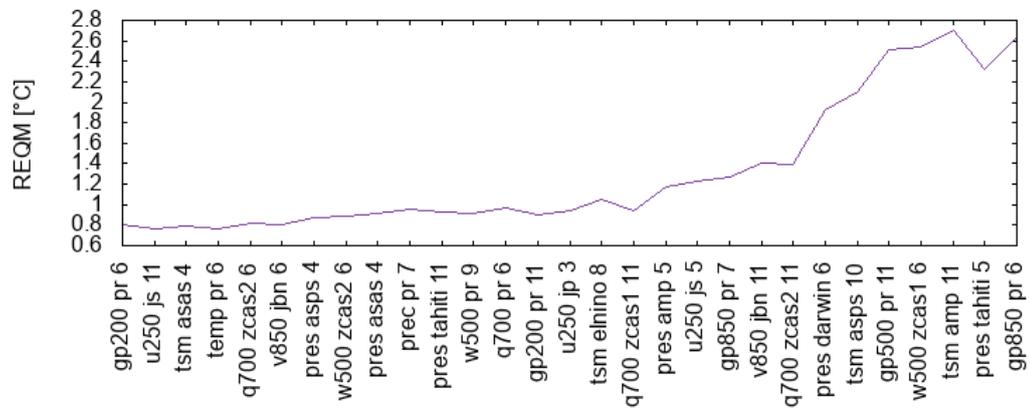


Figura B.8: Resultado do terceiro ciclo do método *stepwise* para a temperatura sazonal: 11 variáveis selecionadas.

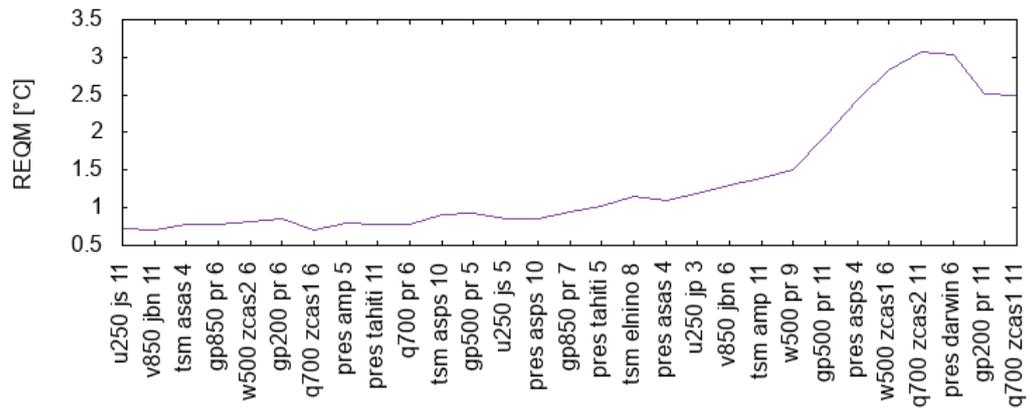


Figura B.9: Resultado do quarto ciclo do método *stepwise* para a temperatura sazonal: 15 variáveis selecionadas.

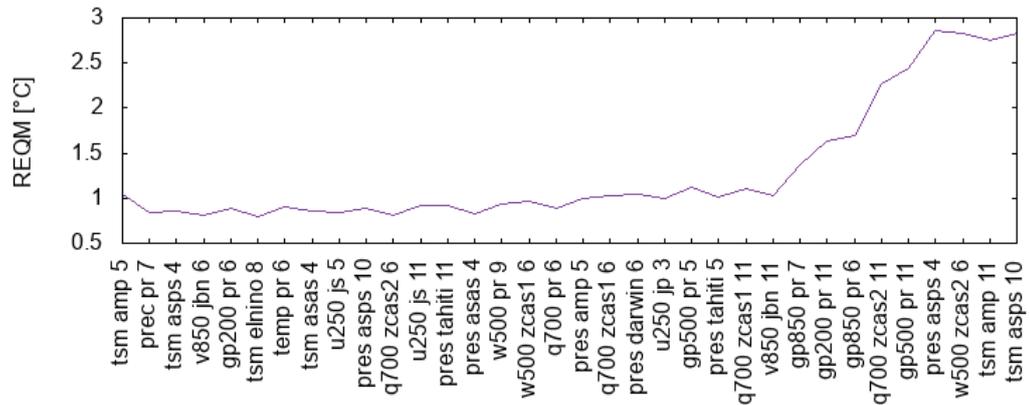


Figura B.10: Resultado do quinto ciclo do método *stepwise* para a temperatura sazonal: 15 variáveis selecionadas.

Nas Figuras B.11, B.12, B.13, B.14 e B.15 são apresentados os resultados dos cinco ciclos do método *stepwise* para a precipitação mensal.

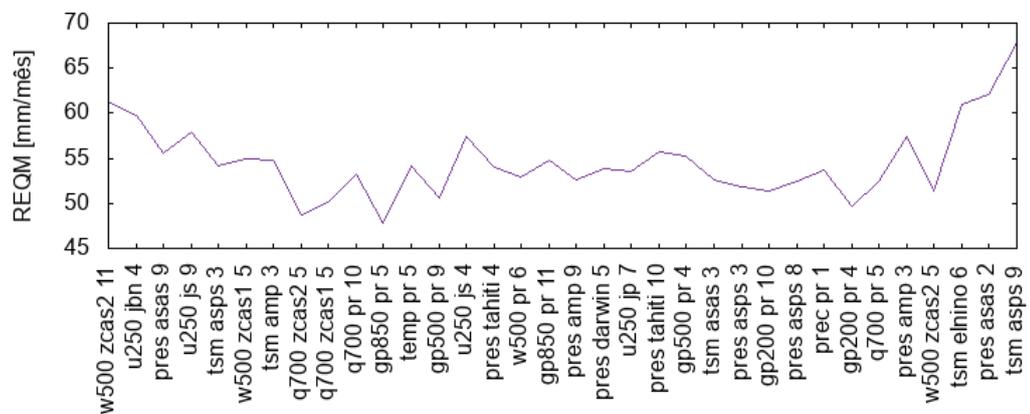


Figura B.11: Resultado do primeiro ciclo do método *stepwise* para a precipitação mensal: 17 variáveis selecionadas.

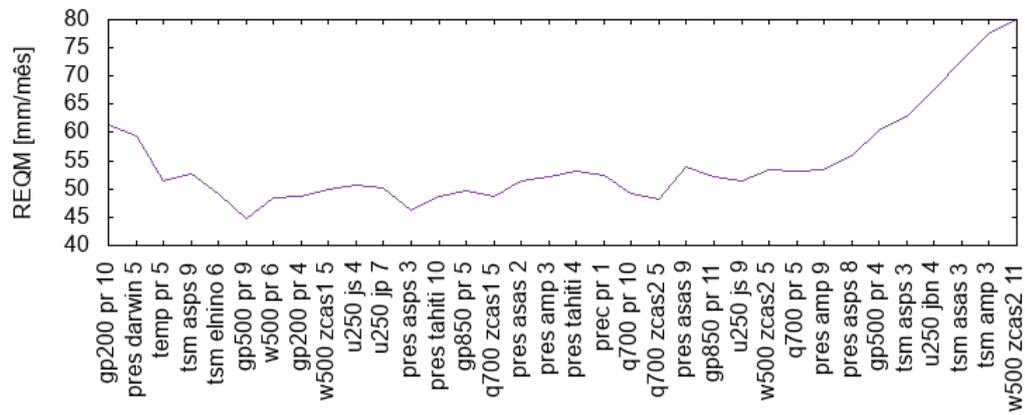


Figura B.12: Resultado do segundo ciclo do método *stepwise* para a precipitação mensal: 17 variáveis selecionadas.

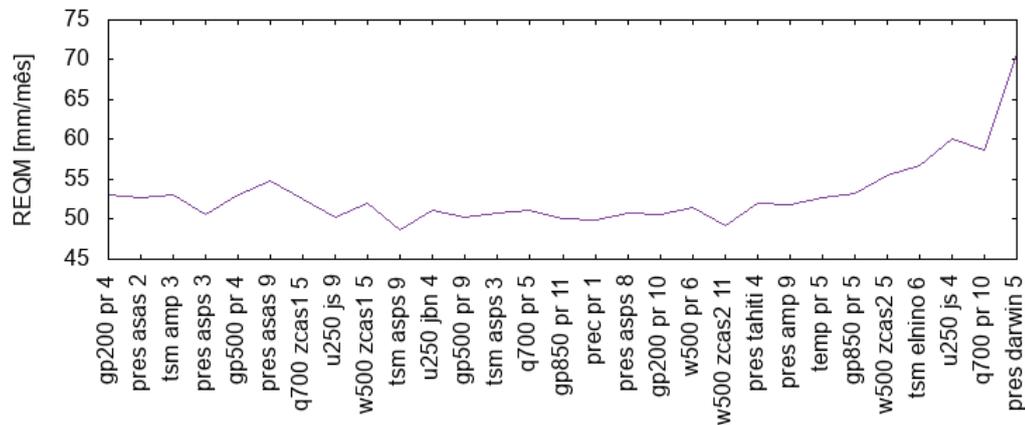


Figura B.13: Resultado do terceiro ciclo do método *stepwise* para a precipitação mensal: 17 variáveis selecionadas.

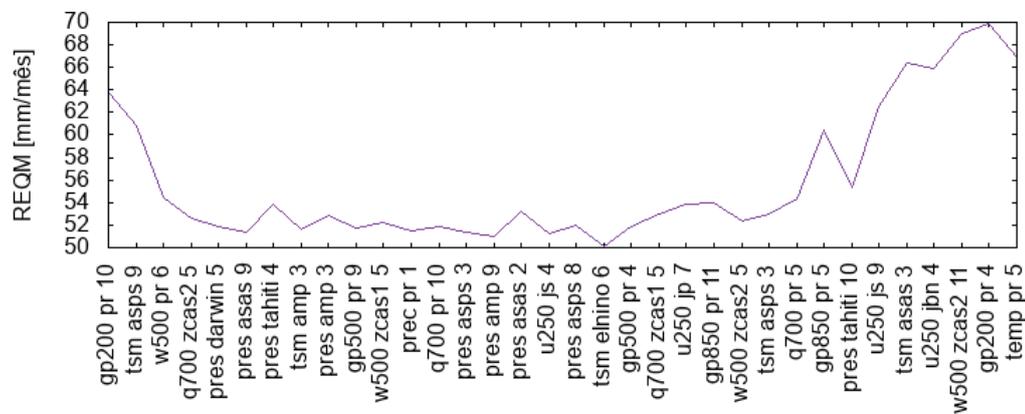


Figura B.14: Resultado do quarto ciclo do método *stepwise* para a precipitação mensal: 17 variáveis selecionadas.

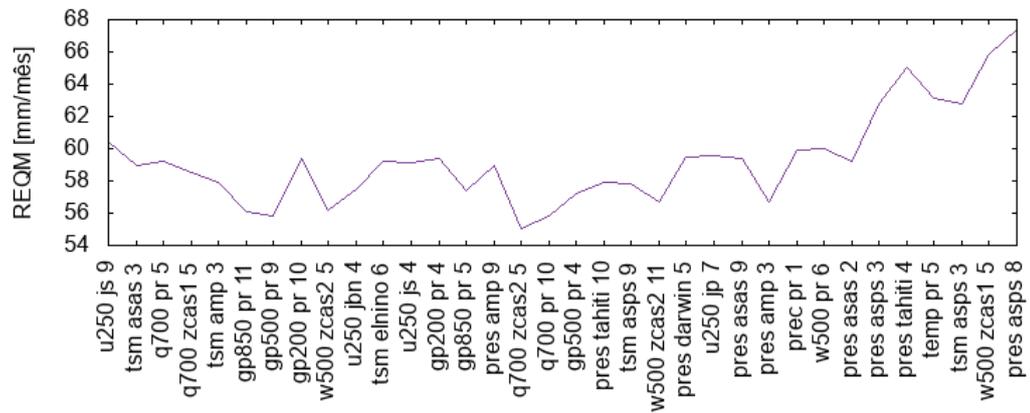


Figura B.15: Resultado do quinto ciclo do método *stepwise* para a precipitação mensal: 17 variáveis selecionadas.

Após a execução dos cinco ciclos do método *stepwise* para a precipitação mensal, foram selecionadas 19 variáveis que serviram de entrada para outros cinco ciclos do método de seleção de variáveis. Os resultados dos novos cinco ciclos do método *stepwise* são mostrados nas Figuras B.16, B.17, B.18, B.19 e B.20.

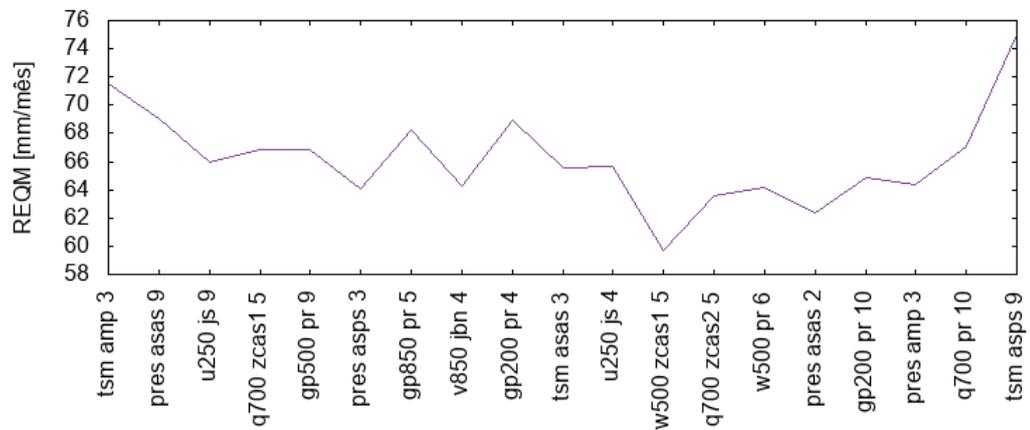


Figura B.16: Resultado do primeiro ciclo da segunda etapa do método *stepwise* para a precipitação mensal: 10 variáveis selecionadas.

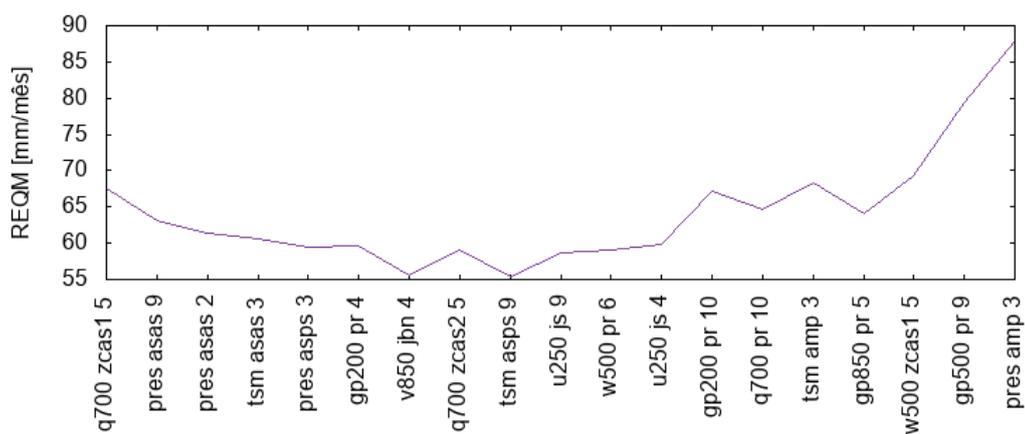


Figura B.17: Resultado do segundo ciclo da segunda etapa do método *stepwise* para a precipitação mensal: 10 variáveis selecionadas.

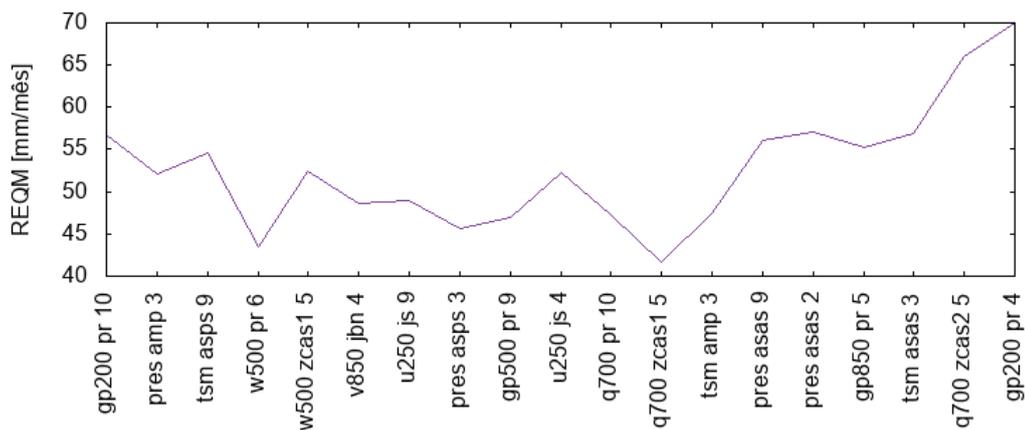


Figura B.18: Resultado do terceiro ciclo da segunda etapa do método *stepwise* para a precipitação mensal: 10 variáveis selecionadas.

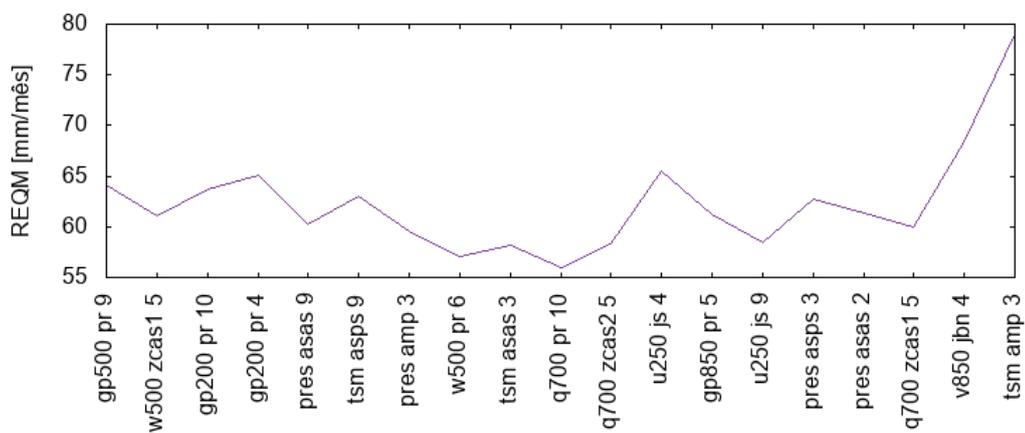


Figura B.19: Resultado do quarto ciclo da segunda etapa do método *stepwise* para a precipitação mensal: 10 variáveis selecionadas.

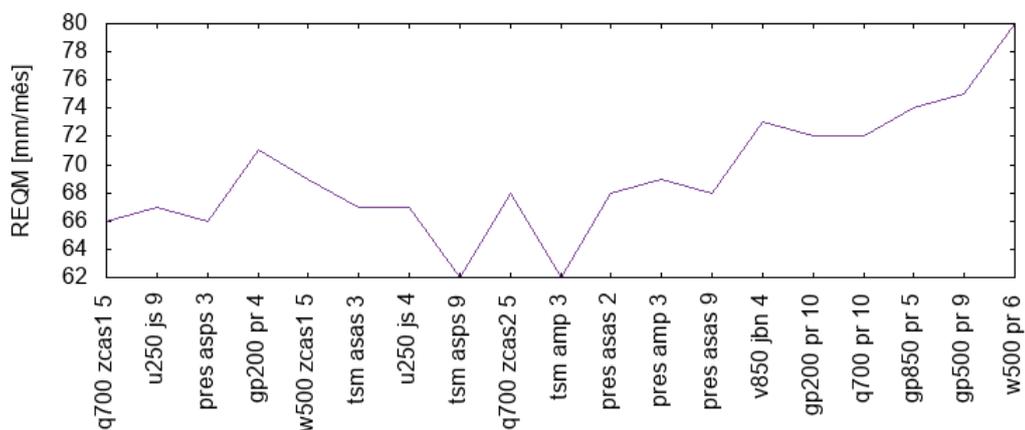


Figura B.20: Resultado do quinto ciclo da segunda etapa do método *stepwise* para a precipitação mensal: 10 variáveis selecionadas.

Os resultados dos cinco ciclos do método *stepwise* para a precipitação sazonal são apresentados nas Figuras B.21, B.22, B.23, B.24 e B.25.

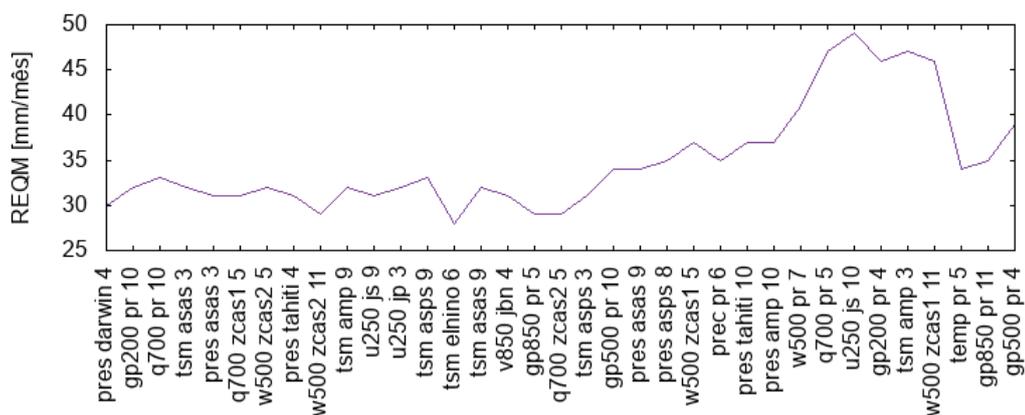


Figura B.21: Resultado do primeiro ciclo do método *stepwise* para a precipitação sazonal: 17 variáveis selecionadas.

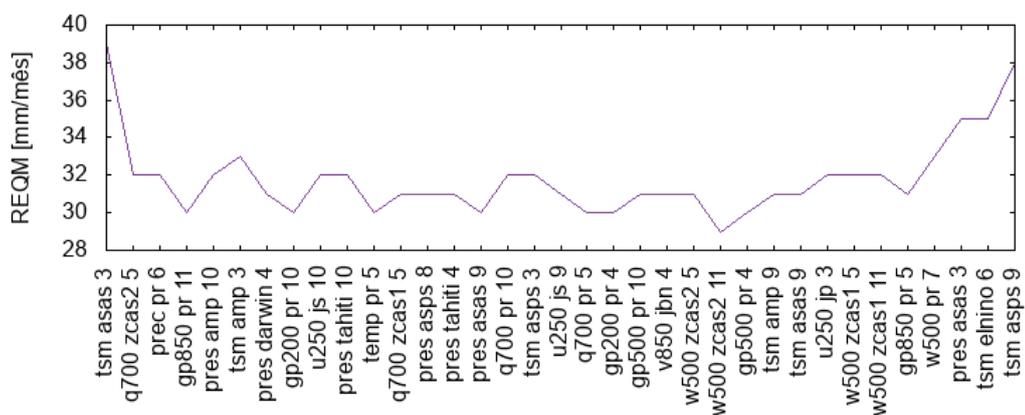


Figura B.22: Resultado do segundo ciclo do método *stepwise* para a precipitação sazonal: 17 variáveis selecionadas.

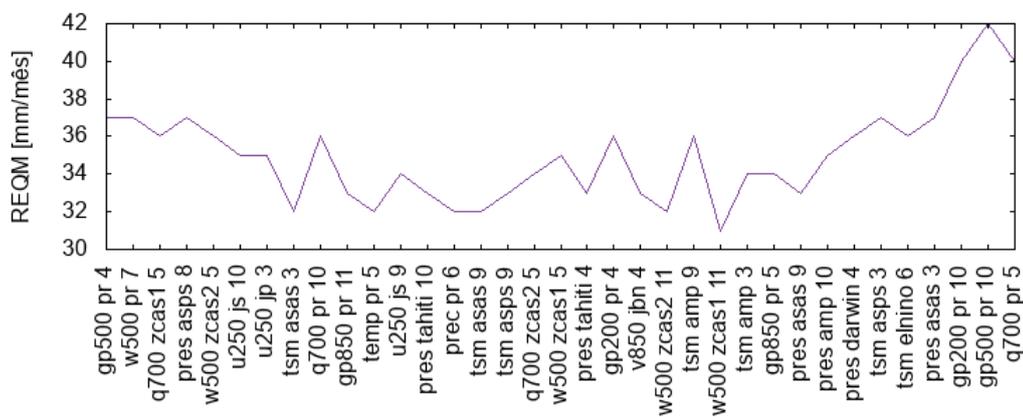


Figura B.23: Resultado do terceiro ciclo do método *stepwise* para a precipitação sazonal: 17 variáveis selecionadas.

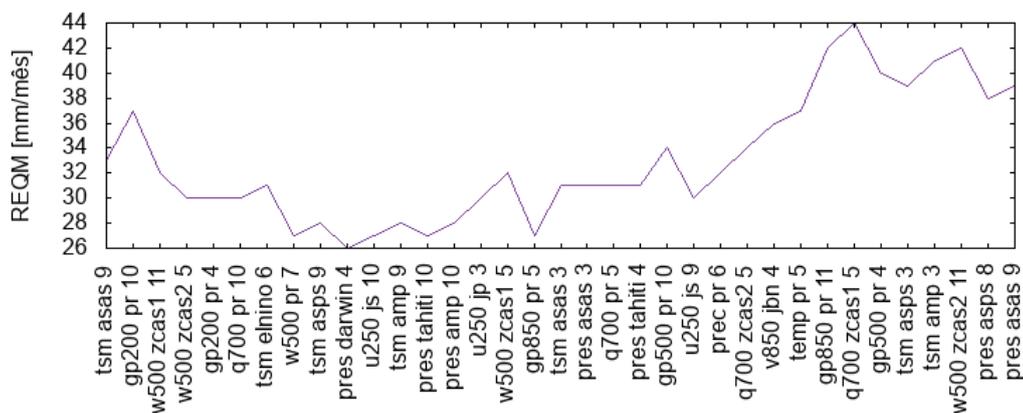


Figura B.24: Resultado do quarto ciclo do método *stepwise* para a precipitação sazonal: 17 variáveis selecionadas.

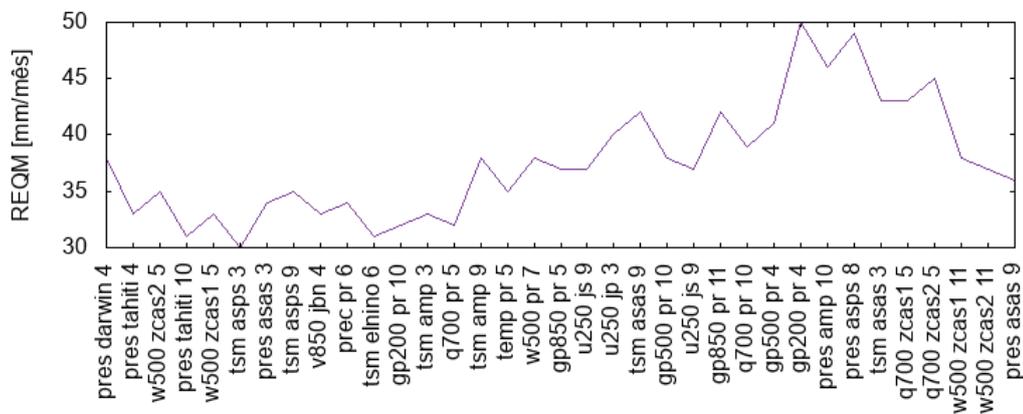


Figura B.25: Resultado do quinto ciclo do método *stepwise* para a precipitação sazonal: 14 variáveis selecionadas.

B.2 REDE NEURAL RBF

Os resultados dos cinco ciclos do método *stepwise* para a temperatura mensal, utilizando a rede neural RBF, são apresentados nas Figuras B.26, B.27, B.28, B.29 e B.30.

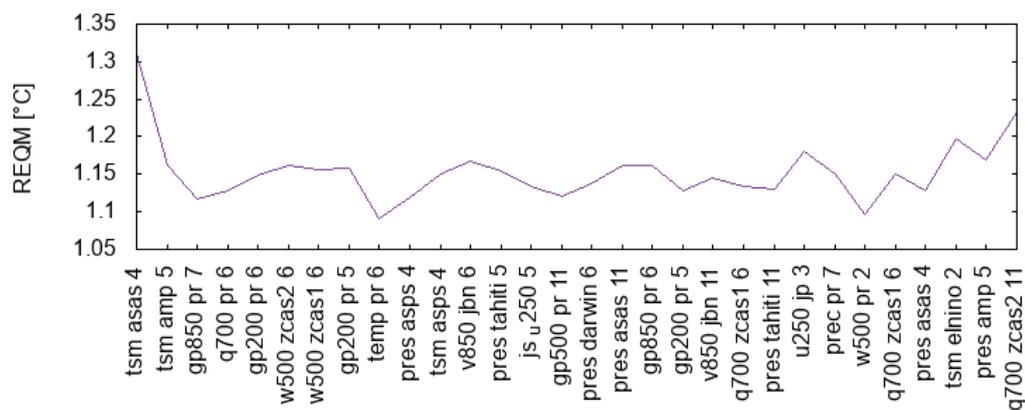


Figura B.26: Resultado do primeiro ciclo do método *stepwise* para a temperatura mensal: 15 variáveis selecionadas.

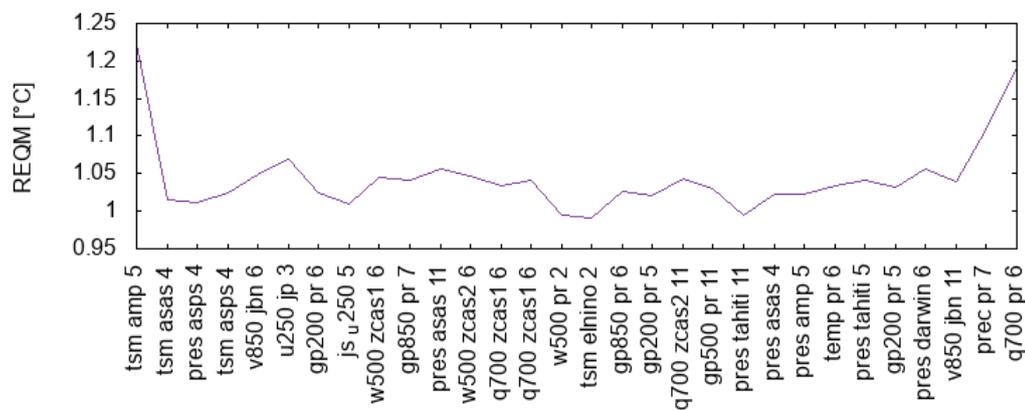


Figura B.27: Resultado do segundo ciclo do método *stepwise* para a temperatura mensal: 15 variáveis selecionadas.

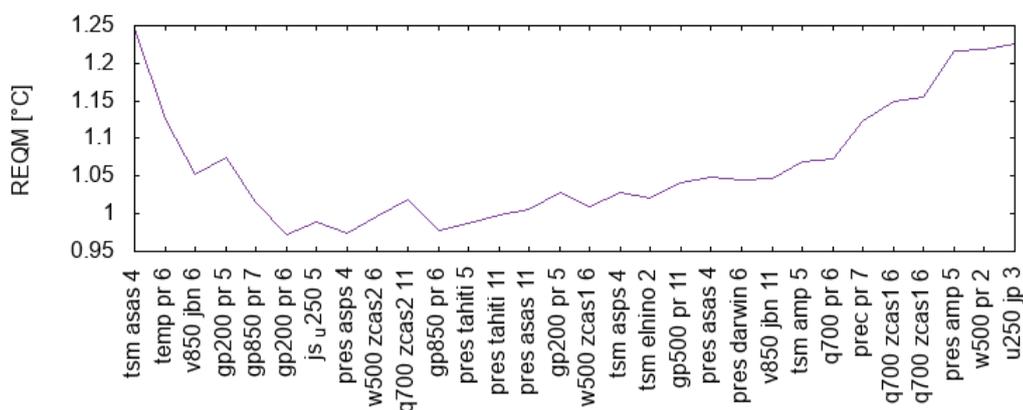


Figura B.28: Resultado do terceiro ciclo do método *stepwise* para a temperatura mensal: 06 variáveis selecionadas.

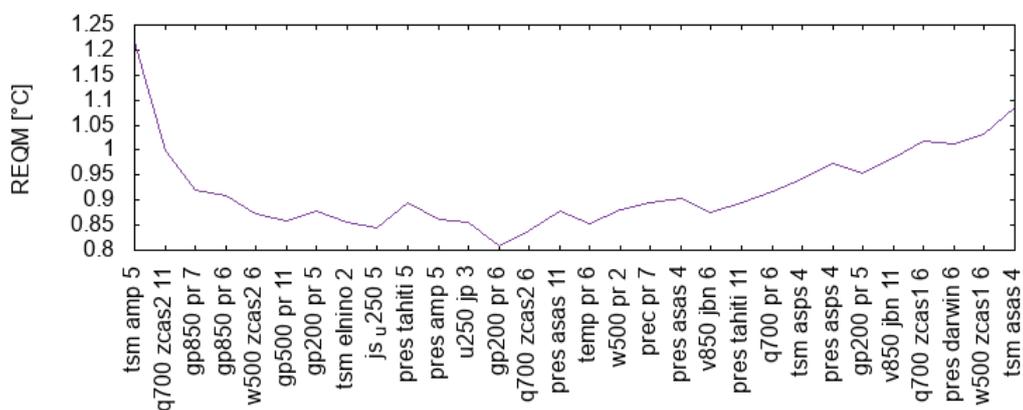


Figura B.29: Resultado do quarto ciclo do método *stepwise* para a temperatura mensal: 13 variáveis selecionadas.

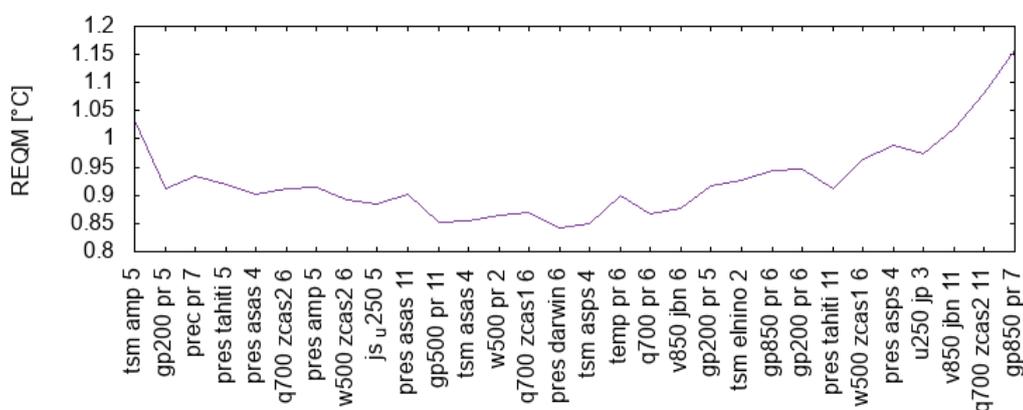


Figura B.30: Resultado do quinto ciclo do método *stepwise* para a temperatura mensal: 15 variáveis selecionadas.

Nas Figuras B.31, B.32, B.33, B.34 e B.35 são apresentados os resultados dos cinco ciclos do método *stepwise* para a temperatura sazonal.

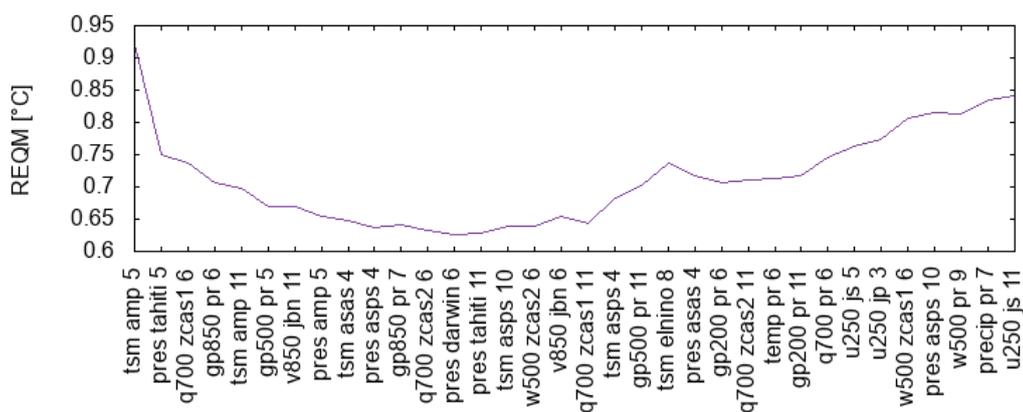


Figura B.31: Resultado do primeiro ciclo do método *stepwise* para a temperatura sazonal: 13 variáveis selecionadas.

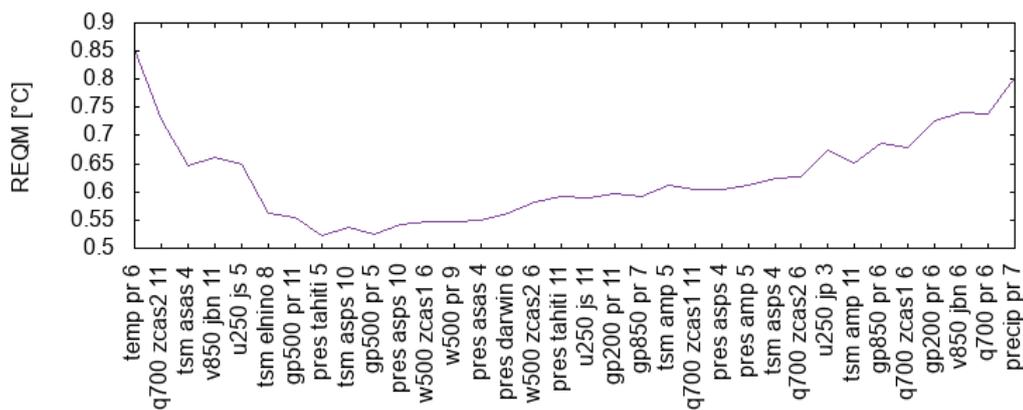


Figura B.32: Resultado do segundo ciclo do método *stepwise* para a temperatura sazonal: 08 variáveis selecionadas.

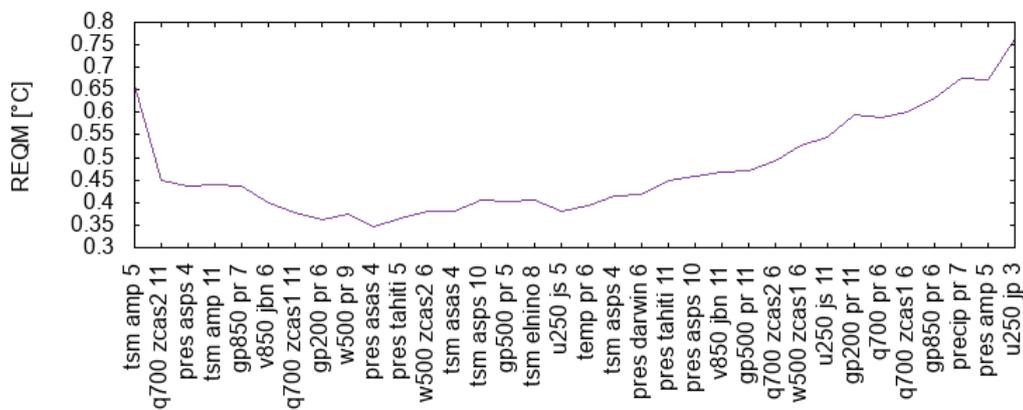


Figura B.33: Resultado do terceiro ciclo do método *stepwise* para a temperatura sazonal: 10 variáveis selecionadas.

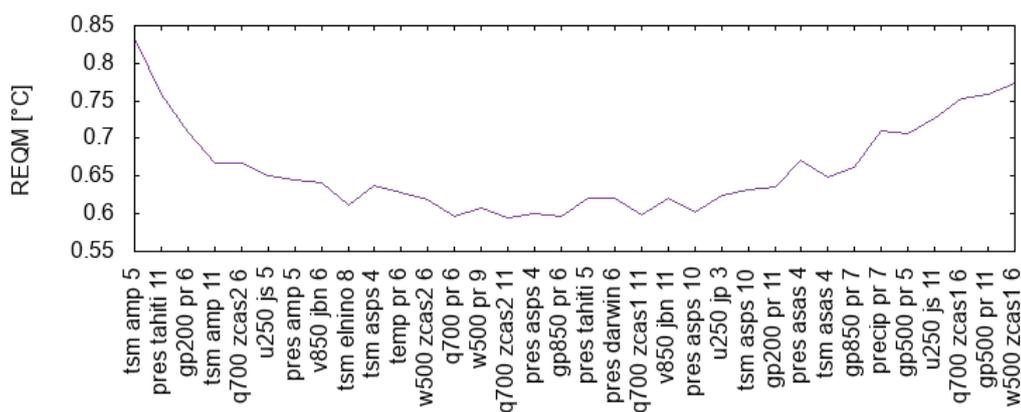


Figura B.34: Resultado do quarto ciclo do método *stepwise* para a temperatura sazonal: 15 variáveis selecionadas.

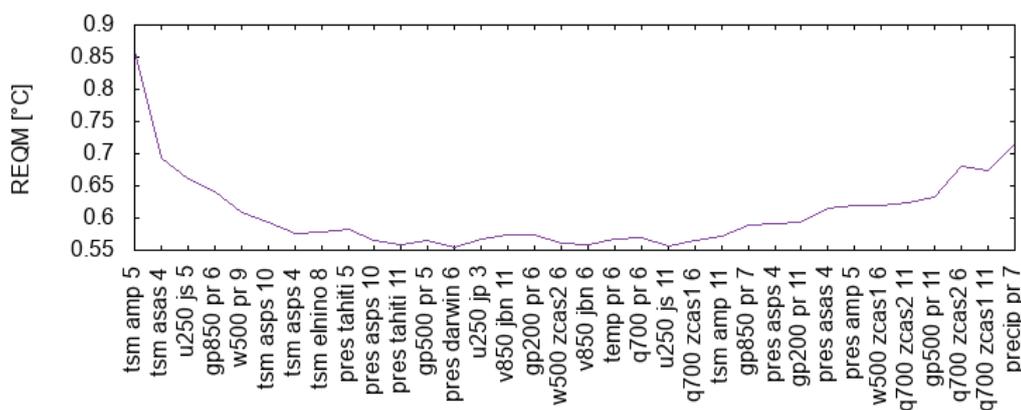


Figura B.35: Resultado do quinto ciclo do método *stepwise* para a temperatura sazonal: 12 variáveis selecionadas.

Para a precipitação mensal, os resultados dos cinco ciclos do método *stepwise* usando a rede RBF são apresentados nas Figuras B.36, B.37, B.38, B.39 e B.40.

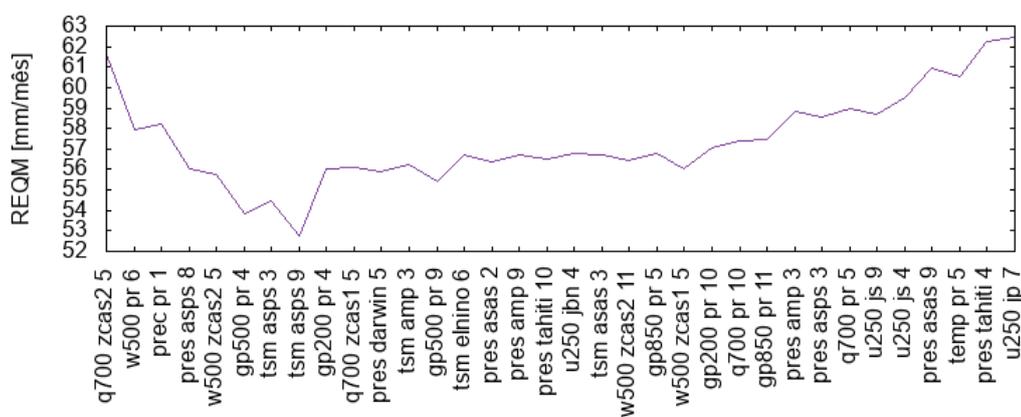


Figura B.36: Resultado do primeiro ciclo do método *stepwise* para a precipitação mensal: 08 variáveis selecionadas.

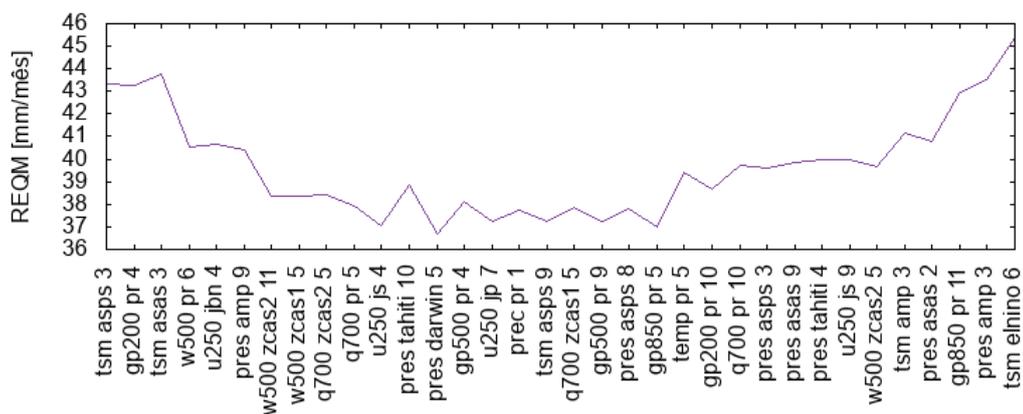


Figura B.37: Resultado do segundo ciclo do método *stepwise* para a precipitação mensal: 15 variáveis selecionadas.

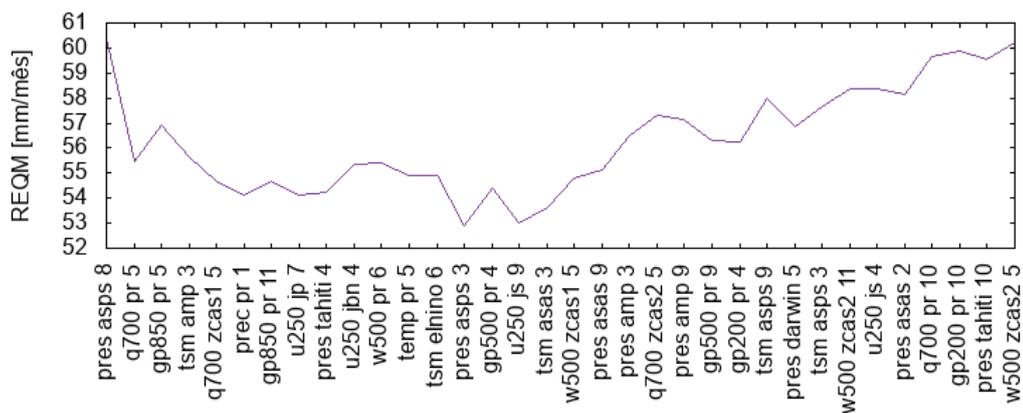


Figura B.38: Resultado do terceiro ciclo do método *stepwise* para a precipitação mensal: 14 variáveis selecionadas.

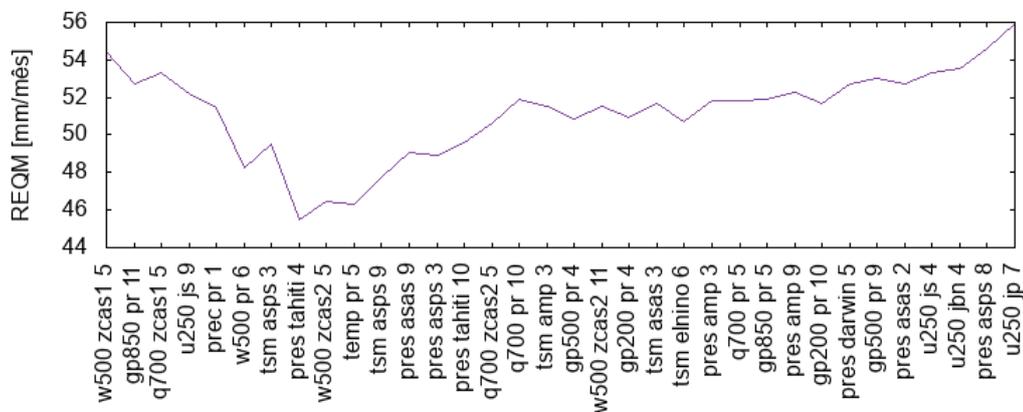


Figura B.39: Resultado do quarto ciclo do método *stepwise* para a precipitação mensal: 08 variáveis selecionadas.

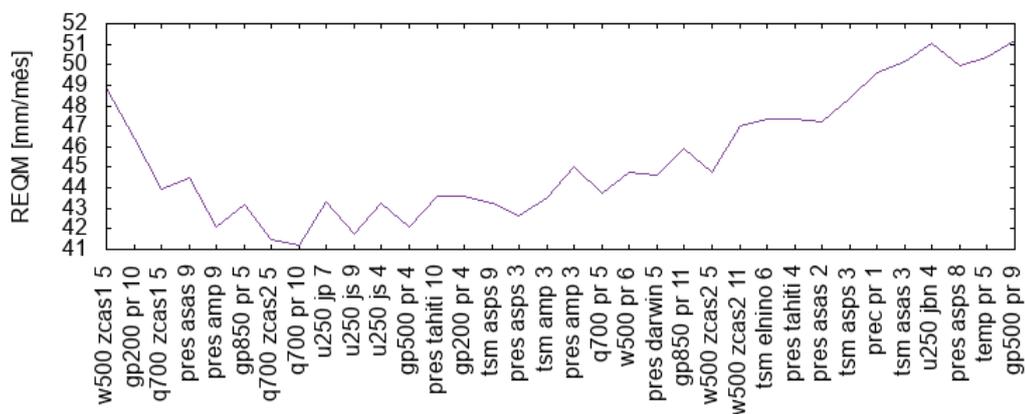


Figura B.40: Resultado do quinto ciclo do método *stepwise* para a precipitação mensal: 08 variáveis selecionadas.

Nas Figuras B.41, B.42, B.43, B.44 e B.45 são mostrados os resultados dos cinco ciclos do método *stepwise* para a precipitação sazonal.

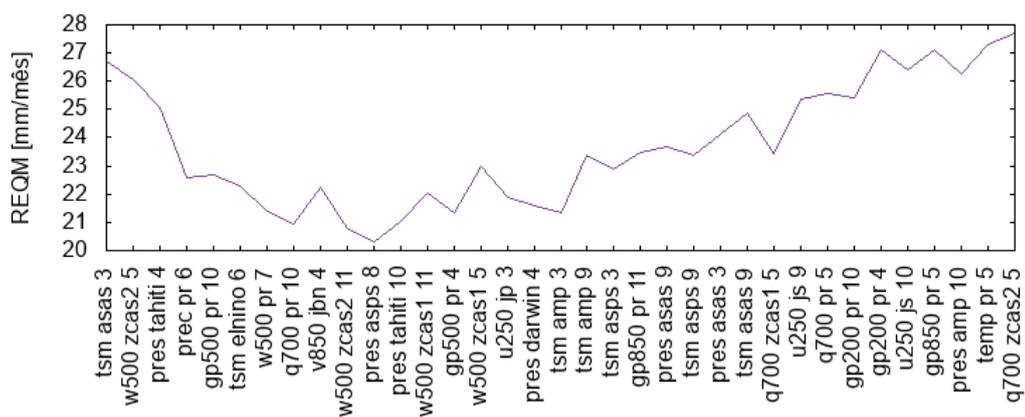


Figura B.41: Resultado do primeiro ciclo do método *stepwise* para a precipitação sazonal: 11 variáveis selecionadas.

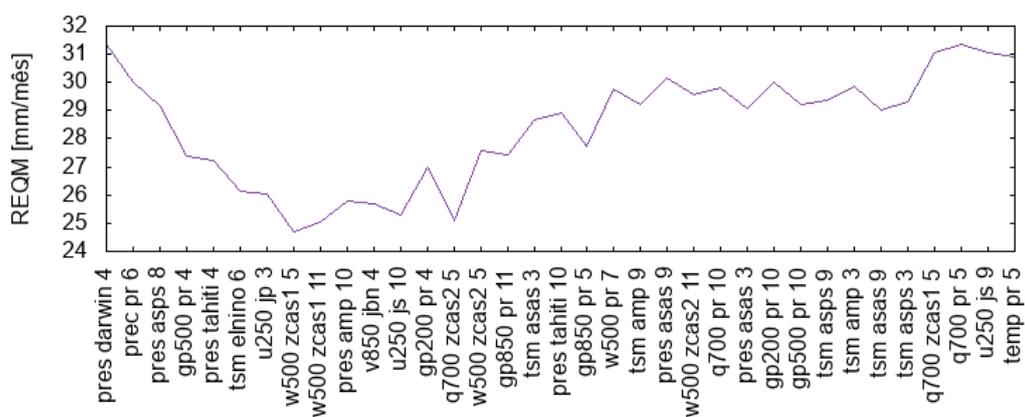


Figura B.42: Resultado do segundo ciclo do método *stepwise* para a precipitação sazonal: 08 variáveis selecionadas.

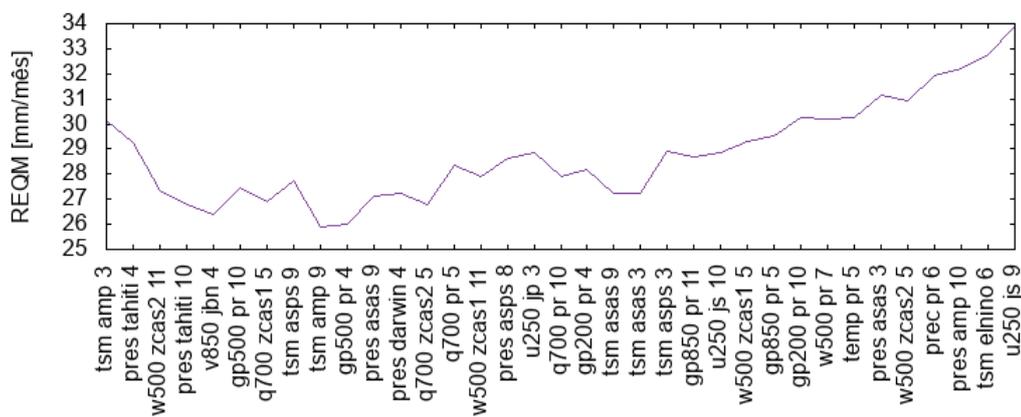


Figura B.43: Resultado do terceiro ciclo do método *stepwise* para a precipitação sazonal: 10 variáveis selecionadas.

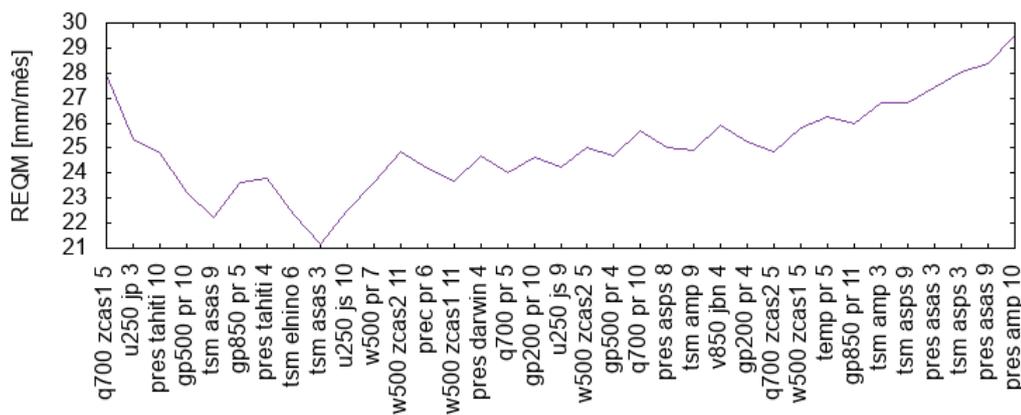


Figura B.44: Resultado do quarto ciclo do método *stepwise* para a precipitação sazonal: 09 variáveis selecionadas.

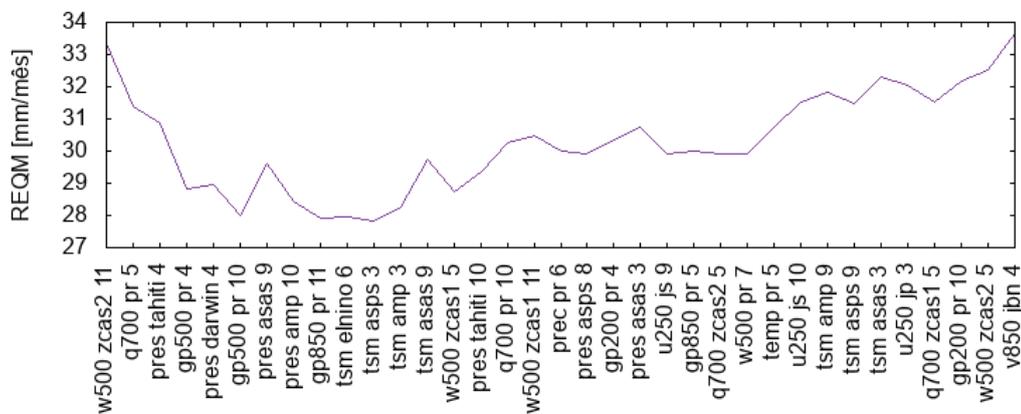


Figura B.45: Resultado do quinto ciclo do método *stepwise* para a precipitação sazonal: 11 variáveis selecionadas.