

UFRRJ
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM
MATEMÁTICA E COMPUTACIONAL

DISSERTAÇÃO

APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA COMPUTACIONAL
NO PREENCHIMENTO DE FALHAS DE SÉRIES TEMPORAIS
CLIMATOLÓGICAS

ELUÃ RAMOS COUTINHO

2014



UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E
COMPUTACIONAL

APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA COMPUTACIONAL NO
PREENCHIMENTO DE FALHAS DE SÉRIES TEMPORAIS CLIMATOLÓGICAS

ELUÃ RAMOS COUTINHO

Sob a Orientação do Professor

Dr. Angel Ramon Sanchez Delgado

e Co-orientação do Professor

Dr. Robson Mariano da Silva.

Dissertação submetida como requisito parcial para obtenção do grau de Mestre em Modelagem Matemática e Computacional, no Curso de Pós-graduação em Modelagem Matemática e Computacional, Área de concentração em Inteligência Computacional e Otimização.

Seropédica, RJ

Abril de 2014

006.3
C871a
T

Coutinho, Eluã Ramos, 1988-

Aplicação de técnicas de inteligência computacional no preenchimento de falhas de séries temporais climatológicas / Eluã Ramos Coutinho - 2014.

148 f.: il.

Orientador: Angel Ramon Sanchez Delgado.
Dissertação (mestrado) - Universidade Federal Rural do Rio de Janeiro, Curso de Pós-Graduação em Modelagem Matemática e Computacional.

Bibliografia: f. 87-92.

1. Inteligência artificial - Teses. 2. Inteligência computacional - Teses. 3. Redes neurais (Computação) - Teses. 4. Algoritmos genéticos - Teses. 5. Climatologia - Teses. I. Delgado, Angel Ramon Sanchez, 1956-. II. Universidade Federal Rural do Rio de Janeiro. Curso de Pós-Graduação em Modelagem Matemática e Computacional. III. Título.

UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS EXATAS
CURSO DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E
COMPUTACIONAL

ELUÃ RAMOS COUTINHO

Dissertação submetida como requisito parcial para obtenção do grau de Mestre em Ciências,
no Curso de Pós-graduação em Modelagem Matemática e Computacional, área de
concentração em Inteligência Computacional e Otimização.

Dissertação aprovada em ____/____/____

Angel Ramon Sanchez Delgado. Prof. Dr. UFRRJ
(Orientador)

Robson Mariano. Prof. Dr. UFRRJ
(Co-orientador)

Felipe Maia Galvão França. Prof. Ph.D. UFRJ

Ednaldo Oliveira dos Santos. Prof. Dr. UFRRJ

Dedico este trabalho aos meus pais Eloi e Jussara, meus irmãos Erlon e Ebert, e a Pollyanna, pela paciência, compreensão durante esta importante etapa da minha vida.

AGRADECIMENTOS

A Deus, pela força, sabedoria, e por ter permitido dar mais esse grande passo em minha vida.

A Universidade Federal Rural do Rio de Janeiro (UFRRJ) por esses anos de muito aprendizado.

Ao Angel Ramon Sanchez Delgado, pela orientação e determinação para que este trabalho fosse concluído. Obrigado pela amizade.

Ao Robson Mariano da Silva por aceitar me coorientar, e auxiliar com muitas barreiras que me deparei até a conclusão deste trabalho. Obrigado pela confiança e amizade.

Ao Programa de Pós-Graduação em Modelagem Matemática e Computacional da UFRRJ, em especial Wanderson José Lambert, Priscila Machado Vieira de Lima e Janaína Gama pela disposição em ajudar sempre que precisei.

A Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro-FAPERJ pela bolsa concedida.

Ao Instituto Nacional de Meteorologia (INMET), pela disponibilização dos dados meteorológicos.

Minha prima Letícia pela amizade ajuda e incentivo, e aos amigos Italo, Charlan, Ronilson, Diogo, Márcio, Arnaldo, Alessandro, Leandro, e todos do grupo do Mestrado pelos momentos bons e difíceis que passamos juntos.

Aos meus familiares, minhas tias Penha, Lúcia e Jacira por todo suporte e segurança que me deram não só nestes dois anos mais em toda minha vida.

Agradeço àqueles que foram a grande força e razão para isso ser possível. Aos meus pais Eloi e Jussara, meus irmãos Erlon e Ebert, e a Pollyanna, por todo amor companheirismo e compreensão prestada em mais esse passo de minha vida.

Por fim a todas as pessoas que por motivo de esquecimento não foram citadas anteriormente, vou deixando neste espaço minhas sinceras desculpas.

RESUMO

COUTINHO, Eluã Ramos. **Aplicação de técnicas de inteligência computacional no preenchimento de falhas de séries temporais climatológicas**. 2014. 163f. Dissertação (Mestrado em Modelagem Matemática e Computacional). Instituto de Ciências Exatas, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2014.

As informações sobre o clima possuem extrema importância para a sociedade e instituições públicas e privadas. Geralmente os dados reais associados com fenômenos climatológicos são obtidos utilizando sensores que estão sujeitos a falhas. Desta forma geram-se dados inválidos que podem prejudicar a análise. O presente estudo visa a aplicação de técnicas de inteligência computacional no desenvolvimento de um modelo híbrido composto por Redes Neurais Artificiais (RNAs) e Algoritmos Genéticos (AGs), para preencher as falhas das variáveis mensais de evaporação, temperatura máxima e umidade relativa do ar pertencentes a seis regiões do estado do Rio de Janeiro. Os resultados obtidos foram comparados com os métodos de Médias, Regressão Linear Múltipla, RNAs *Multilayer Perceptron* (MLP) e *Radial Basis Function* (RBF). A performance do modelo proposto foi avaliada utilizando o coeficiente de correlação, erro médio absoluto, raiz do erro médio quadrático, erro médio percentual, índice de concordância e índice de confiança. A avaliação dos resultados mostraram que na maioria dos casos o modelo proposto apresentou índices superiores aos demais modelos, chegando a obter 0,99 de correlação, e erros médios percentuais entre 10% e 1%. No que tange a performance do preenchimento de falhas reais o modelo proposto obteve valor médio de 100% para o preenchimento de dados referentes a evaporação, 89,2% para temperatura máxima do ar e 75,3% para umidade relativa do ar. Logo, pode-se concluir que o modelo proposto é uma alternativa no preenchimento de falhas de dados meteorológicos.

Palavras chaves: Clima, Redes Neurais Artificiais (RNAs), Algoritmos Genéticos (AGs).

ABSTRACT

COUTINHO, Eluã Ramos. **Application of computational intelligence techniques in gap filling of climatological time series**. 2014. 163f. Dissertation (Master of Science in Mathematical and Computational Modelling). Instituto de Ciências Exatas, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2014.

The climate information have extreme importance for the society and public and private institutions. Generally, the actual data associated with climatological phenomena are obtained using sensors that are subject to failures. Thus, invalid data are generated and may negatively influence the analysis. The present study aims the application of computational intelligence techniques in the development of a hybrid model composed of Artificial Neural Networks (ANN) and Genetic Algorithms (GAs), to fill the gaps of monthly variables, as evaporation, maximum temperature and relative humidity from six regions of the state of Rio de Janeiro - Brazil. The results were compared with the Average data of the regions, Multiple Linear Regression, ANNs Multilayer Perceptron (MLP) and Radial Basis Function (RBF). The performance of the proposed model was evaluated using the correlation coefficient between the estimated data and actual data, mean absolute error, root mean square error, average percentage error, concordance index and confidence index. The evaluation results showed that, in most cases, the proposed model were more accurate than the other models indices, reaching correlation value equal to 0.99, and average percentage errors between 10% and 1%. Regarding the performance of the actual filling of data gaps, the proposed model had a mean value of 100% for the filling of evaporation data, 89.2% for maximum air temperature and 75.3% for relative humidity. Therefore, it can be concluded that the model is an alternative to filling the gaps in meteorological data.

Keywords: Climate, Artificial Neural Networks (ANN), Genetic Algorithms (GAs).

LISTA DE FIGURAS

Figura 1: Estação Meteorológica do Instituto de Astronomia da USP. Fonte: http://www5.usp.br/19536/estacao-meteorologica-do-iag-completa-80-anos-de-operacoes/ ...	6
Figura 2: Estações Meteorológicas Automáticas.....	7
Figura 3: regiões utilizadas para preenchimento de dados de uma região faltante.	8
Figura 4: Estrutura de um neurônio biológico. Fonte: MACHADO, 2005.....	10
Figura 5:Estrutura de um neurônio Artificial. Fonte: BRAGA ; CARVALHO; LUDEMIR, 2012.....	11
Figura 6:Estrutura de um RNA do tipo MLP.	13
Figura 7:Modelo de neurônio não linear . Fonte adaptado de Haykin (2001).	14
Figura 8: Gráficos dos comportamentos das funções de transferência. (a) Função Degrau, (b) Função Linear, (1) Função Logística, (2) Função Tangente Hiperbólica	15
Figura 9: Estrutura de uma RNA do tipo RBF.	19
Figura 10: Exemplo de uma função hipotética de máximo local e global. Fonte adaptada de Lindel(2012).	22
Figura 11: Estrutura básica do Funcionamento de um AG.	23
Figura 12: Esquema de um cromossomo binário.	24
Figura 13: Seleção utilizando o método da roleta. Fonte adaptada de BRAGA; CARVALHO; LUDEMIR, 2012.	24
Figura 14: Exemplo de <i>crossover</i> de 2 pontos.	25
Figura 15: Exemplo da ocorrência do processo de mutação.....	26
Figura 16: Mapa do estado do Rio de Janeiro com o conjunto de regiões utilizadas.	27
Figura 17: modelo proposto para preenchimento de falhas.	29
Figura 18: Percentual de dados registrados corretamente e porcentagem de falhas ocorridas em cada dado.	31
Figura 19: Exemplo do processo de triagem aplicado aos dados.	32
Figura 20: Estrutura da RNA RBF aplica pelo MATLAB. Fonte: http://www.mathworks.com/help/nnet/ug/radial-basis-neural-networks.html	35
Figura 21: Exemplo de cromossomo adotado para configurar a RNA.	36
Figura 22: Dados estimados de evaporação pelo modelo MLP-AG para a estação de Campos dos Goytacazes.	44
Figura 23: Estação de Campos dos Goytacazes , (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.....	45
Figura 24: Estação de Cordeiro, (A) Dados reais e estimados de evaporação com o modelo MLP-AG, (B) Dados reais e estimados de evaporação com o modelo MLP.....	46
Figura 25: Estação de Cordeiro, (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.	47
Figura 26: Dados estimados de evaporação com o modelo MLP-AG para a estação de Itaperuna.....	48
Figura 27: Resultados da estimação de evaporação obtidos com o modelo MLP para a estação do Rio de Janeiro.....	49
Figura 28: Estação do Rio de Janeiro, (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.	50
Figura 29: Resultados de estimativa de evaporação obtidos com o modelo MLP-AG para estação Paty do Alferes.	51
Figura 30: Estação de Paty do Alferes, (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.	52

Figura 31: Dados de evaporação estimados com o modelo MLP-AG para a estação de Resende.	53
Figura 32: Estação de Resende, (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.	54
Figura 33: Dados estimados de temperatura máxima do ar obtidos com MLP-AG para estação de Campos dos Goytacazes.	57
Figura 34: Estação de Campos dos Goytacazes, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.	58
Figura 35: Resultados da estimativa de temperatura máxima do ar obtidos com o modelo MLP-AG para estação de Cordeiro.	59
Figura 36: Estação de Cordeiro, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.	60
Figura 37: Resultados da estimativa de temperatura máxima do ar obtidos com o modelo MLP-AG para estação de Itaperuna.	61
Figura 38: Estação de Itaperuna, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.	62
Figura 39: Dados estimados de temperatura máxima do ar com o modelo MLP-AG para estação do Rio de Janeiro.	63
Figura 40: Estação do Rio de Janeiro, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.	64
Figura 41: Resultados da estimativa de temperatura máxima do ar obtidos com o modelo MLP-AG para estação de Paty do Alferes.	65
Figura 42: Estação de Paty do Alferes, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.	66
Figura 43: Resultados de estimativa de temperatura máxima do ar obtidos com o modelo MLP-AG para estação de Resende.	67
Figura 44: Estação de Resende, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.	68
Figura 45: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Campos dos Goytacazes.	71
Figura 46: Estação de Campos dos Goytacazes, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.	72
Figura 47: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Cordeiro.	73
Figura 48: Estação de Cordeiro, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.	74
Figura 49: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Itaperuna.	75
Figura 50: Estação de Itaperuna, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.	76
Figura 51: Resultados de estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação do Rio de Janeiro.	77
Figura 52: Estação do Rio de Janeiro, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.	78
Figura 53: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Paty do Alferes.	79
Figura 54: Estação Paty do Alferes, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.	80
Figura 55: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Resende.	81

Figura 56: Estação de Resende, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.	82
---	----

LISTA DE TABELAS

Tabela 1: Estações e suas coordenadas geográficas.	28
Tabela 2: Quantidade de falhas ocorridas no período de 31/05/2002 a 31/12/2012.....	30
Tabela 3: Conjunto de estações utilizadas na estimativa ou predição de dados.....	34
Tabela 4: Critérios para avaliação e análise do desempenho de modelos com base no índice de confiança	40
Tabela 5: Índices de correlação apresentados para a variável de evaporação nas estações estudadas.	42
Tabela 6: Conjunto de variáveis estimadoras de evaporação definidas com o AG.	42
Tabela 7: Índices de correlação dos conjuntos de estimadores definidos manualmente para variável de evaporação.....	43
Tabela 8: Índices de correlação dos conjuntos de estimadores definidos pelo AG para variável de evaporação.....	43
Tabela 9: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Campos dos Goytacazes.	44
Tabela 10: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Cordeiro.....	46
Tabela 11: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Itaperuna.	47
Tabela 12: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação do Rio de Janeiro.	49
Tabela 13: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Paty do Alferes.....	51
Tabela 14: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Resende.	52
Tabela 15: Índice de correlação apresentado para as variáveis de temperatura máxima do ar.....	54
Tabela 16: Conjunto de variáveis estimadoras de temperatura máxima do ar definidas com o AG.....	55
Tabela 17: Índices de correlação dos conjuntos de estimadores definidos manualmente para variável de temperatura máxima do ar.....	56
Tabela 18: Índices de correlação dos conjuntos de estimadores definidos pelo AG para variável de temperatura máxima do ar.....	56
Tabela 19: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da estação de Campos dos Goytacazes.	57
Tabela 20: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da região de Cordeiro.....	59
Tabela 21: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da região de Itaperuna.....	61
Tabela 22: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da estação do Rio de Janeiro.	63
Tabela 23: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da estação de Paty do Alferes.....	65
Tabela 24: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da estação de Resende.....	67
Tabela 25: Índice de correlação apresentado para as variáveis de umidade relativa do ar.....	68
Tabela 26: Conjunto de variáveis estimadoras de umidade relativa do ar definidas com o AG.....	69

Tabela 27: Índices de correlação dos conjuntos de estimadores definidos manualmente para variável de umidade relativa do ar.....	70
Tabela 28: Índices de correlação dos conjuntos de estimadores definidos pelo AG para variável de umidade relativa do ar.....	70
Tabela 29: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação de Campos dos Goytacazes.....	71
Tabela 30: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação de Cordeiro.	73
Tabela 31: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação de Itaperuna.....	75
Tabela 32: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação do Rio de Janeiro.....	77
Tabela 33: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação de Paty do Alferes.	79
Tabela 34: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação de Resende.	81

SUMÁRIO

1. INTRODUÇÃO	1
2. JUSTIFICATIVA.....	3
3. OBJETIVOS	4
3.1. Objetivo Geral	4
3.2. Objetivos Específicos	4
4. FUNDAMENTAÇÃO TEÓRICA	5
4.1. Estações Meteorológicas Convencionais	5
4.2. Estações Meteorológicas Automáticas	6
4.3. Homogeneidade das Séries Temporais Climatológicas.....	7
4.4. Preenchimento de Falhas em Séries Temporais Climatológicas.....	7
4.5. Médias de Regiões Próximas	8
4.6. Regressão Linear Múltipla	8
4.7. Rede Neurais Artificiais (RNAs).....	9
4.7.1. Estrutura Básica de um Neurônio Artificial	10
4.7.2. Formas de Aprendizagem	12
4.7.3. Redes Perceptrons de Múltiplas Camadas (MLP)	13
4.7.3.1. Back-propagation	16
4.7.4. Redes de Função de Base Radial (RBF).....	18
4.8. Algoritmo Genético (AG).....	21
5. MATERIAIS E MÉTODOS	27
5.1. Base de Dados	27
5.2. Modelo Proposto para Preenchimento de Falhas	28
5.2.1. Levantamento Quantitativo de Falhas.....	29
5.2.2. Triagem de Dados	32
5.2.3. Normalização	33
5.2.4. Definição das Estações Utilizadas na Estimativa dos Dados	33
5.2.5. Definição do Conjunto de Dados de Treino e Validação	34
5.2.6. Aplicação das Técnicas de Preenchimento de Falhas	34
5.2.7. Configuração da RNA MLP	34
5.2.8. Configuração da RNA RBF.....	35
5.2.9. Configuração da RNA MLP com AG.....	36
5.2.10. Interpretação dos Resultados	38
5.2.11. Avaliação do Desempenho	38
5.2.12. Preenchimento de Falhas.....	40
6. RESULTADOS	41
6.1. Configurações.....	41
6.2. Resultados de Evaporação	41
6.2.1. Análise do Conjunto de Dados Estimadores para Variável de Evaporação	42
6.2.2. Resultados e Preenchimento das Falhas de Evaporação para estação de Campos dos Goytacazes.....	43
6.2.3. Resultados e Preenchimento das Falhas de Evaporação para estação de Cordeiro.....	45
6.2.4. Resultados e Preenchimento das Falhas de Evaporação para estação de Itaperuna.....	47
6.2.5. Resultados e Preenchimento das Falhas de Evaporação para estação do Rio de Janeiro	48
6.2.6. Resultados e Preenchimento das Falhas de Evaporação para estação de Paty do Alferes.....	50
6.2.7. Resultados e Preenchimento das Falhas de Evaporação para estação de Resende	52

6.3. Resultados de Temperatura Máxima do Ar	54
6.3.1. Análise do Conjunto de Dados Estimadores para Variável de Temperatura Máxima do Ar	55
6.3.2. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Campos dos Goytacazes.....	56
6.3.3. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Cordeiro.....	58
6.3.4. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Itaperuna.....	60
6.3.5. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação do Rio de Janeiro	62
6.3.6. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Paty do Alferes	64
6.3.7. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Resende	66
6.4. Resultados de Umidade Relativa do Ar	68
6.4.1. Análise do Conjunto de Dados Estimadores para Variável de Umidade Relativa do Ar	69
6.4.2. Resultados e Preenchimento das falhas de Umidade Relativa do Ar para estação de Campos dos Goytacazes.....	70
6.4.3. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação de Cordeiro.....	72
6.4.4. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação de Itaperuna	74
6.4.5. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação do Rio de Janeiro	76
6.4.6. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação de Paty do Alferes	78
6.4.7. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação de Resende	80
7. DISCUSSÃO	83
8. CONCLUSÕES.....	85
9. TRABALHOS FUTUROS	86
10. REFERÊNCIAS BIBLIOGRÁFICAS	87
ANEXOS	93
ANEXO A- Dados Meteorológicos da Estação de Campos dos Goytacazes	93
ANEXO B- Dados Meteorológicos da Estação de Cordeiro	96
ANEXO C- Dados Meteorológicos da Estação de Itaperuna.....	99
ANEXO D- Dados Meteorológicos da Estação de Paty do Alferes	102
ANEXO E- Dados Meteorológicos da Estação de Resende	105
ANEXO F- Dados Meteorológicos da Estação do Rio de Janeiro	108
APÊNDICES	111
APÊNDICE A - Configurações das RNAs utilizadas na estimativa dos dados de evaporação	111
APÊNDICE B - Configurações das RNAs utilizadas na estimativa dos dados de temperatura máxima do ar.	113
APÊNDICE C - Configurações das RNAs utilizadas na estimativa dos dados de umidade relativa do ar.	115
APÊNDICE D – Dados estimados de evaporação para estação de Campos Goytacazes com o modelo MLP-AG	117

APÊNDICE E – Dados estimados de evaporação para estação de Cordeiro com o modelo MLP-AG.....	118
APÊNDICE F – Dados estimados de evaporação para estação de Itaperuna com o modelo MLP-AG.....	119
APÊNDICE G – Dados estimados de evaporação para estação do Rio de Janeiro com o modelo MLP.....	120
APÊNDICE H – Dados estimados de evaporação para estação de Paty do Alferes com o modelo MLP-AG.....	121
APÊNDICE I – Dados estimados de evaporação para estação de Resende com o modelo MLP-AG.....	122
APÊNDICE J – Dados estimados de temperatura máxima do ar para estação de Campos dos Goytacazes com o modelo MLP-AG.....	123
APÊNDICE L – Dados estimados de temperatura máxima do ar para estação de Cordeiro com o modelo MLP-AG.....	124
APÊNDICE M – Dados estimados de temperatura máxima do ar para estação de Itaperuna com o modelo MLP-AG.....	125
APÊNDICE N – Dados estimados de temperatura máxima do ar para estação do Rio de Janeiro com o modelo MLP-AG.....	126
APÊNDICE O – Dados estimados de temperatura máxima do ar para estação de Paty do Alferes com o modelo MLP-AG.....	127
APÊNDICE P – Dados estimados de temperatura máxima do ar para estação de Resende com o modelo MLP-AG.....	128
APÊNDICE Q – Dados estimados de umidade relativa do ar para estação de Campos dos Goytacazes com o modelo MLP-AG.....	129
APÊNDICE R – Dados estimados de umidade relativa do ar para estação de Cordeiro com o modelo MLP-AG.....	130
APÊNDICE S – Dados estimados de umidade relativa do ar para estação de Itaperuna com o modelo MLP-AG.....	131
APÊNDICE T – Dados estimados de umidade relativa do ar para estação do Rio de Janeiro com o modelo MLP-AG.....	132
APÊNDICE U – Dados estimados de umidade relativa do ar para estação de Paty do Alferes com o modelo MLP-AG.....	133
APÊNDICE V – Dados estimados de umidade relativa do ar para estação de Resende com o modelo MLP-AG.....	134

1. INTRODUÇÃO

Os fenômenos climáticos e processos da atmosfera são estudados há milênios. A curiosidade sobre tais fenômenos eram tamanhas que levaram Aristóteles a dar origem ao estudo sobre a meteorologia. Este sábio, mesmo sem ferramentas adequadas, estudou nuvens, chuva, vento, trovão, orvalho e condições de tempo associadas, chegando a escrever um livro chamado “Meteorologia”, cujo significado é “coisas acima da Terra” (TEIXEIRA; TAVARES; ARAÚJO, 2013).

Hoje, há mais de dois milênios à frente e com auxílio de muitos instrumentos, o estudo sobre a meteorologia continua sendo um trabalho complexo, fazendo-se necessária a coleta de dados meteorológicos através de equipamentos como satélites, balões, radares, sensores e estações meteorológicas. Tal estudo pode necessitar de um longo tempo, levando em consideração o número de informações que seguem um regime de coleta contínuo, sendo obtidos em períodos de minutos, horas, dias, ou meses, e de forma constante, gerando um vasto volume de dados (VENTURA, 2012).

Essas informações possuem um grande valor histórico, podendo ser utilizadas por empresas privadas, organizações governamentais e instituições acadêmicas e de pesquisas. Tal importância se deve ao imenso campo de aplicação dos dados meteorológicos, que podem ser empregados em muitas áreas. Na segurança civil, os mesmos podem reduzir danos materiais e humanos através de alertas sobre tempestades, inundações e ressacas. Na área agrícola, os dados podem ser empregados na agricultura de precisão para acompanhamento da cultura, determinação de área e época ideal para plantio e colheita, realização do cálculo do balanço hídrico, entre outros. No campo energético, os dados podem auxiliar no controle de níveis de reservatório de usinas e na busca de informações de novas fontes alternativas de energia. Essas informações ainda podem ser aplicadas em outras áreas como transportes, ecologia, saúde, e na previsão do tempo, auxiliando assim as tomadas de decisões estratégicas e a resolução de problemas (COUTINHO; VALE, 2011).

Contudo, se faz necessário a realização de análises das informações registradas antes que os dados possam ser aplicados em qualquer área. Isso ocorre, porque muitas vezes a análise e o emprego dos dados são dificultados principalmente devido às instalações dos equipamentos e estações meteorológicas ocorrerem em locais distantes e ao ar livre. Consequentemente, os aparelhos estão sujeitos às condições naturais do meio ambiente, ausência de observadores, ação de animais próximos aos equipamentos e fortes mudanças no tempo, podendo assim atrapalhar as leituras dos equipamentos, causando medições inconsistentes e gerando falhas no armazenamento dos dados. Além disso, alguns destes são aparelhos eletrônicos e estão sujeitos a ocorrência de problemas. As condições naturais do meio ambiente e problemas técnicos podem interferir então na leitura e armazenamento dos dados (VENTURA, 2012).

Na maioria das vezes as séries de dados climáticos com perda de informações ou falhas não podem ser utilizadas em grandes pesquisas que necessitam de séries contínuas, o que cria a necessidade do emprego de metodologias que possam auxiliar no preenchimento dessas séries temporais meteorológicas. Esse preenchimento pode ser realizado através da utilização dos dados das séries históricas junto a métodos estatísticos e matemáticos que possuem uma forma de implementação fácil, como médias, interpolações, regressões lineares e outros.

No entanto, para poder realizar o preenchimento através desses métodos, há a necessidade de utilizar dados de outras regiões que possuam características climáticas similares ou de um grande número de dados ininterruptos da própria estação. Mesmo assim, algumas dessas técnicas são estritamente lineares, não considerando a não linearidade dos

dados meteorológicos e ignoram a variação espacial e local das variáveis em estudo, o que pode gerar erros grotescos (ROLIM; CATALÃO; TEIXEIRA, 2010; HUTH; NEMESOVÁ, 1995). Para fazer uso dessas técnicas de preenchimento em falhas de dados meteorológicos, é necessário ter conhecimento das características climáticas do ambiente em questão, o que acaba tornando o trabalho uma operação extremamente difícil.

Estes fatores impulsionam uma busca contínua por novos métodos que não necessitem de tantas informações e possam ter como soluções dados muito próximos dos reais. Isto possibilita a utilização de técnicas de inteligência computacional conhecidas como meta-heurísticas na solução de problemas de preenchimento de falhas, pois obtêm resultados de forma rápida não necessitando de um grande número de informações e se caracterizam por considerar casos lineares e não lineares. A busca por metodologias que possam aprimorar o processo de estimação e preenchimento de falhas em dados de estações meteorológicas é um grande desafio para os pesquisadores.

2. JUSTIFICATIVA

É primordial destacar que o estudo da Meteorologia é essencial na atualidade devido possuir aplicações em variadas áreas, o que acaba incentivando pesquisadores a desenvolverem um grande número de pesquisas com a utilização dos dados meteorológicos.

Contudo, a grande demanda por informações contínuas e confiáveis é um dos principais problemas de se utilizar dados meteorológicos. Tais problemas podem ser ocasionados pela ausência ou inconsistências dos dados gerados por uma determinada falha. Estes fatores, se não tratados, podem prejudicar ou mesmo inviabilizar pesquisas que demandem de um grande número de informações.

Em vista disso, a correção desses dados vem se tornando cada vez mais comum através da utilização de métodos matemáticos e estatísticos. Entretanto, algumas dessas técnicas não conseguem representar com qualidade uma estimativa dos valores reais. Isso ocorre, principalmente porque cada variável meteorológica possui características intrínsecas aos processos de origem, o que dificulta o emprego de algumas técnicas.

Contudo, a não existência de uma metodologia padrão para corrigir tais erros acaba gerando uma busca por inúmeras técnicas com a intenção de melhorar os resultados. Assim, está se tornando cada vez mais comum a utilização de meta-heurísticas, como as técnicas de Redes Neurais Artificiais (RNAs) e os Algoritmos Genéticos (AG), no auxílio da estimativa de valores para preenchimento de falhas.

As técnicas de RNAs são inspiradas na estrutura neural de organismos inteligentes, se caracterizando por reconhecimento de padrões e generalização de informações e pela capacidade de aprender e adquirir conhecimento através da experiência (HAYKIN, 2001). Já os AG são métodos matemáticos que se caracterizam por problemas de busca e otimização em sistemas complexos, baseando-se na teoria de Darwin da evolução biológica (BECKAMANN, 2010).

Cada uma dessas técnicas possuem um imenso campo de aplicação, e a junção das mesmas com o intuito de aplicar uma meta-heurística híbrida e assim melhorar resultados, é um grande incentivo para pesquisadores que procuram à inovação e aprimoramento de técnicas.

Alem disso, a utilização conjunta dessas técnicas ainda pode auxiliar na aceleração e na melhor utilização dos dados disponíveis para o processo de preenchimento de falhas. Para tanto, o presente estudo aplicou uma heurística híbrida no preenchimento de dados ausentes.

3. OBJETIVOS

3.1. Objetivo Geral

Aplicar uma metodologia híbrida, composta por Redes Neurais Artificiais e Algoritmos Genéticos no preenchimento de falhas em séries históricas de dados climatológicos.

3.2. Objetivos Específicos

- Avaliar a técnica de pré-processamento utilizada.
- Verificar a capacidade do Algoritmo Genético na seleção de variáveis estimadoras.
- Verificar a capacidade de diferentes metodologias no preenchimento de falhas de séries temporais climatológicas.
- Integrar os Algoritmos Genéticos as Redes Neurais Artificiais.
- Comparar o desempenho das metodologias aplicadas.

4. FUNDAMENTAÇÃO TEÓRICA

A influência do clima em variadas áreas, como o setor econômico da agricultura que em 2011 foi responsável por 23,82% do produto interno bruto (PIB) do Brasil (CEPEA, 2013), e as constantes ocorrências de catástrofes na atmosfera terrestre, demonstram claramente a importância de se monitorar as informações ambientais.

Esse monitoramento pode auxiliar na diminuição de perdas, maximização de lucros ou em pesquisas agrícolas. A Sociedade Brasileira de Meteorologia destaca que a coleta e a análise de dados geofísicos (meteorológicos, oceânicos, hidrológicos, sísmicos, etc.) e ambientais são de extrema importância para um melhor entendimento do nosso ambiente natural e suas variações no tempo e no espaço, proporcionando um suporte científico para decisões governamentais voltadas para a preservação da vida em nosso Planeta (SBMET, 2008).

No entanto, para realizar estudos na área de meteorologia, há a necessidade de coletar dados ambientais, e essas informações podem ser adquiridas através do monitoramento do meio ambiente feito por estações de coleta de dados. Estas estações podem ser meteorológicas, hidrometeorológicas e agrometeorológicas, que são divididas em convencionais ou automáticas. Nas convencionais é necessária a presença de um observador para fazer a leitura dos aparelhos e o armazenamento dos dados, sendo que este observador deve ser uma pessoa capacitada e responsável para que os dados sejam confiáveis. Nas automáticas os dados são enviados eletronicamente para uma central de processamento (FERRAZ, 2010).

4.1. Estações Meteorológicas Convencionais

As estações meteorológicas convencionais são compostas por vários aparelhos isolados que medem e registram parâmetros meteorológicos de forma contínua. Elas devem seguir normas para sua criação, como ter uma área mínima de 4x5m, ser cercada com tela de arame de 1,2m de altura, com a dimensão maior no sentido norte-sul, seu piso deve ser coberto com grama mantida sempre verde e aparada, ter um abrigo meteorológico feito de madeira com laterais de venezianas e piso vazado, a porta do abrigo deve abrir para o sul, e o piso deve ficar 1,2m de altura em relação do solo (Figura 1). Os aparelhos indispensáveis são o pluviômetro, termômetro de máxima e mínima, anemômetro, psicrômetro, termo-higrógrafo e outros (HELDWEIN et al., 2007).



Figura 1: Estação Meteorológica do Instituto de Astronomia da USP. Fonte: <http://www5.usp.br/19536/estacao-meteorologica-do-iag-completa-80-anos-de-operacoes/>

O monitoramento das informações meteorológicas é realizado por um observador humano em intervalos de tempo padronizados, o mesmo tem a obrigação de enviar as informações a um centro coletor por meio de uma comunicação qualquer (INMET, 2013).

Contudo, com o passar dos anos o registro de informações foi facilitado com o advento das estações meteorológicas automáticas, que tem como pretensão inicial realizar um maior número de registros, tendo assim um monitoramento mais apurado e minimizando problemas oriundos ocorridos pela falta ou ineficiência de um observador.

4.2. Estações Meteorológicas Automáticas

Uma estação meteorológica automática, que é conhecida como plataforma de coleta de dados (PCD), é um dispositivo eletrônico que monitora o meio ambiente fazendo registros dos dados e realizando a transmissão automática de dados meteorológicos.

Os dispositivos acoplados as PCDs, como carregadores de bateria, painéis solares, modem de linha discada, modem celular GSM, modem de linha privativa, rádio UHF, Comunicação TCP/IP e diversos tipos de sensores, facilitam a telemetria dos dados e o registro de diferentes parâmetros meteorológicos, auxiliando as instalações das estações em variadas localidades do território nacional (WMO, 2003).

Dependendo da necessidade das informações, as PCDs ainda podem ser programadas para efetuar medições meteorológicas em diversos espaços de tempo, podendo assim facilitar o monitoramento das informações em tempo real. A Figura 2 exibe alguns tipos de PCD's.



Figura 2: Estações Meteorológicas Automáticas

No entanto, a manutenção das PCDs pode ser prejudicada devido a dificuldade de acesso às localidades onde se encontram instaladas, isso somado a ocorrência de determinados problemas nos sensores, prejudicam a coleta e consequentemente a homogeneidade dos dados.

4.3. Homogeneidade das Séries Temporais Climatológicas

O manuseio dos aparelhos por pessoas não qualificadas, ou mesmo a falta de pessoas que possam realizar o registro em estações convencionais junto aos inúmeros problemas que podem ocorrer em uma PCD, como falha dos sensores, a falta de energia, ações do meio ambiente como animais próximos aos equipamentos e fortes mudanças no tempo, podem gerar falhas e prejudicar a análise dos elementos climáticos, como a temperatura do ar, umidade relativa do ar, precipitação e outros. Essas variáveis possuem variabilidade tanto no tempo quanto no espaço, tornando-se necessário um tratamento apurado para que a base de dados possa se tornar homogênea e possibilite a realização de análises climáticas mais precisas a longo prazo (LONGOBARDI; VILLANI, 2009).

A ocorrência desses problemas causa inconsistência na base de dados, fazendo com que ela deixe de ser homogênea e necessite do emprego de métodos e técnicas que possam auxiliar no preenchimento das falhas e na validação dos dados.

4.4. Preenchimento de Falhas em Séries Temporais Climatológicas

O preenchimento de séries falhadas consiste em gerar uma série sintética de dados através de métodos matemáticos, estatísticos e computacionais. Segundo Santos (2012), não existe uma forma padronizada para preencher essas falhas, o que permite o emprego de diferentes técnicas. Pode-se assim utilizar estações próximas e aplicar formas básicas como médias dos dados observados e geradores de dados sintéticos, ou utilizar métodos avançados como interpolação, regressões lineares e redes neurais artificiais, que se caracteriza por sua

capacidade de aprender através de um conjunto reduzido de exemplos, mapeando relações de entrada e saída, dando respostas para dados não conhecidos (CHIBANA et al., 2005; GONÇALVES et al., 2010).

4.5. Médias de Regiões Próximas

A utilização de dados de regiões que possuem características climáticas similares e que sejam consideravelmente próximas para preencher as falhas ocorridas em estações meteorológicas, são cada vez mais comum. Através desta seleção a média tenta estimar a variável faltante com qualidade (Figura 3).

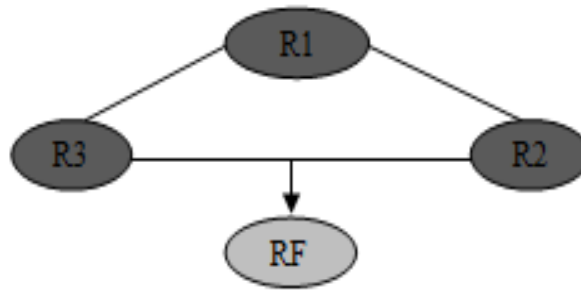


Figura 3: regiões utilizadas para preenchimento de dados de uma região faltante.

A Figura 3 demonstra três regiões estimadoras expressas por R1, R2 e R3, que são utilizadas para realizar o cálculo da media aritmética simples (Equação 1) e estimar um valor para preenchimento da falha ocorrida na estação que possui o dado faltante (RF).

$$MED = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

Em que x_i representa o conjunto de dados e n o número de dados amostrados.

4.6. Regressão Linear Múltipla

A regressão linear múltipla é uma técnica que tem por finalidade analisar ou relacionar uma variável dependente à diversas variáveis independentes também conhecidas como variáveis de previsão (FERNANDEZ, 2007; FONSECA; MARTINS; TOLEDO, 2012). Segundo Sousa, Dantas e Limeira (2007), a relação entre uma variável dependente Y e outras variáveis independentes $(X_1, X_2, X_3, \dots, X_k)$, é formulada pelo seguinte modelo linear (Equação 2):

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} + U_i \quad (2)$$

Na equação, $\alpha, \beta_1, \beta_2, \dots, \beta_k$ são os coeficientes de regressão determinados pelas observações do preditando e dos preditores, e U_i são os distúrbios aleatórios ou erros aleatórios independentes e com distribuição normal, de média zero e variância constante (SOUSA; DANTAS; LIMEIRA, 2007; LYRA; SOUZA; VIOLA, 2011).

O problema de resolução deste modelo linear com mais de duas variáveis está atrelado a estimação dos valores dos diversos parâmetros $\alpha, \beta_1, \beta_2, \dots, \beta_k$, o que pode ser realizado a partir de vários métodos. Contudo, um método comumente utilizado é o método dos mínimos quadrados que tem a finalidade de determinar os valores de α e β minimizando a soma dos desvios ao quadrado (FONSECA; MARTINS; TOLEDO, 2012).

A minimização da soma dos desvios ao quadrado é obtida através da resolução das n equações múltiplas lineares. No entanto, o grande número de constantes desconhecidas $\alpha, \beta_1, \beta_2, \dots, \beta_k$, pode tornar a utilização deste método muito trabalhosa devido o método de mínimos quadrados exigir a resolução de equações normais, no mesmo número que existem as constantes desconhecidas (FERNANDEZ, 2007).

Assim a estimação desses parâmetros, em um caso onde a variável dependente Y é postulada pela função entre duas variáveis independentes X_1 e X_2 , é dada pelo seguinte modelo de regressão (Equação 2.1).

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i \quad (2.1)$$

A partir de uma amostra de n observações das variáveis Y, X_1 e X_2 , estima-se os valores de a, b_1 e b_2 dos parâmetros α, β_1 e β_2 compondo assim o modelo estimador (Equação 2.2).

$$Y = a + b_1 X_1 + b_2 X_2 \quad (2.2)$$

Para realização do ajuste desta equação, deve-se resolver as três equações normais que irão determinar a, b_1 e b_2 (Equações 2.3, 2.4, 2.5).

$$\begin{cases} \Sigma Y = na + b_1 \Sigma X_1 + b_2 \Sigma X_2 & (2.3) \\ \Sigma Y X_1 = a \Sigma X_1 + b_1 \Sigma X_1^2 + b_2 \Sigma X_1 X_2 & (2.4) \\ \Sigma Y X_2 = a \Sigma X_2 + b_1 \Sigma X_1 X_2 + b_2 \Sigma X_2^2 & (2.5) \end{cases}$$

Onde tem-se os somatórios dos valores de X_1 e X_2 expressos por ΣX_1 e ΣX_2 , o produto de cada valor de X_1 pelo seu correspondente em X_2 dado por $\Sigma X_1 X_2$, a soma de todos os produtos de cada valor de X_1 pelo produto de Y expresso por $\Sigma Y X_1$ e a soma de todos os produtos de cada valor de X_2 pelo produto de Y expresso por $\Sigma Y X_2$.

4.7. Rede Neurais Artificiais (RNAs)

Sabe-se que o cérebro humano contém milhões de neurônios, e cada um desses neurônios se comunica com outros milhares paralelamente para processar informações.

Além disso a comunicação entre os neurônios é realizada através das sinapses por meio de breves impulsos elétricos, e esses impulsos passam por ligações chamadas dendritos. A soma de todos os impulsos que chegam até o neurônio representa uma energia que gera um grau de ativação, e partir desse processo é gerada uma resposta na forma de impulso que é transmitido para o próximo neurônio pelo axônio (MACHADO, 2005; AGGARWAL; SONG, 1998) (Figura 4). Assim acredita-se que é através deste processo que o cérebro consegue aprender e ser capaz de executar tarefas complexas.

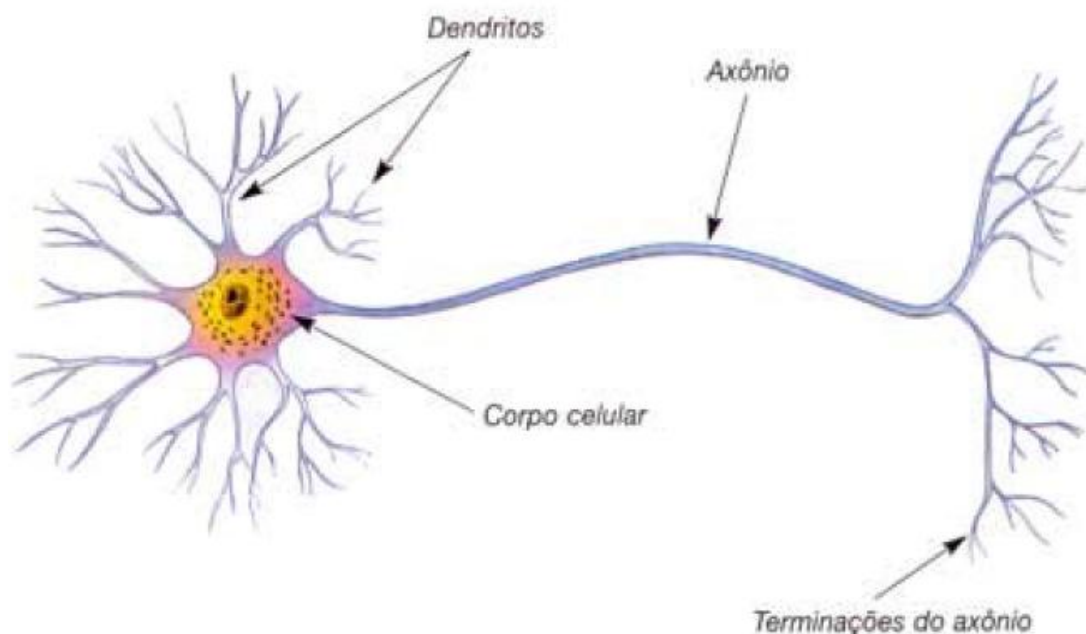


Figura 4: Estrutura de um neurônio biológico. Fonte: MACHADO, 2005.

Estas questões incentivaram o pioneiro trabalho de criação do primeiro modelo artificial de um neurônio biológico, desenvolvido através da pesquisa conjunta do psicólogo e neurofisiologista Warren McCulloch, e do matemático Walter Pitts. Em 1943 eles publicaram o trabalho "*A Logical Calculus of the Ideas Immmanent in Nervous Activity*", que concentrou-se em descrever um modelo artificial de um neurônio e apresentar suas capacidades computacionais (BRAGA; CARVALHO; LUDEMIR, 2012; ROCHA; CORTEZ; NEVES, 2008).

O trabalho de McCulloch e Pitts originou uma das principais ferramentas do ramo da Inteligência Artificial (AI), as Redes Neurais Artificiais (RNAs), que tentam emular os neurônios biológicos do cérebro humano (HEATON, 2010), através de um processamento massivamente paralelo e distribuído capaz de aprender por meio de exemplos e de generalizar a informação aprendida. As RNAs calculam funções matemáticas e têm uma propensão natural para armazenar o conhecimento proveniente da experiência e torná-lo útil. Desta forma, assemelhando ao cérebro humano (HAYKIN, 2001).

4.7.1. Estrutura Básica de um Neurônio Artificial

A primeira estrutura de um neurônio artificial proposta pelos pesquisadores McCulloch e Pitts, possuía n entradas que recebem valores $x_1, x_2, \dots, x_n$ e apenas uma camada de saída y (representado o axônio). Para representar o comportamento das sinapses os terminais de entrada dos neurônios possuem pesos acoplados $w_1, w_2, \dots, w_n$, cujos valores podem ser positivos ou negativos, dependendo se as sinapses correspondentes forem inibitórias ou excitatórias (BRAGA ; CARVALHO; LUDEMIR, 2012). A Figura 5 demonstra o modelo proposto McCulloch e Pitts.

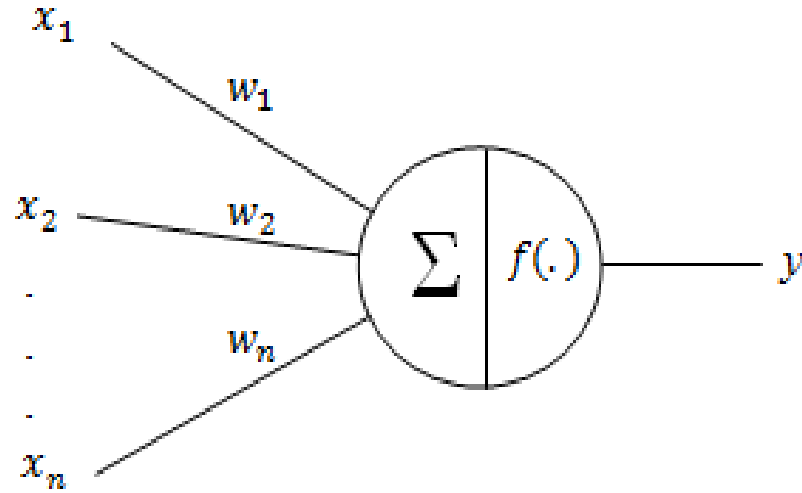


Figura 5: Estrutura de um neurônio Artificial. Fonte: BRAGA ; CARVALHO; LUDEMIR, 2012.

Entretanto, este modelo de RNA tinha uma grande limitação, pois seus pesos eram fixos e não ajustáveis, o que possibilitava apenas a resolução de funções linearmente separáveis (TORRES JUNIOR; MACHADO; SOUZA, 2005). Ou seja, o processamento de uma informação neste modelo é dado pela atuação de uma sinapse w_i em uma entrada x_i expressa pela multiplicação $w_i x_i$, onde os pesos determinam em que grau o neurônio deve considerar sinais de disparo ou ativação. O disparo do neurônio ocorre através da aplicação de uma função de ativação que ativa ou não a saída, dependendo do valor da soma ponderada das entradas pelos pesos (AGGARWAL; SONG, 1998) (Equação 3).

$$\sum_{i=1}^n x_i w_i \quad (3)$$

A função de ativação $f(.)$ é responsável por gerar a saída y do neurônio i , a partir do valor de ativação. O seguinte modelo utilizava uma função de ativação degrau deslocada do limiar (threshold) de ativação θ em relação à origem (Equação 3.1), o que impossibilitava a resolução de problemas de maior complexidade como os casos não lineares (ROCHA CORTEZ; NEVES, 2008).

$$f(.) = \begin{cases} 1 & \sum_{i=1}^n x_i w_i \geq \theta \\ 0 & \sum_{i=1}^n x_i w_i < \theta \end{cases} \quad (3.1)$$

Essas características fizeram desse modelo alvo de críticas por não conseguir resolver problemas mais complexos. Contudo, o trabalho de outros pesquisadores como Donald Hebb, Frank Rosenblatt, Minsky, Papert, John Hopfield, Rumelhart, Hitnton, Williams e outros, recriaram o paradigma das RNAs através da criação de metodologias de aprendizagem e treinamento, dando origem ao grande número de estruturas de RNAs utilizadas hoje, que podem ser empregadas tanto na solução de problemas de baixa complexidade como os casos

lineares, quanto na solução de problemas de alta complexidade como os casos não lineares (LIPPMANN, 1987).

4.7.2. Formas de Aprendizagem

Uma RNA possui a característica de aprender por meio de exemplos, estraiendo conhecimento de um determinado conjunto de dados. O conhecimento é adquirido a partir do processo pelo qual os parâmetros livres de uma rede neural são ajustados por meio de uma forma continuada de estímulo pelo ambiente externo. Este processo é definido como aprendizado que é obtido por meio do treinamento das RNAs.

Para tanto, existem muitos algoritmos de treinamentos que podem ser definidos em dois conjuntos, aprendizado supervisionado, que necessita de um supervisor, e o aprendizado não supervisionado, onde não há um supervisor (HAYKIN, 2001). Dentro do contexto de aprendizado supervisionado, um dos mais típicos processos é aprendizagem por correção de erros, que tem por finalidade ajustar os pesos das sinapses por meio do cálculo da diferença do valor esperado ou desejado pela saída de um neurônio i $y_{di}(t)$ no instante t , menos o valor predito pela RNA $y_{pi}(t)$ no instante t produzindo desta forma um erro $e(t)$ (Equação 4) (BRAGA; CARVALHO; LUDEMIR, 2012).

$$e(t) = y_{di}(t) - y_{pi}(t) \quad (4)$$

Este resultado $e(t)$ auxiliará no cálculo do ajuste aplicado aos pesos, que tem como objetivo minizar o erro da próxima iteração ($t+1$) (Equação 4.1).

$$w_{ij}(t + 1) = w_{ij}(t) + \Delta w_{ij}(t) \quad (4.1)$$

Na equação, $w_{ij}(t)$ representa o valor do peso no tempo t , e $\Delta w_{ij}(t)$ o valor do ajuste a ser aplicado ao peso, que de acordo com Yi e York (2004), é alcançado através da minimização da função de custo $E(t)$ dada pelo erro médio quadrático (Equação 4.2):

$$E(t) = \frac{1}{2} \sum_{i=1}^n e_i^2(t) \quad (4.2)$$

Na equação 4.2, n é o numero de amostras do treinamento, enquanto que $e_i^2(t)$ é o erro ao quadrado originado pela saída esperada $y_{di}(t)$ menos a predita pela RNA $y_{pi}(t)$. Esta minimização resulta na regra do delta ou regra de *Widrow-Hoff*, em que o valor do ajuste $\Delta w_{ij}(t)$ é originado da multiplicação de uma constante positiva que determina a taxa de aprendizado η , pelo produto do erro $e(t)$ e a entrada da sinapse x_i (Equação 4.3) (BRAGA; CARVALHO; LUDEMIR, 2012; MULLUHI; BAYOUMI; RAO, 1993).

$$\Delta w_{ij}(t) = \eta e_i(t) x_i(t) \quad (4.3)$$

4.7.3. Redes *Perceptrons* de Múltiplas Camadas (MLP)

A RNA do tipo MLP ou em inglês *multilayer perceptron*, pertencente a uma classe de RNAs conhecida como *feedforward*, é uma aproximadora universal de funções que inicialmente foi implementada para a resolução do problema não linear do XOR. Devido ao seu sucesso, ela vem sendo aplicada em diferentes problemas combinatórios e na solução de diferentes tarefas como processamento de informações, reconhecimento de padrões, previsões de tempo, problemas de classificação, processamento de imagens, previsões de atividades sísmicas e outros (SHAH; GHAZALI, 2011).

A estrutura de uma MLP é formada por um conjunto de neurônios dispostos em camadas, contendo uma camada de entrada, uma ou mais camadas intermediárias ou escondidas, e uma camada de saída. Cada um dos neurônios da camada de entrada está conectada a todos os neurônios da camada intermediária. Da mesma forma, cada neurônio da camada intermediária está conectado com todos os neurônios da camada de saída (MACHADO, 2005; ASADUZZAMAN et al., 2010). Esta estrutura pode ser vista na Figura 6.

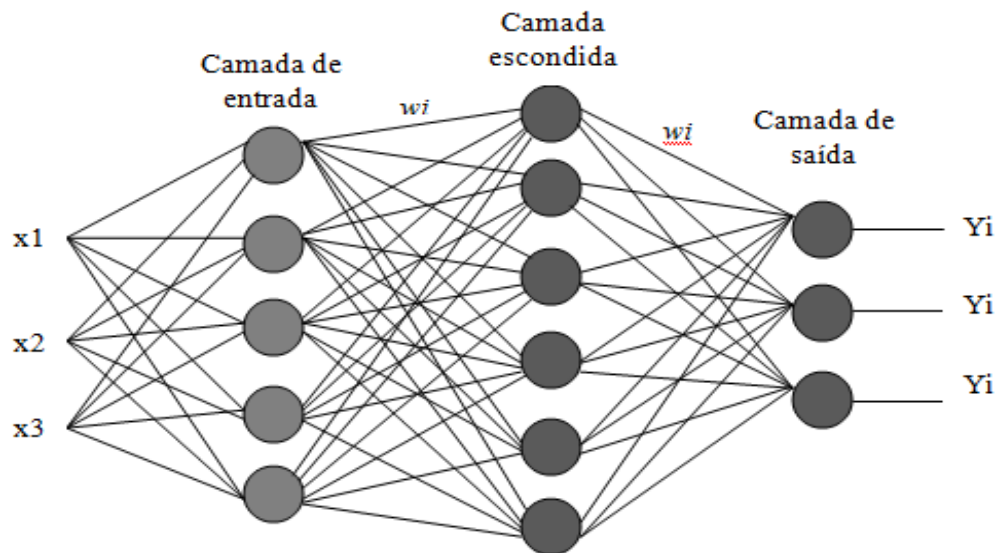


Figura 6: Estrutura de um RNA do tipo MLP.

Diferente do modelo proposto por McCulloch e Pitts onde o neurônio simplesmente comparavam o somatório ponderado das entradas pelos pesos a um limiar que gerava a saída binária ou seja zero ou um. A RNA do tipo MLP possui neurônios capazes de gerar qualquer saídas (HAYKIN, 2001; GUARNIERI, 2006). Desta forma, cada neurônio da RNA do tipo MLP deve ser interpretado como modelo de um neurônio não linear (Figura 7).

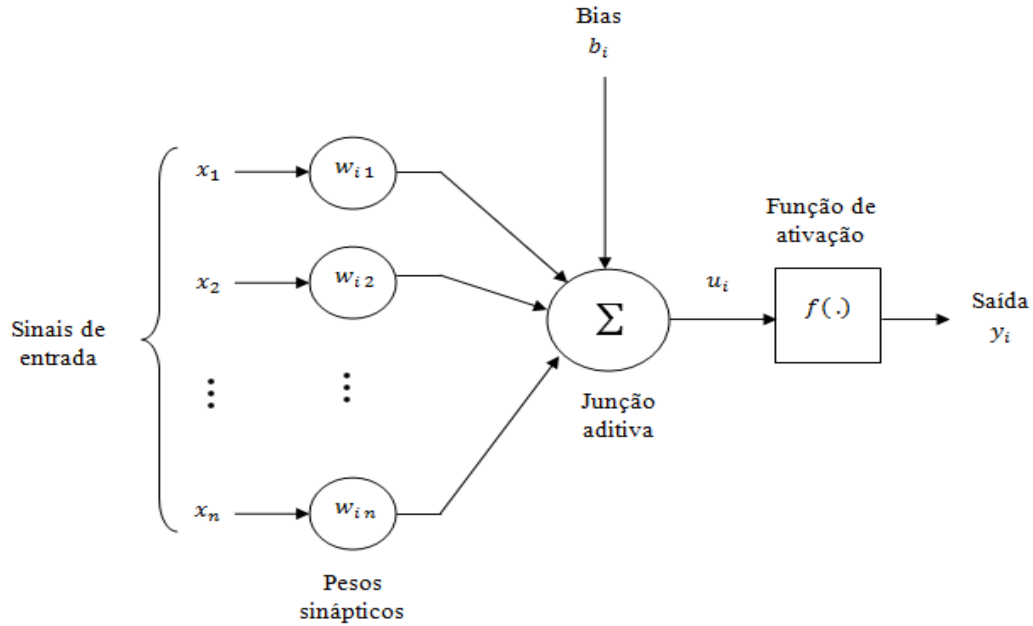


Figura 7:Modelo de neurônio não linear . Fonte adaptado de Haykin (2001).

Na Figura 7, u_i é a saída da combinação linear do somatório das entradas ponderadas pelos pesos w_i , mais o *bias* (b_i). A saída do neurônio y_i é originada pela aplicação de uma função $f(.)$ no resultado originado da combinação linear do somatório das entradas ponderadas pelos pesos w_i , mais o *bias* (b_i) que tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativo (HAYKIN, 2001). Desta forma, o neurônio i é expresso pelas (Equações 5, 5.1):

$$u_i = \sum_{i=1}^n w_{ij}x_i + b_i \quad (5)$$

$$y_i = f(u_i) \quad (5.1)$$

Existem muitos tipos de funções de ativação. Contudo, as mais empregadas são as funções degrau, linear, sigmóides (Equações 5.2, 5.3, 5.4, 5.5) (HAYKIN, 2001; BRAGA; CARVALHO; LUDEMIR, 2012; AGGARWAL; SONG, 1998).

a) Função de ativação degrau:

$$f(u_i) = \begin{cases} 1 & \text{se } u_i \geq 0 \\ 0 & \text{se } u_i < 0 \end{cases} \quad (5.2)$$

b) Função Linear :

$$f(u_i) = u_i \quad (5.3)$$

c) Funções Sigmóides:

1) Função Logística:

$$f(u_i) = \frac{1}{(1 + \exp(-u_i))} \quad (5.4)$$

2) Função Tangente Hiperbólica:

$$f(u_i) = \tanh\left(\frac{u_i}{2}\right) = \frac{1 - \exp(-u_i)}{1 + \exp(-u_i)} \quad (5.5)$$

O comportamento das funções está expresso na Figura 8, onde as entradas dos neurônios são representadas pelo eixo x e as saídas pelo eixo y, como demonstrado acima a função degrau pode originar saídas no intervalo entre 0 e 1 (AGGARWAL; SONG, 1998; LIPPMANN, 1987; SKODZIK et al., 2013). A função linear tem como saída mesmo valor de u_i , o que resulta num gradiente igual a 1, enquanto que as funções sigmóides, como a logística e a tangente hiperbólica produzem saídas entre dois intervalos 0 a 1 e -1 e 1.

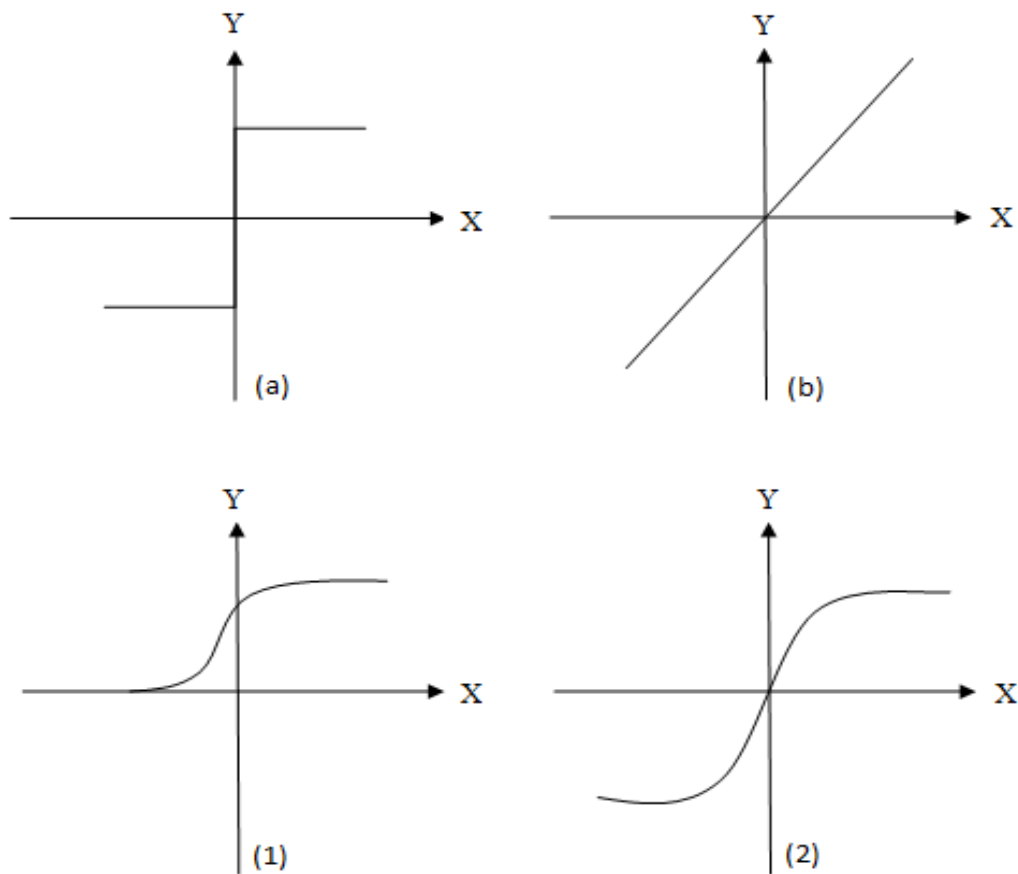


Figura 8: Gráficos dos comportamentos das funções de transferência. (a) Função Degrau, (b) Função Linear, (1) Função Logística, (2) Função Tangente Hiperbólica

No entanto, a saída final da RNA ou resultado é obtido após o processo de treinamento que faz o ajuste dos pesos através de um algoritmo. O principal algoritmo para o treinamento de uma RNA do tipo MLP, é o algoritmo *Back-propagation*.

4.7.3.1. *Back-propagation*

O algoritmo de *Back-Propagation* ou algoritmo de retropropagação é uma técnica de treinamento supervisionada que realiza o ajuste dos pesos através do cálculo da diferença entre o valor predito pela rede e o valor esperado.

Este algoritmo é executado em duas fases definidas como *forward* e *backward* (YI; YORK, 2004; MULLUHI; BAYOUMI; RAO, 1993; ALBARAKATI; KECMAN, 2013). Na fase *forward*, as entradas são submetidas a rede e propagam-se até a geração da saída, sem a alteração dos valores dos pesos. Na fase *backward*, calcula-se a diferença entre o valor predito $y_{pi}(t)$ pela rede e o esperado $y_{di}(t)$ (Equação 6) para que possa realizar o ajuste dos pesos a partir desta diferença propagando-a até a camada de entrada.

$$e(t) = y_{di}(t) - y_{pi}(t) \quad (6)$$

Este algoritmo é uma generalização da regra do delta ou como é chamado regra do delta generalizado. Dado um conjunto de treinamento, a cada passo t deste treinamento é calculado o sinal de erro para um determinado neurônio i da camada de saída de uma RNA do tipo MLP. Desta forma, a função de custo para uma RNA que tenha n neurônios em sua camada de saída é definida como o somatório de $\frac{1}{2} e_i^2(t)$ (GUARNIERI, 2006) (Equação 6.1).

$$E(t) = \frac{1}{2} \sum_{i=1}^n e_i^2(t) \quad (6.1)$$

A partir desta equação é calculado o ajuste dos pesos em direção contrária ao gradiente do erro, ou seja, aplicando-se a técnica do gradiente descendente. O ajuste de cada peso do neurônio da camada de saída é proporcional à derivada parcial de $E(t)$ em relação a cada um dos pesos w_{ij} da camada de saída da rede, que são calculados por meio da aplicação da regra da cadeia. Assim é obtida a equação do ajuste de pesos de um neurônio i qualquer da camada de saída (Equação 6.2) (BOUGHRARA; CHTOUROU; AMAR, 2012).

$$\frac{\partial E(t)}{\partial w_{ij}}(t) = -e_i(t) \cdot f'(u_i(t)) \cdot h_j(t) \quad (6.2)$$

Na equação acima, $e_i(t)$ é o erro originado pela saída da rede, $f'(u_i(t))$ é a derivada da função aplicada nos valores ponderados pelos pesos do neurônio da camada de saída e $h_j(t)$ é o valor de entrada de uma conexão do neurônio i . Logo, o ajuste dos pesos da camada de saída da RNA é conseguido por meio da regra do delta (Equação 6.3) (BOUGHRARA; CHTOUROU; AMAR, 2012; TASPINAR; SEYMAN, 2010).

$$\Delta w_{ij}(t) = -\eta \frac{\partial E(t)}{\partial w_{ij}} \quad (6.3)$$

Em que η é a taxa de aprendizado do algoritmo de retropropagação que deve ser mantida entre 0 a 1, onde irá determinar a velocidade de mudança dos pesos. Por isso, a mesma não deve ter um valor muito alto, pois pode acabar saltando ao redor da superfície do erro, e nem muito baixo porque o algoritmo pode demandar de um longo tempo para convergir (TASPINAR; SEYMAN, 2010). O sinal negativo indica a descida do gradiente no espaço de pesos, ou seja, busca uma direção para minimização do $E(t)$.

A equação acima pode ser reescrita como (Equação 6.4):

$$\Delta w_{ij}(t) = \eta \delta_i(t) \cdot h_j(t) \quad (6.4)$$

Em que o gradiente local $\delta_i(t)$ é originado do produto do erro $e_i(t)$ pela derivada da função de ativação $f'(u_i(t))$ (Equação 6.5) :

$$\delta_i(t) = e_i(t) \cdot f'(u_i(t)) \quad (6.5)$$

Esta medida de erro é executada de forma direta devido cada saída da RNA ser comparada com uma saída desejada. No entanto, a camada oculta de uma RNA do tipo MLP, mesmo não possuindo acesso direto ao cálculo do erro, seus neurônios compartilham a responsabilidade pelo erro cometido na saída da rede.

Segundo Haykin (2001), quando um neurônio i está localizado na camada oculta da rede, não existe uma resposta desejada para aquele neurônio. Desta maneira, um sinal de erro para o neurônio oculto deve ser determinado recursivamente, em termos dos sinais de erro de todos os neurônios aos quais o neurônio oculto está diretamente conectado. Pode-se então redefinir o gradiente local $\delta_i(t)$ para o neurônio oculto i (Equações 6.6, 6.7):

$$\delta_i(t) = - \frac{\partial E(t)}{\partial y_{pi}(t)} \frac{\partial y_{pi}(t)}{\partial u_i(t)} \quad (6.6)$$

$$= - \frac{\partial E(t)}{\partial y_{pi}(t)} f'_i(u_i(t)) \quad (6.7)$$

Consequentemente, deve-se dar sequência ao cálculo da derivada parcial $\partial E(t) / \partial y_{pi}(t)$. A partir de um determinado neurônio k da camada de saída, pode-se calcular o sinal de $E(t)$ (Equação 6.8):

$$E(t) = \frac{1}{2} \sum_{k=1}^n e_k^2(t) \quad (6.8)$$

Diferenciando a equação 6.7 em relação ao valor predito $y_{pi}(t)$ e aplicando a regra da cadeia para a derivada parcial de $\partial e_k(t)/\partial y_{pi}(t)$, chega-se a solução (Equação 6.9, 6.10).

$$\frac{\partial E(t)}{\partial y_{pi}(t)} = - \sum_k e_k(t) f'_k(u_k(t)) w_{ki}(t) \quad (6.9)$$

$$= - \sum_k \delta_k(t) w_{ki}(t) \quad (6.10)$$

Na equação (6.9), $e_k(t)$ é o erro originado pela saída da rede, $f'(u_k(t))$ é a derivada da função aplicada nos valores ponderados pelos pesos do neurônio e as entradas da camada seguinte ou de saída, e $w_{ki}(t)$ é o peso da ligação sináptica que conecta cada neurônio k da camada seguinte ao neurônio i . Na equação 6.10 reescreve-se a equação com o gradiente local $\delta_k(t)$.

Através desta equação, chega-se na fórmula de retropropagação para o gradiente local $\delta_i(t)$ (Equação 6.11).

$$\delta_i(t) = f'_i(u_i(t)) \sum_k \delta_k(t) w_{ki}(t) \quad (6.11)$$

No qual $f'_i(u_i(t))$ é a derivada da função aplicada nos valores ponderados pelos pesos das sinapses do neurônio e as entradas da camada oculta. No final, a atualização dos pesos da camada oculta é conseguido por meio da regra do delta (Equação 6.12)

$$\Delta w_{ij}(t) = \eta \delta_i(t) \cdot x_j(t) \quad (6.12)$$

Segundo Braga, Carvalho e Ludemir (2012), o treinamento da RNA do tipo MLP por meio do algoritmo *Back-propagation* padrão pode ser lento para várias aplicações, e seu desempenho ainda pode piorar para problemas mais complexos. Por isso, desde a criação deste algoritmo, várias modificações foram propostas tanto para acelerar seu tempo de treinamento quanto para melhorar seu desempenho. Por meio destas variações foram desenvolvidos os seguintes métodos de treinamento: *Quickprop*, *Levenberg-Marquardt*, *Back-propagation* com Newton, *Rprop* e outros. Contudo, as resoluções desses métodos não serão abordadas neste estudo.

4.7.4. Redes de Função de Base Radial (RBF)

As Redes de Função de Base Radial ou *Radial Basis Function* (RBF) pertencem a uma classe de redes neurais artificiais, que são consideradas, assim como as redes MLP, aproximadores universais de funções (CASTRO, 2001).

Contudo, existem várias diferenças entre esses dois modelos. Primeiramente, uma das principais diferenças, está relacionada a ativação dos neurônios, onde um neurônio de uma MLP é ativado por meio do produto escalar do vetor de entrada e do vetor de pesos do respectivo neurônio, enquanto que uma RBF é ativada por meio da função da distância entre seus vetores de entrada e seus centros (BRAGA; CARVALHO; LUDEMIR, 2012).

Em sua forma básica, a arquitetura de uma RNA do tipo RBF é composta de três camadas, contendo uma camada de entrada, uma camada intermediária e uma camada de saída (EBEID, 2011; SONG; KASABOV, 2004) (Figura 9).

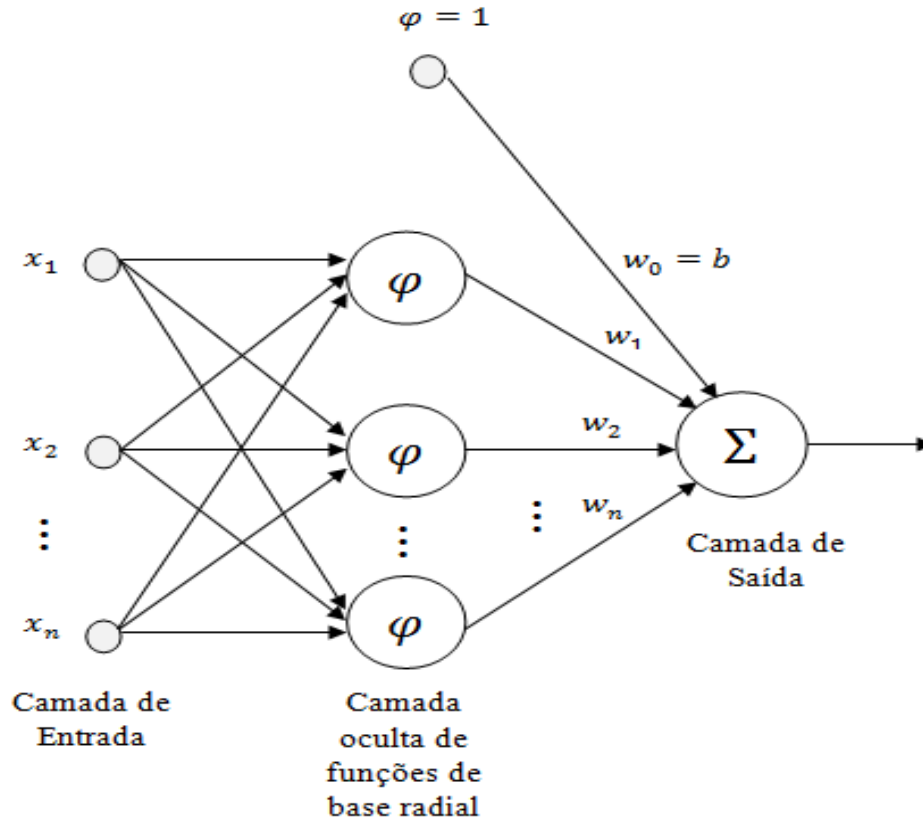


Figura 9: Estrutura de uma RNA do tipo RBF.

A camada intermediária de uma RBF utiliza funções de base radiais, e tem a finalidade de agrupar os dados de entrada em *clusters* e transformar um conjunto de padrões de entrada não linearmente separáveis em um conjunto de saídas linearmente separáveis, e a camada de saída tem a função de classificar os padrões recebidos da camada anterior através da combinação linear das saídas das funções (BRAGA; CARVALHO; LUDEMIR, 2012).

Dentro deste contexto, o processo de aprendizagem em uma rede RBF é equivalente a ajustar uma superfície não linear ao conjunto de dados em um espaço multidimensional, de acordo com algum critério estatístico, e a generalização equivale a usar esta superfície multidimensional para interpolar os dados de teste (CASTRO, 2001; BISHOP, 1997).

Segundo Cover (1965), a transformação não linear por uma transformação linear é justificada no teorema de separabilidade de padrões, onde é demonstrado que um problema complexo de classificação de padrões dispostos não linearmente em um espaço de alta dimensionalidade tem maior probabilidade de ser linearmente separável do que em um espaço de baixa dimensionalidade.

As funções de base radiais utilizadas nas camadas intermediárias das redes RBF são uma classe de funções em que sua resposta decresce ou cresce monotonicamente com o distanciamento de um ponto central. Para tanto, diferentes tipos de funções de base radial têm sido utilizadas em redes RBF, e as mais comuns são gaussiana, multiquadrática e *thin-plate-spline* (Equações 7, 7.1, 7.2) (MOREIRA et al., 2010, BRAGA; CARVALHO; LUDEMIR, 2012).

a) Função gaussiana:

$$f(u) = \exp\left(-\frac{v^2}{2\sigma^2}\right) \quad (7)$$

b) Função multiquadrática:

$$f(u) = \sqrt{(u^2 + \sigma^2)} \quad (7.1)$$

c) Função *thin-plate-spline*:

$$f(u) = v^2 \log(v) \quad (7.2)$$

Nas equações anteriores, $v = (\|x - t_i\|)$, é dado pela distância euclidiana, onde x é o valor de entrada da rede e t_i e σ correspondem ao centro e a largura da função radial. Dessa maneira a resolução de um determinado problema por uma RNA do tipo RBF consiste na procura por uma função que tenha a seguinte forma (Equação 7.3):

$$f(x) = \sum_{i=1}^k w_{ji} \varphi(\|x - t_i\|) \quad (7.3)$$

Na equação 7.3, w_{ji} são os pesos e $\varphi(\|x - t_i\|)$ é o conjunto de N funções de base radial e $\| \cdot \|$ é norma euclidiana. Assim a resposta originada por um neurônio da camada de saída y_j de uma RBF é dada pela (Equação 7.4) :

$$y_j = \sum_{i=1}^k w_{ji} \phi(\|x - t_i\|) + w_{j0} \quad (7.4)$$

Na equação 7.4, w_{ji} são os pesos de cada conexão sináptica, ϕ é a matriz de interpolação originada do conjunto de N funções de base radial aplicadas nas entradas x da RNA e dos seus respectivos centros t_i , e que o *bias* é representado por w_{j0} . Desta forma, as saídas y podem ser reescritas como (Equação 7.5):

$$\begin{bmatrix} \varphi(\|x_1 - t_1\|) & \dots & \varphi(\|x_1 - t_2\|) & \varphi(\|x_1 - t_N\|) \\ \varphi(\|x_2 - t_1\|) & \dots & \varphi(\|x_2 - t_2\|) & \varphi(\|x_2 - t_N\|) \\ \vdots & \ddots & \vdots & \vdots \\ \varphi(\|x_N - t_1\|) & \dots & \varphi(\|x_N - t_2\|) & \varphi(\|x_N - t_N\|) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (7.5)$$

Contudo, da mesma maneira que uma RNA do tipo MLP, o resultado S de uma RBF é conseguido através do ajuste dos pesos obtido por meio da minimização dos erros quadráticos

que comparam as saídas desejadas y_{di} com as saídas preditas y_{pi} (CERQUEIRA; ANDRADE; POPPI, 2001) (Equação 7.6).

$$S = \sum_{i=1}^n (y_{di} - y_{pi})^2 \quad (7.6)$$

Tal ajuste é conseguido por meio do processo de treinamento, que em uma RNA RBF pode acontecer de forma híbrida tendo dois estágios, o primeiro pode-se utilizar de técnicas não supervisionadas como o algoritmo *K-means-clustering*, RNAs do tipo SOM e outros, para seleção dos centros das funções de base radiais. O segundo estágio pode se utilizar de técnicas supervisionadas para definição dos pesos como a Regra do Delta, ou do Método de Decomposição em Valores Singulares (HU; HWANG, 2002).

4.8. Algoritmo Genético (AG)

A evolução é o ato pelo qual um indivíduo tende a se aprimorar com o objetivo de suprir as necessidades impostas pelo ambiente ao seu redor. Contudo, esse processo não tem o objetivo de obter a perfeição do indivíduo, e sim de torna-lo apto ao ambiente que foi imposto.

Tendo como referência esta situação, os Algoritmos Genéticos (AGs) fazem parte do ramo dos algoritmos evolucionários, e são um tipo de heurística de otimização global que se baseia na seleção natural, evolução biológica e na sobrevivência do indivíduo mais apto, baseando-se principalmente na teoria da evolução de Darwin (XUE et al., 2009; XIAN et al., 2011). O AG foi desenvolvido na década de 1970 por John Holland, através de um profundo estudo sobre a evolução das espécies, onde o mesmo propôs um modelo heurístico computacional que poderia gerar boas soluções para problemas extremamente difíceis que eram insolúveis computacionalmente até aquela época (LINDEN, 2012; BARCELLOS, 2000).

Com o passar dos anos, esse algoritmo se tornou extremamente popular devido a sua grande capacidade de gerar boas soluções para problemas complexos. Atualmente, ele é aplicado em muitas áreas, como descoberta de conhecimento, otimização combinatória, aprendizagem de máquina, processamento de imagens, mineração de dados, controle adaptativo e vida artificial, e outros (WANKHADE; RANE; THOOL, 2013).

Diferente dos algoritmos de busca e otimização convencionais que trabalham de forma sequencial avaliando a cada instante uma possível solução, os AGs trabalham com conjunto de soluções originadas das populações de indivíduos (cromossomos) que são submetidas aos operadores genéticos (ALVES et al., 2011). Estes operadores não fazem uso da derivada de uma função, mas utilizam a caracterização da qualidade de cada indivíduo como solução do problema, e geram um processo de evolução natural dos indivíduos que podem ter como resultado uma boa solução para o problema (BRAGA; CARVALHO; LUDEMIR, 2012).

Desta forma, um AG não está preso a métodos como gradiente de (*hill climbing*) que seguem a derivada de uma função para encontrar o máximo de uma função, ficando facilmente estagnados em um ponto de máximo local, em que os métodos de busca e otimização mais comuns geralmente se destacam, tendo um espaço de busca onde se encontram as possíveis soluções para o problema e pela função de avaliação ou de custo que avalia por meio de nota cada possível solução presente no espaço de busca (Figura 10).

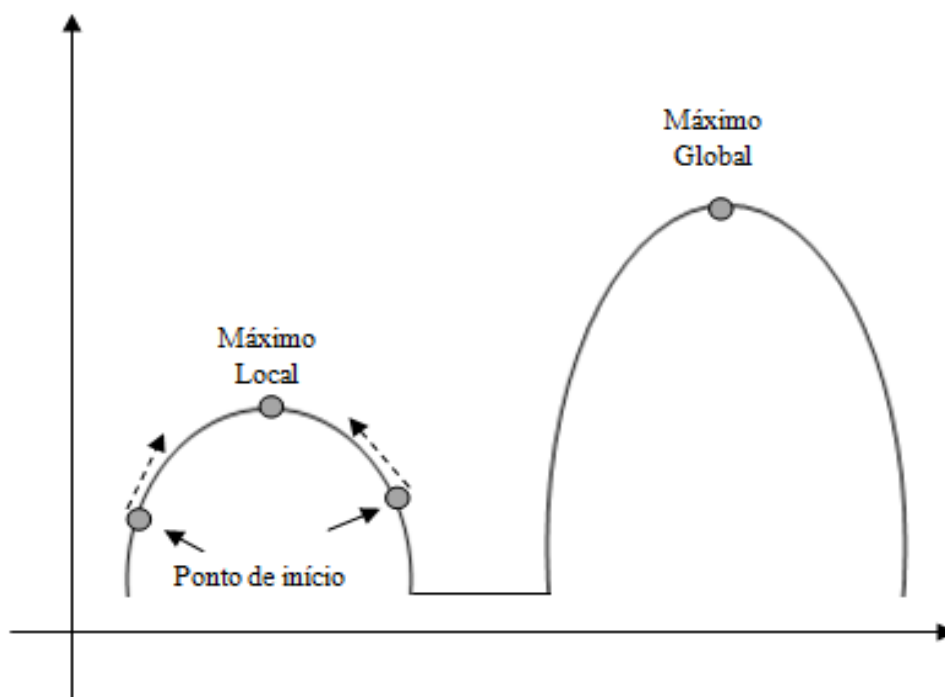


Figura 10: Exemplo de uma função hipotética de máximo local e global. Fonte adaptada de Lindel(2012).

No entanto, (LINDEN, 2012) destaca que apesar dos AGs tentarem chegar a um ponto ótimo (máximo ou mínimo), o mesmo não garante a obtenção deste resultado em todas suas execuções e que só devem ser usados na resolução de problemas complexos que demandem de um longo tempo de execução (Problemas NP - completos), afirmando ainda que se houver um algoritmo exato para resolução do problema. Logo, deve-se fazer uso dele.

Partindo destes pontos (GOLDBERG, 1989), definiu quatro aspectos principais que diferem os AGs dos métodos tradicionais de busca e otimização, sendo eles:

- Utilizam-se de uma codificação do conjunto de parâmetros, e não dos próprios parâmetros.
- Realiza a busca de uma possível solução em um conjunto de candidatos (População) e não em apenas um.
- Não fazem uso de conhecimento específico do problema, utilizando-se apenas de uma forma de avaliação do indivíduo.
- Fazem uso de regras de transição probabilísticas e não de determinísticas

Através desses aspectos pode se definir o foco principal de um AG, que é evoluir uma população até que possa chegar em uma possível solução ótima. Isto ocorre por meio de sua execução, onde gera-se uma população, avalia seus indivíduos quanto a sua aptidão para resolver o problema, partindo da aptidão seleciona os indivíduos mais aptos e cria uma nova população por meio dos operadores de recombinação e mutação. Este ciclo chamado de geração é repetido até que o critério de otimização seja alcançado, e no final o indivíduo que estiver apto para solucionar o problema é tomado como solução (ALVES, 2006; COHOON; PARIS, 1987) (Figura 11).

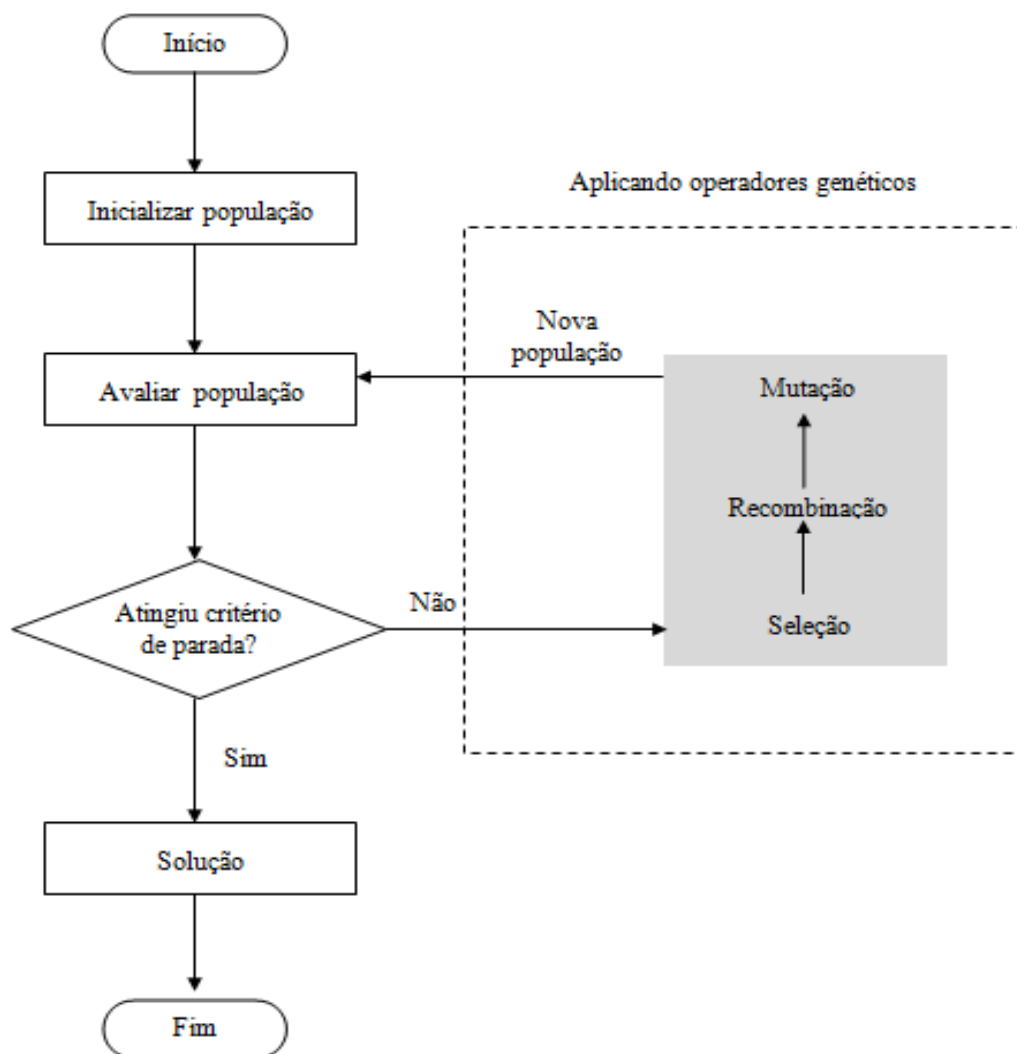


Figura 11: Estrutura básica do Funcionamento de um AG.

Basicamente, este é o processo pelo qual um algoritmo genético chega a uma solução válida. Para tanto, deve-se conhecer cada fase deste processo, sabendo-se que o primeiro passo de um AG está na geração de uma população. Entretanto, cada indivíduo desta população deve ser representado de maneira que possa traduzir informações do problema de forma viável para tratamento pelo computador. A representação fundamental para um AG é a cromossômica, sendo que este tipo de representação permite que cada pedaço indivisível do cromossomo, ou seja o gene, possa ser representado por números binários (0 ou 1), inteiros ou reais (SRINIVAS; PATNAIK, 1994; RODRIGUES et al., 2004).

Na representação binária que foi adotada no presente estudo, o cromossomo é representado por uma sequência de *bits* e cada gene é composto por somente um bit, que pode conter informações sobre valores e características do problema.(Figura 12)

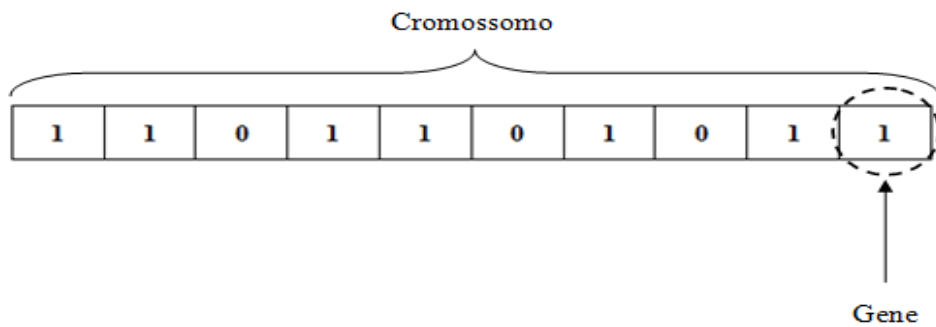


Figura 12: Esquema de um cromossomo binário.

Durante o processo de evolução, cada cromossomo da população é avaliado por meio de uma função, onde é atribuída uma nota a cada cromossomo que reflete sua aptidão para resolver o problema em questão. Por meio das notas obtidas por cada cromossomo, é executado um método de módulo de população e depois o método de seleção, que tem a finalidade de escolher os indivíduos que participarão do processo de reprodução (BRAGA; CARVALHO; LUDEMIR, 2012).

Um dos métodos de módulo de população mais comum é o elitismo. Ele previne que aconteça a eliminação dos melhores indivíduos em favor de indivíduos piores, e isto acontece porque o mesmo define que os melhores indivíduos de uma geração não devem morrer com a geração atual, e sim serem propagados para próxima geração. Desta forma, é garantido que o desempenho do AG sempre cresça com o decorrer das gerações (LINDEN, 2012; BRAGA; CARVALHO; LUDEMIR, 2012).

Entretanto, o método de seleção deve ser realizado com cuidado, privilegiando indivíduos com maiores valores de aptidão, e não desprezando os indivíduos com baixos valores. Isto evita que características genéticas contidas apenas em um indivíduo de baixa aptidão sejam perdidas e que se crie uma população sem diversidade com indivíduos semelhantes. Para tanto, existem vários métodos de seleção, onde os mais comuns são roleta e torneio.

O método da roleta toma por base o valor da aptidão de cada indivíduo, e cria uma roleta artificial dividida por partes diferentes que são proporcionais às suas avaliações. Desta forma, é criado um fator probabilístico em cada sorteio da roleta onde cada indivíduo tem uma certa probabilidade de ser selecionado (CAVALCANTE JUNIOR, 2013; BARCELLOS, 2000) (Figura 13).

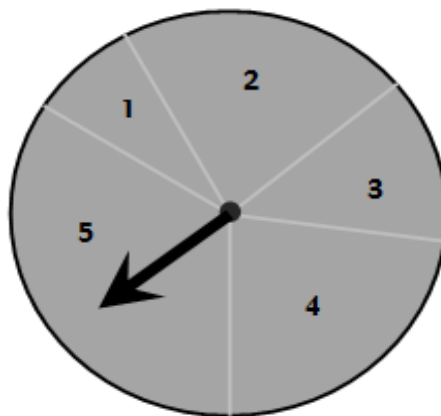


Figura 13: Seleção utilizando o método da roleta. Fonte adaptada de BRAGA; CARVALHO; LUDEMIR, 2012.

Com isto, consegue-se selecionar os indivíduos mais aptos sem deixar de selecionar os indivíduos menos aptos que podem conter características importantes para resolução do problema, evitando desta forma que aconteça o caso de convergência genética onde os indivíduos menos aptos nunca são escolhidos como possíveis pais, o que pode dificultar a convergência para uma solução.

Já o método de torneio seleciona um certo número de indivíduos da população e os fazem entrar em uma competição pelo direito de ser pai, usando como arma o seu valor de aptidão. Neste método, existe o parâmetro denominado "tamanho do torneio" que define quantos indivíduos serão selecionados aleatoriamente dentro da população para competir. Uma vez definido os competidores, o que tiver o maior valor de aptidão é selecionado como pai para aplicação do operador genético de cruzamento (*crossover*) (LINDEN, 2012).

O operador genético de cruzamento (*crossover*) é responsável por selecionar aleatoriamente partes genéticas de dois pais, e recombiná-las gerando um novo indivíduo (filho) permitindo que características sejam herdadas para as próximas gerações (COHOON; PARIS, 1987; THIRUGNANAM et al., 2014; RODRIGUES et al., 2004). Ressalta-se que, existem muitas maneiras de utilizar esse operador, onde as mais comuns são a de um ponto, dois pontos, multipontos e uniforme.

No presente estudo foi utilizado somente o *crossover* de dois pontos. Seu funcionamento consiste na realização de um sorteio aleatório de dois pontos de corte que irão originar os dois filhos. O primeiro filho será composto pelos genes do início até o ponto de corte do primeiro pai, pelos genes do núcleo do segundo pai e do segundo ponto de corte até os genes finais do primeiro pai, da mesma forma que ocorrido na geração do primeiro filho. O segundo é composto pelos genes do início até o ponto de corte do segundo pai, pelos genes do núcleo do primeiro pai e do segundo ponto de corte até os genes finais do segundo pai. Isto pode ser visualizado na Figura 14.

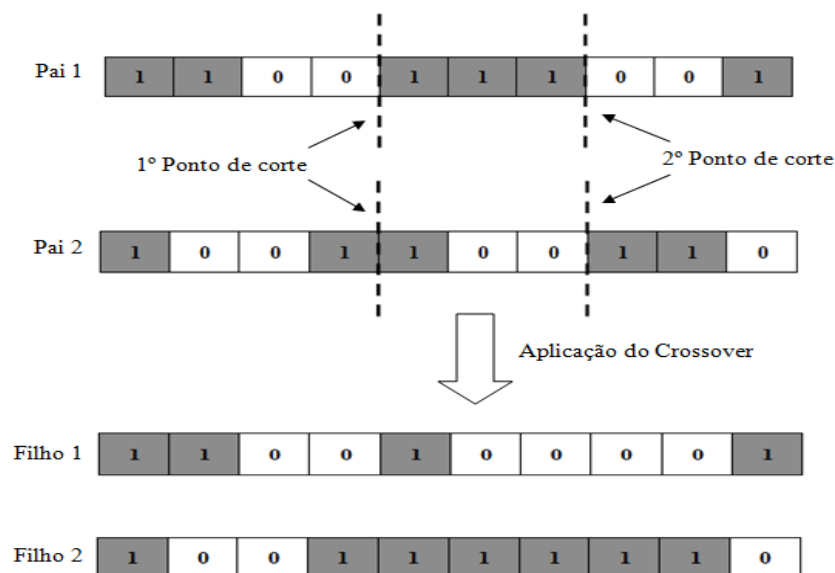


Figura 14: Exemplo de *crossover* de 2 pontos.

Por fim, após selecionar os indivíduos da população e recombiná-los pelos operadores de *crossover*, o AG impõe aos indivíduos dessa população um mecanismo de variação, a fim de tornar mais diverso o universo de busca e, conseqüentemente, menor a probabilidade de convergência precoce para ótimos locais (ALVES, 2006). Esse mecanismo de variação é a

mutação, que tem a finalidade de alterar o valor de um gene de um indivíduo. Contudo, para preservar a instabilidade do algoritmo genético, é aconselhável manter uma porcentagem de mutação pequena.

O processo de mutação pode ser observado na Figura 15, onde é realizada uma demonstração de como os valores dos *bits* são alterados.

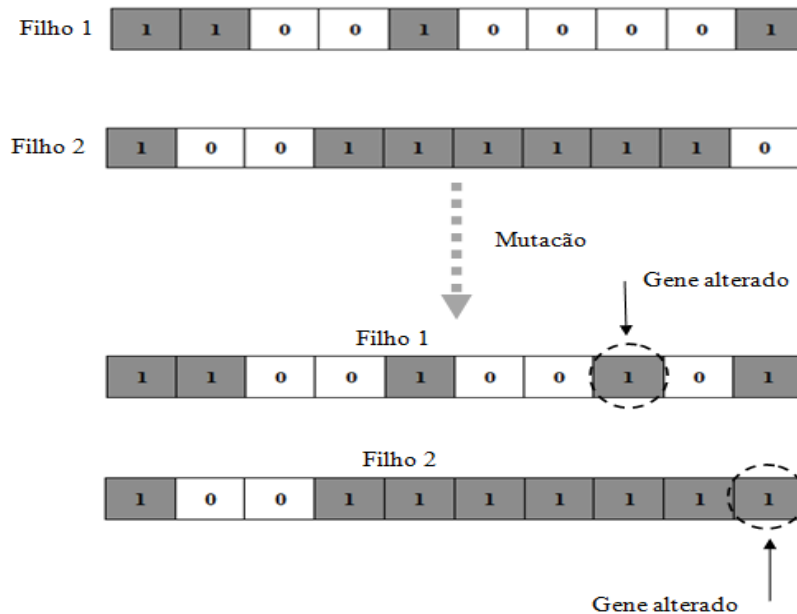


Figura 15: Exemplo da ocorrência do processo de mutação.

Após esses passos, uma nova geração da população é formada e deverá ser avaliada novamente, reiniciando o ciclo do algoritmo genético. Essa sequência continuará até que um dos critérios de parada seja alcançado (VENTURA, 2012).

5. MATERIAIS E MÉTODOS

A qualidade dos dados meteorológicos é uma questão preocupante. Por isso, a realização de um preenchimento adequado dos dados faltosos é necessário. Esse preenchimento pode ser feito de formas variadas devido não haver um padrão. Para tanto, este tópico tem por objetivo descrever as metodologias adotadas no presente estudo para preencher as falhas ocorridas nas estações meteorológicas, destacando fonte dos dados, metodologia proposta para realizar o preenchimento dos dados com falhas e como foi feita a análise dos dados gerados.

5.1. Base de Dados

O presente estudo utilizou uma série de dados meteorológicos registrados por seis estações pertencentes ao Instituto Nacional de Meteorologia (INMET). Essas estações são oriundas de regiões do estado do Rio de Janeiro, pertencentes a seis municípios que se destacam pela diversidade climática, econômica, produção agrícola e extração de minérios (petróleo) (PERES; MOREIRA, 2007).

As estações estão localizadas nos municípios de Paty do Alferes (região Centro Sul Fluminense), Campos dos Goytacazes (região Norte Fluminense), Cordeiro (região Serrana), Itaperuna (região Noroeste Fluminense), Resende (região do Médio Paraíba), e Rio de Janeiro pertencente a região Metropolitana do estado. Os municípios e as respectivas coordenadas geográficas das estações podem ser observadas na Figura 16 e na Tabela 1.

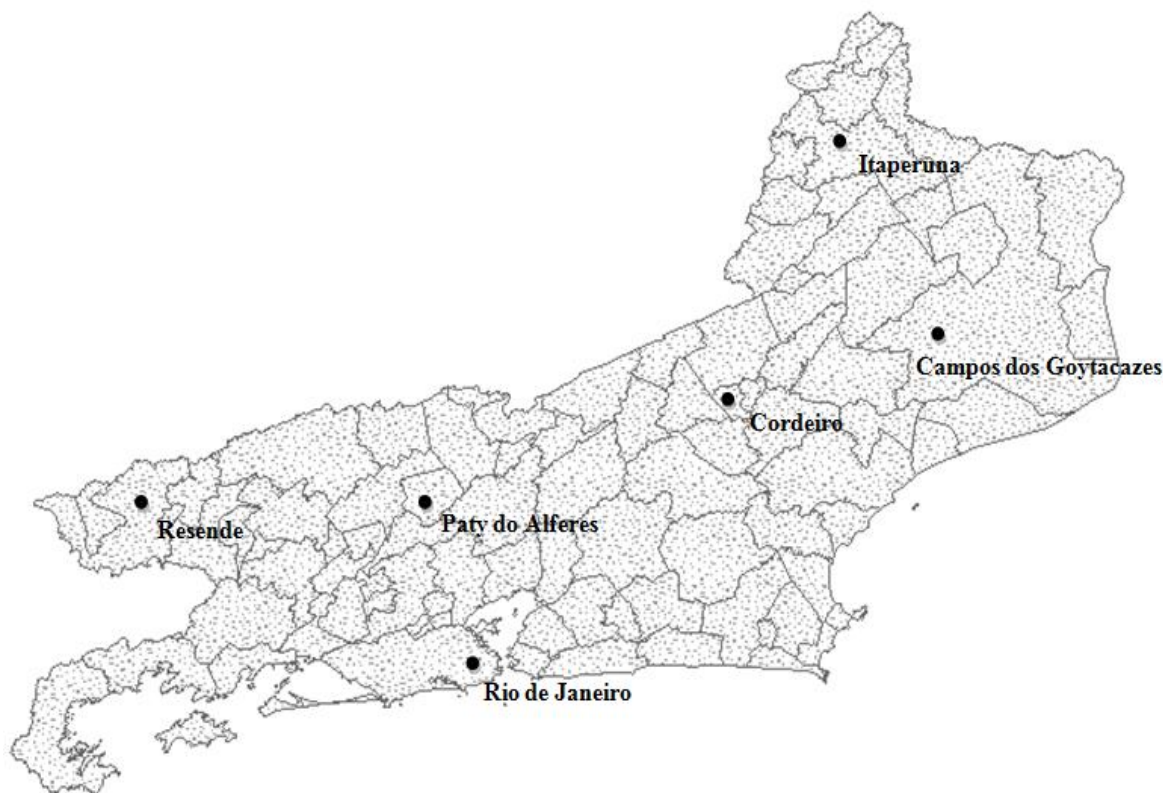


Figura 16: Mapa do estado do Rio de Janeiro com o conjunto de regiões utilizadas.

Tabela 1: Estações e suas coordenadas geográficas.

<i>Estações</i>	<i>Latitude</i>	<i>Longitude</i>	<i>Altitude (m)</i>
Campos	-21,74 °S	-41,33 °W	11,20
Cordeiro	-22,02 °S	-42,36 °W	505,92
Itaperuna	-21,2 °S	-41,9 °W	123,59
Rio de Janeiro	-22,89 °S	-43,18 °W	11,10
Paty do Alferes	-22,35 °S	-43,41 °W	507
Resende	-22,45 °S	-44,44 °W	439,89

Os dados utilizados são médias mensais registradas no período de 31/05/2002 a 31/12/2012. Estes dados são medidas de evaporação, temperatura máxima do ar e umidade relativa do ar, que foram escolhidos por possuírem uma série histórica de maior estabilidade que os demais dados registrados.

5.2. Modelo Proposto para Preenchimento de Falhas

A proposta principal para preenchimento das falhas das seis estações meteorológicas escolhidas foi a aplicação de um modelo híbrido formado por uma RNA do tipo MLP treinada com auxílio do AG. Esse modelo foi comparado com outras técnicas de preenchimento de dados, a fim de avaliar qual modelo possuía maior adequação para solucionar o problema das falhas de cada variável pertencentes as regiões utilizadas. A seguinte metodologia consistiu na realização de quatro fases distintas, que vão desde a preparação das informações através de um pré-processamento dos dados, até a avaliação dos resultados obtidos pelos métodos de preenchimento de falhas adotados.

Na primeira fase desta metodologia foi aplicado um pré-processamento nos dados, que tem como finalidade encontrar ruídos que podem prejudicar a pesquisa (SOMASUNDARAM; NEDUNCHEZHIAN, 2013). Esta fase foi responsável por preparar os dados que foram aplicados nas técnicas de preenchimento, onde foram identificados os dados faltantes e aplicado um processo de triagem, junto à normalização nos dados das estações.

A segunda fase definiu o conjunto de variáveis que foram empregadas no treinamento das RNAs e na estimativa dos dados faltantes. Contudo, esse conjunto de variáveis não foi utilizado na técnica de inteligência computacional híbrida (MLP-AG), pois a mesma define seu próprio conjunto de variáveis de treino e estimação.

A terceira fase foi responsável por tomar o conjunto de dados definidos pela segunda fase e os aplicar nas técnicas de preenchimento como média, regressão linear múltipla e RNAs (MLP, RBF).

Já a última fase teve a finalidade de comparar e avaliar todas as técnicas aplicadas no presente estudo. As fases da metodologia podem ser visualizadas na Figura 17.

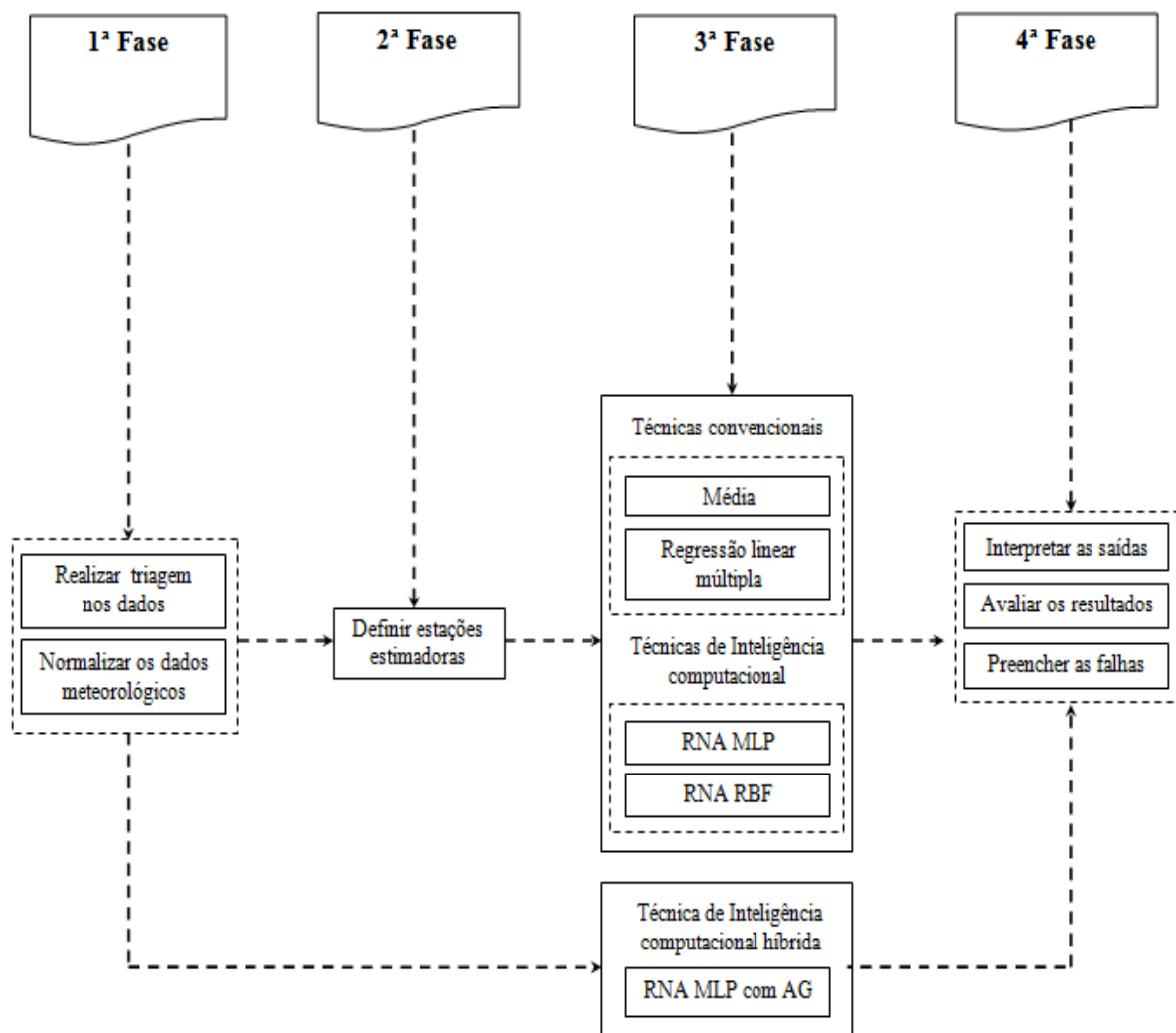


Figura 17: modelo proposto para preenchimento de falhas.

5.2.1. Levantamento Quantitativo de Falhas

Como foi apresentado anteriormente, as estações meteorológicas estão sujeitas as falhas originadas pelo mau manuseio, falha de equipamentos e até influências externas como, animais e mudanças drásticas de clima (VENTURA, 2012).

Com a finalidade de analisar quantitativamente a qualidade dos registros e identificar os dados que necessitam de um preenchimento, foi feito o levantamento da quantidade de falhas ocorridas no período de 31/05/2002 a 31/12/2012 de cada variável pertencente as regiões que foram utilizadas (Tabela 2).

Tabela 2: Quantidade de falhas ocorridas no período de 31/05/2002 a 31/12/2012.

<i>Estações</i>	<i>Numero de meses que ocorreram falhas</i>		
	<i>Evap</i>	<i>Temp.Max.</i>	<i>Umid. Relat.</i>
Campos	3	7	5
Cordeiro	3	4	23
Itaperuna	0	4	7
Rio de Janeiro	3	7	7
Paty do Alferes	1	7	11
Resende	10	17	27

- Evap = Evaporação
- Temp.Max= Temperatura Máxima do Ar
- Umid.Relat = Umidade Relativa do Ar

A partir do levantamento do número de falhas, foi realizado o cálculo percentual em cima do total de dados que corresponde a 128 meses, para verificar a influência dos dados falhados no total de registros (Figura 18).

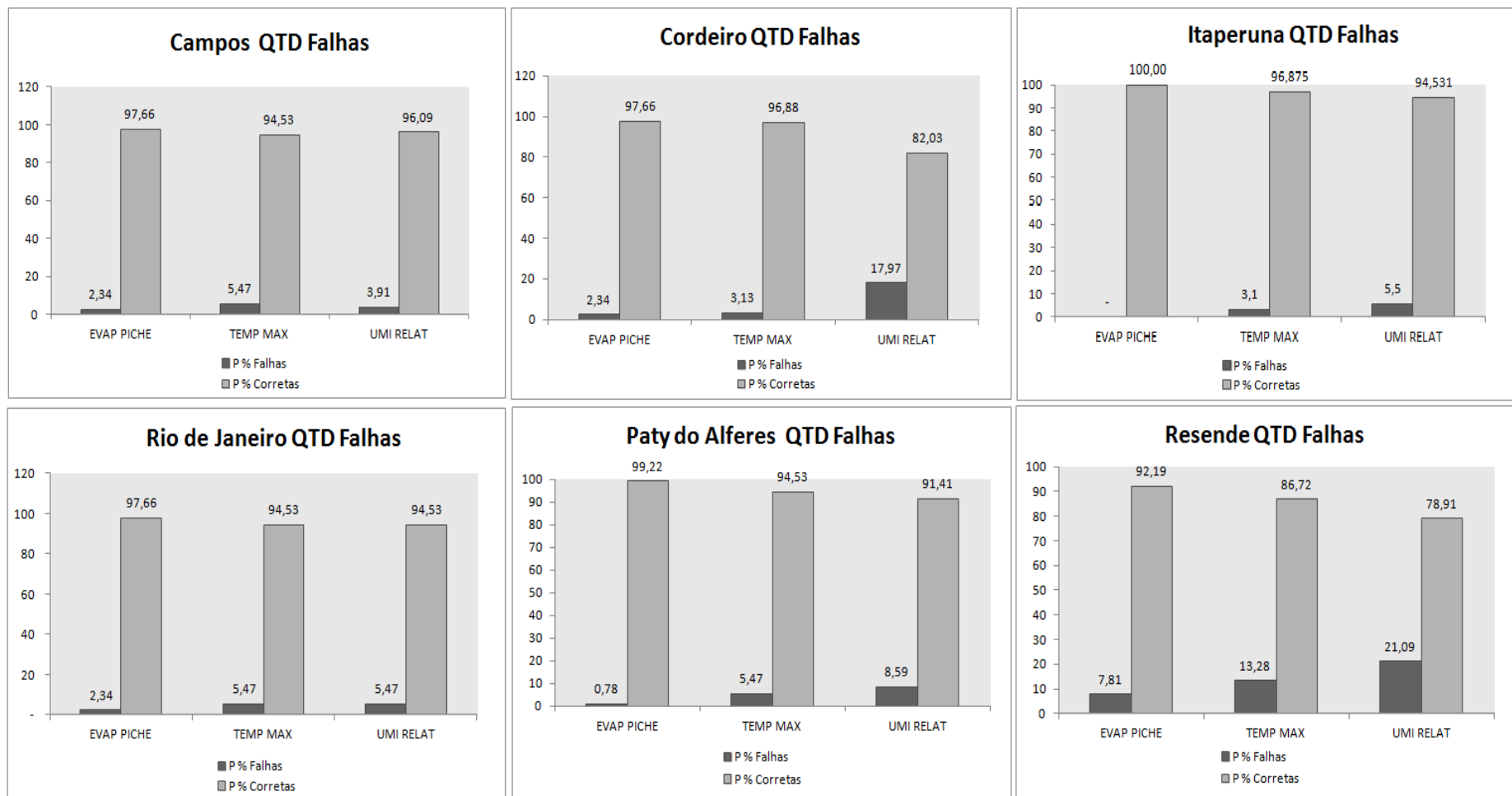


Figura 18: Percentual de dados registrados corretamente e porcentagem de falhas ocorridas em cada dado.

5.2.2. Triagem de Dados

Primeiramente, os dados coletados pelas estações foram inseridos em um programa de planilha eletrônica, seguido pelos dados de evaporação, temperatura máxima do ar e umidade relativa do ar, e em seguida passaram por um processo de identificação de falhas e inconsistências.

Após esta fase foram retirados os dados faltantes de todas as estações, ou seja, se uma estação *x* do conjunto de estações não tivesse o registro da média mensal da umidade relativa do ar do período de 30/04/2008, o mesmo deveria ser retirado de todas as outras estações. Isso garantiu a criação de uma série homogênea fazendo com que todas as estações tivessem o mesmo número de dados e os mesmos meses registrados (Figura 19).

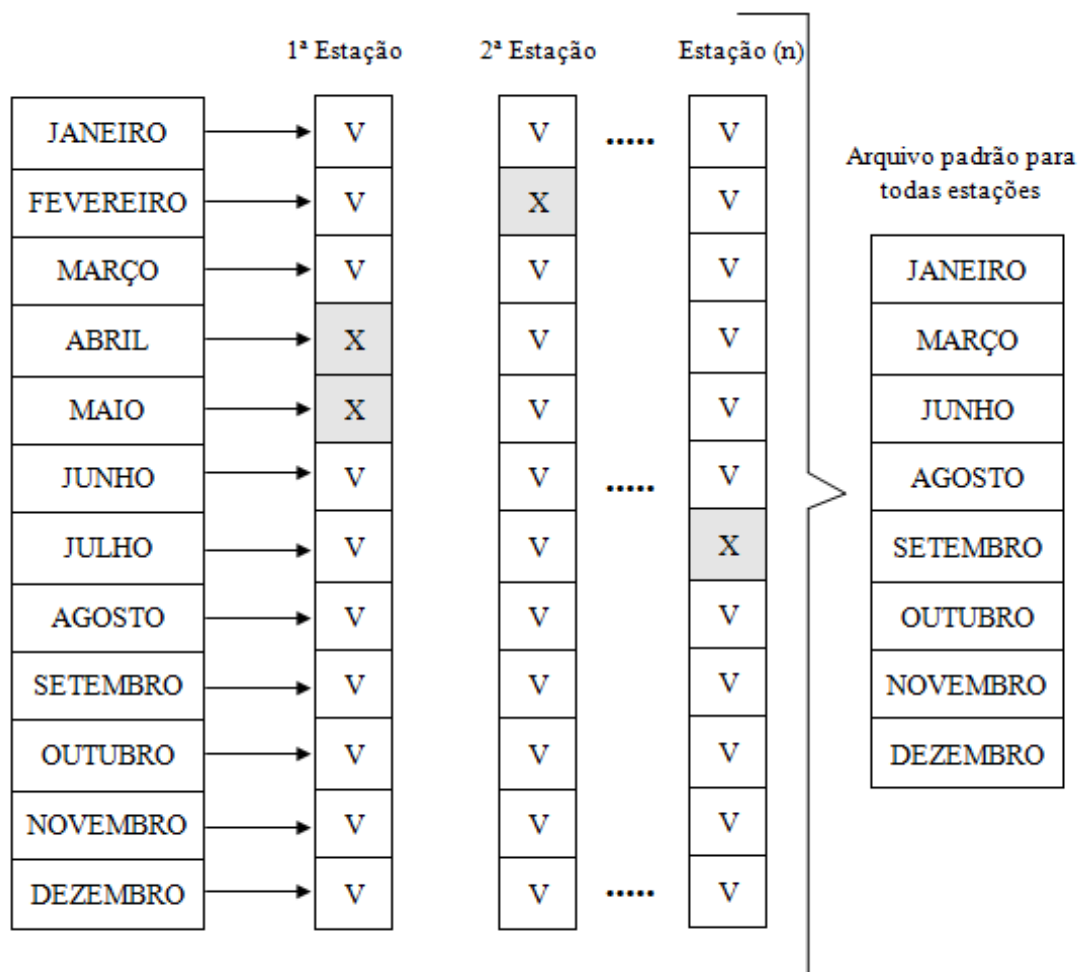


Figura 19: Exemplo do processo de triagem aplicado aos dados.

Após a triagem, o conjunto de dados passou pela etapa de normalização que foi responsável por preparar os dados para fase de determinação de variáveis que serão utilizadas na estimação dos dados meteorológicos.

5.2.3. Normalização

Segundo Rocha, Cortez e Neves (2008), o procedimento de normalização também pode ser designado por escalonamento, que é responsável por realizar uma transformação nos dados de forma que o algoritmo de treinamento ou ajuste possa obter melhores resultados. Esse procedimento codifica todos atributos em intervalos semelhantes fazendo com que todos dados tenham a mesma importância.

Esta semelhança de valores é obtida pela mudança da escala original dos valores para um intervalo, podendo ser entre 0 a 1, -1 a 1 ou outros. Isso aumenta a compatibilidade dos dados com as funções de transferência (MACHADO, 2005; NETO; PELLI, 2004).

Para tanto, neste estudo foi adotada a metodologia de normalização das variáveis que altera a escala real dos valores de entrada para o intervalo entre 0 a 1 (Equação 8).

$$x_j^{norm} = \frac{x_j - x^{min}}{x^{max} - x^{min}} \quad (8)$$

- Onde x_j^{norm} é a variável normalizada.
- x_j é a variável na posição j.
- x^{min} valor mínimo observado entre as variáveis.
- x^{max} valor máximo observado entre as variáveis.

5.2.4. Seleção das Estações Utilizadas na Estimativa dos Dados

Um dos grandes problemas em se utilizar dados de outras estações para preencher as falhas de uma estação x é como realizar a escolha das estações estimadoras. Existem varias maneiras de defini-las, podendo ser por regiões mais próximas, métodos estatísticos como regressão passo a passo (*stepwise linear regression*), coeficiente de correlação entre seus dados, ou outros (FERNANDEZ, 2007; GUARNIERI, 2006).

No entanto, Serrano et al. (2010) destaca que não existe um critério geral para seleção das estações adequadas para reconstrução dos dados com falhas em vista disso, neste estudo, num primeiro momento foram testadas e analisadas possíveis combinações onde foi definido a utilização de três estações para estimativa dos dados, adotando em parte o critério de escolha baseado na proximidade das regiões com a estação a ser preenchida. Contudo, não foram utilizadas todas as três regiões mais próximas de uma estação que tem variáveis a ser preenchida, pois a intenção era comparar a capacidade de escolha de regiões da técnica híbrida (MLP-AG) com a escolha definida para as demais técnicas, e verificar se a mesma conseguiria chegar em uma melhor escolha, ou na escolha ótima.

Para tanto, foram definidas as seguintes regiões estimadoras (Tabela 3) a serem aplicadas nas técnicas de média, regressão linear múltipla, MLP, RBF.

Tabela 3: Conjunto de estações utilizadas na estimativa ou predição de dados.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>
Campos dos Goytacazes	Cordeiro	Itaperuna	Rio de janeiro
Cordeiro	Rio de janeiro	Itaperuna	Campos dos Goytacazes
Itaperuna	Rio de janeiro	Cordeiro	Campos dos Goytacazes
Rio de Janeiro	Paty do Alferes	Resende	Cordeiro
Paty do Alferes	Rio de janeiro	Resende	Cordeiro
Resende	Rio de Janeiro	Paty do Alferes	Cordeiro

5.2.5. Seleção do Conjunto de Dados de Treino e Validação

Após as fases de preparação dos dados e seleção das estações estimadoras, o conjunto de informações de cada variável foi dividido em duas partes, sendo 75% para treino/ajuste e 25% para validação.

Esse critério de divisão de dados foi aplicado a fim de realizar uma validação cruzada, que é obtida pela comparação entre os valores observados e estimados pelas técnicas de preenchimento (LEENAERS; OHX; BURROUGH, 1990). Segundo Robinson e Metternicht (2006), o processo de validação cruzada é obtido por meio da estimativa de um conjunto de dados amostrais, geralmente estima-se um valor de cada vez. Em seguida, calcula-se a diferença entre o valor real e valor estimado para cada dado do conjunto amostral.

5.2.6. Aplicação das Técnicas de Preenchimento de Falhas

Neste estudo foram empregadas as técnicas de preenchimento de falhas (média, regressão linear múltipla, RNAs MLP, RBF e MLP-AG). A implementação computacional de todos modelos foi realizada com a utilização do software MATLAB, junto a seu conjunto de ferramentas.

A técnica de preenchimento baseada na média dos dados das estações utilizadas na estimação foi aplicada apenas no conjunto de dados de validação. Já as técnicas de regressão linear múltipla, MLP, RBF, MLP-AG, utilizaram o conjunto de dados de treinamento para extrair os padrões ou características da série de dados, e em seguida, com a configuração obtida, os modelos foram validados com o conjunto de dados de validação.

No entanto, existe algumas diferenças. As técnicas de médias, regressão linear múltipla, MLP e RBF utilizaram o conjunto de dados das estações estimadoras já determinadas, para prever as variáveis de uma estação x . Diferente dessas técnicas, o modelo de MLP-AG define seu próprio conjunto de treinamento, podendo variar sua escolha entre um ou mais conjuntos de variáveis das estações para treinamento e validação.

5.2.7. Configuração da RNA MLP

Nesta etapa foram testadas várias estruturas com diferentes números de camadas, neurônios, funções de transferência e tipos de treinamento para RNA MLP. Através desses testes foi verificado que o tipo de treinamento *Quasi-Newton Back-propagation*, que ajusta os pesos e o viés de acordo com o método de *Quasi-Newton*, apresentou resultados satisfatórios levando um pequeno tempo para convergir. Ainda foi possível observar que uma RNA MLP

com duas camadas intermediárias apresenta melhores resultados na estimação do conjunto de dados do que uma RNA MLP com apenas uma camada intermediária.

Para tanto, foram utilizadas as seguintes funções de ativação, na primeira camada intermediária, adotou-se a função tangente hiperbólica, na segunda camada optou-se pela função logística, e na última camada foi utilizada uma função linear.

O número de neurônios de cada camada também foi definido através de testes, onde foi adotado para primeira camada intermediária 30 neurônios, para segunda camada intermediária 15 neurônios e para camada de saída apenas 1 neurônio.

Esta determinação é justificada por diversos autores (CAUDILL, 1990; RODRIGUES, 2007), onde se apresenta o número de unidades das camadas intermediárias como uma incógnita, afirmando que o tamanho da camada intermediária deve ser escolhido usando o bom senso. Deve-se apenas ter o cuidado para que a mesma não possua um tamanho muito grande, porque assim a RNA estará memorizando os padrões e acabará perdendo a capacidade de generalização, e nem que a mesma possua um número de camadas muito pequeno, porque desta forma ela levará muito mais tempo para atingir a precisão desejada. Além disso, pode-se afirmar que, quando a RNA memoriza um certo conjunto de dados, ela vai apresentar pequenos erros de previsão para o período amostral, mas grandes erros para previsões fora da amostra.

5.2.8. Configuração da RNA RBF.

A configuração da RNA RBF por meio do software MATLAB é restrita, não permitindo um controle de parâmetros mais abrangente como a MLP. O software usa por padrão um conjunto de funções *radbas* em sua camada intermediária.

No treinamento desta RNA, primeiro é calculada a distância euclidiana entre as entradas (p) e o vetor de peso (w). Em seguida, o produto do módulo desta distância com o valor de *bias* (b) é submetido a função *radbas*, gerando uma saída que será a entrada da camada seguinte, que utiliza uma função linear (MATHWORKS, 2014). A Figura 20 demonstra a RNA RBF aplicada pelo software MATLAB.

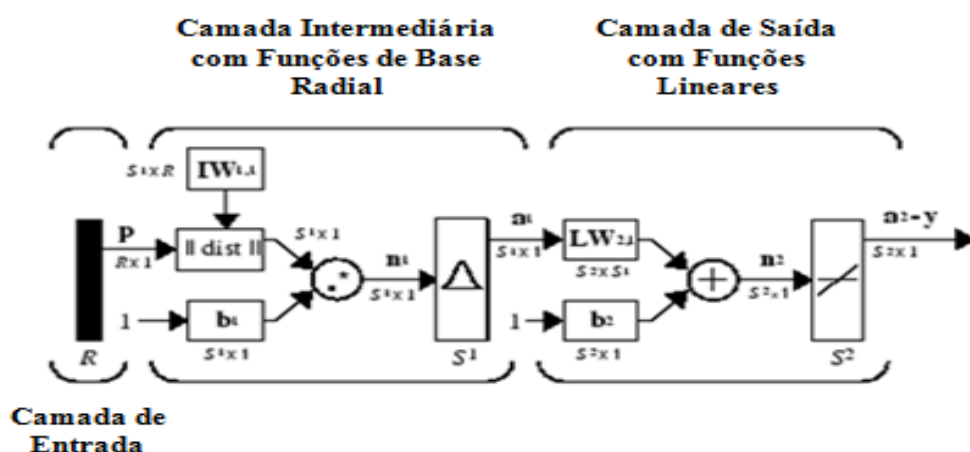


Figura20: Estrutura da RNA RBF aplicada pelo MATLAB. Fonte: <http://www.mathworks.com/help/nnet/ug/radial-basis-neural-networks.html>

Foram feitas alterações apenas no espalhamento da função de base radial (*spread*), que se deu de forma aleatória variando de 0.1 a 0.98, e no número de neurônios utilizados na camada intermediária da RNA, que também teve critério aleatório variando 1 a 30. Esse

procedimento foi adotado para realizar diferentes combinações e encontrar o melhor ajuste para um determinado problema.

5.2.9. Configuração da RNA MLP com AG.

Nesta fase, o algoritmo genético foi aplicado juntamente com a RNA MLP para definir as variáveis estimadoras de cada conjunto de dados a ser estimado, e também para fixar seus principais parâmetros.

Dentro deste contexto, o AG foi responsável por realizar diversas combinações avaliando a cada momento o conjunto de dados de treinamento e os parâmetros fixados para RNA do tipo MLP. Similar ao uso feito por Ventura (2012), o AG foi utilizado para determinar o conjunto de dados para treinamento e validação, função da primeira camada intermediária, função da segunda camada intermediária, algoritmo de treinamento, taxa de aprendizagem, taxa de momentum e um fator de auxílio (λ) para ajuste da rede, que aumenta o conjunto de dados de treinamento através da inserção de um conjunto genérico obtido pela média dos dados utilizados no treinamento.

Desta maneira, cada cromossomo ou indivíduo de uma população de um AG foi então codificado em 14 *bits* (Figura 21).

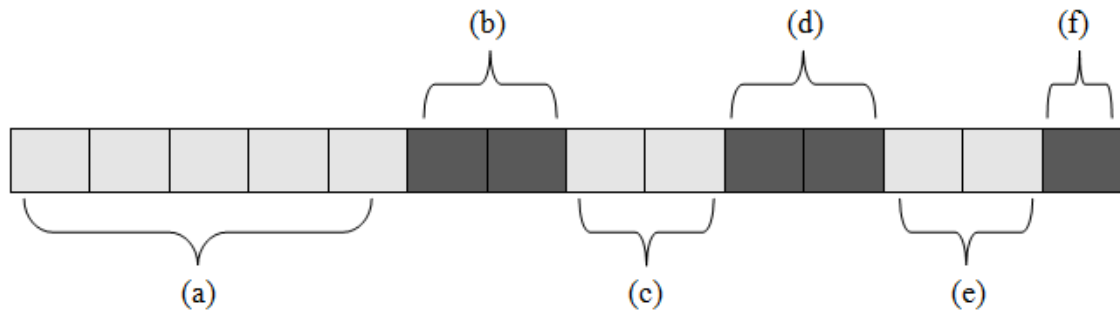


Figura 21: Exemplo de cromossomo adotado para configurar a RNA.

Os 5 *bits* iniciais pertencentes a (a) estão associados às variáveis pertencentes a cada uma das estações que podem ser utilizadas no preenchimento de uma falha. Por exemplo, se este modelo estivesse sendo executado para preencher uma falha na variável de umidade relativa do ar pertencente a estação de Campos dos Goytacazes, então esses cinco *bits* indicam se devem ou não utilizar os dados de umidade relativa do ar das regiões de Paty do Alferes, Cordeiro, Itaperuna, Resende e Rio de Janeiro. Isso é conseguido por meio da configuração binária, em que cada gene que possui o valor igual a 1 identifica que o conjunto de dados pertencentes a uma estação x está ativo e deve ser utilizado, e o valor 0 identifica que esta é inativa e deve ser ignorado.

A parte do cromossomo pertencente a (b) e (c) identifica qual será a função aplicada na primeira e na segunda camada intermediária. Esses *bits* permitem a escolha entre quatro tipos de função de ativação, apresentadas a seguir (MATHWORKS, 2014):

- Função de ativação degrau ou *Saturating linear transfer function* (*bits* 11) :

$$satlin(u_i) = \begin{cases} 1 & \text{se } u_i \geq 0 \\ 0 & \text{se } u_i < 0 \end{cases} \quad (9)$$

- Função Linear ou *Linear transfer function* (bits 0 0):

$$purelin(u_i) = u_i \quad (9.1)$$

- Função Logística ou *Log-sigmoid transfer function* (bits 1 0):

$$logsig(u_i) = \frac{1}{(1 + \exp(-u_i))} \quad (9.2)$$

- Função Tangente Hiperbólica ou *Hyperbolic tangent sigmoid transfer function* (bits 0 1):

$$tansig(u_i) = \frac{2}{(1 + \exp(-2 * u_i))} - 1 \quad (9.3)$$

O espaço pertencente a (d) tem a responsabilidade de fazer a escolha entre os tipos de treinamento, podendo optar pelo algoritmo *Back-propagation* ou por suas variações, como MATHWORKS (2014):

- *Gradient descent Back-propagation - traingd* (0 1): Esse algoritmo foi descrito na literatura, seu treinamento atualiza os pesos e os valores de *bias* da RNA de forma contrária ao valor do gradiente.
- *Quasi-Newton Back-propagation - trainbfg* (1 0): Diferente do método de Newton que é uma alternativa ao método do gradiente conjugado, possuindo uma convergência mais rápida, no entanto mais complexa. Esse algoritmo atualiza os pesos e o valor do *bias* de acordo com o método de Quasi-Newton que é baseado no método de Newton, mas não exige o cálculo da segunda derivada (HAGAN; MENHAJ, 1994).
- *Resilient Back-propagation - RProp - trainrp* (0 0): Esse algoritmo oferece uma convergência rápida e geralmente mais eficaz, se tratando de escapar de pontos mínimos locais. A propagação resiliente é um eficiente esquema de aprendizagem que realiza uma adaptação direta do peso, com base em informações locais do gradiente (RIEDMILLER; BRAUN, 1993). Seu objetivo é eliminar os efeitos negativos das magnitudes das derivadas parciais na definição do ajuste dos pesos. Segundo (BRAGA; CARVALHO; LUDEMIR, 2012), essa influência negativa ocorre quando a saída de um neurônio for próxima de 0 e a saída desejada for 1. Logo, a derivada será próxima de 0 fazendo com que o ajuste dos pesos deste neurônio receba um valor mínimo. Sendo assim, o RProp elimina esse problema utilizando apenas o sinal da derivada, para indicar a direção do ajuste dos pesos e não o seu valor .
- *Levenberg-Marquardt Back-propagation - trainlm* (1 1): Da mesma forma que o algoritmo de *Back-propagation* é baseado no método do gradiente descendente, o *Levenberg-Marquardt* é uma aproximação do método de Newton, feito através de

melhorias no método de Gauss-Newton por meio da utilização de uma taxa de aprendizado variável. Esse algoritmo pode convergir de forma rápida, porém demanda de uma grande quantidade de memória e de alta complexidade computacional (HAGAN; MENHAJ, 1994).

Já a taxa de aprendizado e de momentum adotadas são obtidas através dos *bits* de (e). A taxa de aprendizado pode variar entre 0,01 até 0,86, e a taxa de momentum é obtida por meio de uma multiplicação em cima dos valores da taxa de aprendizado. Os multiplicadores podem variar entre 0,8, 0,15, 0,12 e 0,10 conforme a escolha dos *bits* (1 0), (0 0), (1 1) e (0 1).

Já o último bit identificado por (f) tem a finalidade de criar um fator de auxílio (λ) para ajuste da rede, que aumenta o conjunto de dados de treinamento através da inserção de um conjunto genérico obtido pela média dos dados utilizados no treinamento.

Outros fatores importantes dessa metodologia, como a configuração do algoritmo genético, foram fixados em 40 indivíduos, número máximo de gerações igual a 60, uso do elitismo para propagar cerca de 15 indivíduos com avaliações altas para próxima geração, sorteio aleatório de 5 indivíduos com avaliações baixas para propagar para próxima geração, a fim de evitar que características genéticas contidas apenas em um indivíduo de baixa aptidão sejam perdidas, criando uma população sem diversidade com indivíduos semelhantes.

Além desses parâmetros, o AG ainda utilizou o método de torneio com seu tamanho fixado em 5, tanto para escolha do primeiro pai quanto para escolha do segundo pai, e operador de *crossover* de dois pontos para geração dos novos indivíduos.

A função objetivo do AG se deu pela minimização do erro global absoluto. Essa função foi calculada criando-se, primeiramente, uma RNA com os parâmetros definidos por um indivíduo a ser avaliado. Após o treinamento dessa rede, avalia-se a aptidão da mesma através da estimação do conjunto de dados de validação. A minimização do valor originado pela soma das diferenças entre os valores estimados e os esperados para o conjunto de dados de validação é a função objetivo.

5.2.10. Interpretação dos Resultados

Os valores gerados pelas saídas de cada metodologia adotada no presente estudo se encontram na escala entre 0 e 1 devido à normalização aplicada. Então, para que os resultados possam se analisados, faz-se necessário a aplicação de uma função que mude a escala dos valores para a escala real (Equação 10).

$$V_j^{Est} = (x_j * x^{max}) - (x_j * x^{min}) + x^{min} \quad (10)$$

- Onde V_j^{Est} é a variável desnormalizada.
- x_j é a variável predita pela rede.
- x^{min} valor mínimo observado entre as variáveis.
- x^{max} valor máximo observado entre as variáveis.

5.2.11. Avaliação do Desempenho

O desempenho das técnicas foi verificado por meio de índices estatísticos que estão se tornando cada vez mais comuns na avaliação de variáveis meteorológicas (PEZZOPANE et al., 2012).

O coeficiente de correlação de Pearson (r) foi utilizado em duas partes, primeiro para avaliar a diferença da qualidade do conjunto de regiões estimadoras definidas manualmente com o conjunto definidas pelo modelo híbrido. Em seguida, foi utilizado para avaliar a medida de associação entre os dados estimados e os observados.

O coeficiente de correlação (r) mede o grau de relação das variáveis na amostra, e seu valor pode variar entre -1 a 1, sendo 1 uma correlação perfeita (FONSECA; MARTINS; TOLEDO, 2012; DESHMUKH; GHATOL, 2010) (Equação 11).

$$r = \frac{\frac{\sum_{j=1}^N (x_j - \bar{x})(o_j - \bar{o})}{N}}{\sqrt{\frac{\sum_{j=1}^N (x_j - \bar{x})^2}{N}} \cdot \sqrt{\frac{\sum_{j=1}^N (o_j - \bar{o})^2}{N}}} \quad (11)$$

- n número de dados utilizados
- O_j é o valor observado.
- x_j é o valor predito pelas técnicas empregadas.
- $\bar{O}$ é a média dos dados observados.
- $\bar{x}$ é a média dos dados preditos.

Cálculo do erro médio absoluto (EMA): essa medida representa a magnitude dos desvios propiciados pelo método utilizado (Equação 12).

$$EMA = \frac{\sum_{j=1}^n |O_j - x_j|}{n} \quad (12)$$

Cálculo da raiz do erro médio quadrático (REMQ): Quanto menor for seu resultado melhor será a performance do modelo aplicado (Equação 13).

$$REMQ = \sqrt{\frac{\sum_{j=1}^n (O_j - x_j)^2}{n}} \quad (13)$$

Cálculo do erro médio percentual (EMP)(Equação 14).

$$EMP = \frac{\sum_{j=1}^n \frac{|O_j - x_j|}{O_j}}{n} * 100 \quad (14)$$

Cálculo do índice de confiança (C) (Equação 16): essa medida permite analisar conjuntamente a precisão e a exatidão dos resultados obtidos, e ele é obtido através do produto do coeficiente de correlação (r) pelo índice de concordância (D) (Equação 15). Seus

valores variam de zero (0) para nenhuma concordância a um (1) para a concordância perfeita (PEZZOPANE et al., 2012). A Tabela 4 demonstra os critérios para avaliação de desempenho.

$$D = 1 - \frac{\sum_{j=1}^n (O_j - x_j)^2}{\sum_{j=1}^n (|x_j - \bar{O}| + |O_j - \bar{O}|)^2} \quad (15)$$

$$C = (r * D) \quad (16)$$

Tabela 4: Critérios para avaliação e análise do desempenho de modelos com base no índice de confiança

Valor de IC	Desempenho
> 0,85	Ótimo
0,76 a 0,85	Muito bom
0,66 a 0,75	Bom
0,61 a 0,65	Mediano
0,51 a 0,60	Sofrível
0,41 a 0,50	Mau
≤ 0,40	Péssimo

5.2.12. Preenchimento de Falhas

Para realizar o preenchimento das falhas foram feitas 20 simulações para cada modelo de RNA aplicado, onde cada simulação variou entre 30 minutos a 4 horas e o melhor resultado obtido na estimação do conjunto de dados de validação da região avaliada foi tomado como solução.

Após a fase acima, o preenchimento das falhas reais de cada estação foi realizado somente através da utilização dos dados pertencentes ao conjunto de estimadores registrados no mesmo período, não considerando o uso de valores preenchidos, devido não terem sido avaliados por um especialista da área de meteorologia.

6. RESULTADOS

Neste tópico, são apresentados e comparados os resultados obtidos por cada modelo aplicado no presente estudo. Os resultados foram obtidos por meio de análise das simulações realizadas para cada conjunto de treinamento e validação das variáveis de evaporação, temperatura máxima do ar e umidade relativa do ar. As simulações que apresentaram os menores índices de erro foram consideradas soluções.

Primeiramente, foram geradas as matrizes dos coeficientes de correlação das estações estimadoras definidas manualmente e pelo AG. Isso ocorreu a fim de verificar o grau de relação entre os dados a serem preenchidos com o conjunto de estimadores. Após esta fase, o desempenho das técnicas foi avaliado através de análises dos dados estimados para o conjunto de dados de validação. Por último, foram preenchidos os conjuntos de dados através da aplicação das técnicas que obtiveram os melhores resultados.

6.1. Configurações

O presente estudo fez uso de várias metodologias, onde para cada uma delas foram definidos parâmetros como funções das camadas intermediárias, tipos de treinamento, taxas de aprendizado e outros. A fim de um melhor entendimento, segue abaixo as siglas adotadas neste tópico e informações sobre os parâmetros:

- MED: Média.
- RLM: Regressão linear múltipla.
- MLP: *Multilayer perceptron*.
- RBF: *Radial Basis Function*.
- MLP-AG: *Multilayer perceptron* treinada com Algoritmo Genético.
- (λ): Fator de auxílio (inserção da média dos dados para treinamento e ajuste).

Outros parâmetros são as funções de ativação das camadas intermediárias da RNA MLP-AG, que podem ser *Saturating linear transfer function* (*satlin*), *Linear transfer function* (*purelin*), *Log-sigmoid transfer function* (*logsig*) e *Hyperbolic tangent sigmoid transfer function* (*tansig*). Já os treinamentos foram escolhidos entre os tipos *Gradient Descent Back-propagation* (*traingd*), *Quasi-Newton Back-propagation* (*trainbfg*), *Resilient Back-propagation* (*trainrp*) e *Levenberg-Marquardt Back-propagation* (*trainlm*).

6.2. Resultados de Evaporação

Como inferido anteriormente, a fim de comparar o conjunto de estimadores escolhidos de forma manual para estimativa de evaporação com conjunto definidos pelo modelo híbrido, foram realizadas as correlações entre as estações. A matriz de correlações apresentada para variável de evaporação é demonstrada na Tabela 5.

Tabela 5: Índices de correlação apresentados para a variável de evaporação nas estações estudadas.

Índice de correlação das variáveis de Evaporação						
<i>Estações</i>	<i>CAM</i>	<i>COR</i>	<i>ITA</i>	<i>PAT</i>	<i>RES</i>	<i>RJ</i>
Campos dos Goytacazes	1					
Cordeiro	0,67	1				
Itaperuna	0,69	0,79	1			
Paty do Alferes	0,53	0,71	0,81	1		
Resende	0,48	0,63	0,70	0,73	1	
Rio de Janeiro	0,63	0,54	0,56	0,63	0,39	1

Onde :

- CAM= Correlação entre as regiões e Campos dos Goytacazes.
- COR= Correlação entre as regiões e Cordeiro.
- ITA= Correlação entre as regiões e Itaperuna.
- PAT= Correlação entre as regiões e Paty do Alferes.
- RES= Correlação entre as regiões e Resende.
- RJ= Correlação entre as regiões e Rio de Janeiro.

Os resultados apresentados para estimativa de evaporação foram gerados a partir do conjunto de dados de treinamento e validação. Respectivamente, cada conjunto de dados pertencentes a variável de evaporação, após o pré-processamento, ficou com 108 dados sendo 81 utilizados para treinamento que corresponde à cerca de 75%, e 27 para validação cerca de 25%.

6.2.1. Análise do Conjunto de Dados Estimadores para Variável de Evaporação

Inicialmente, para cada estação, foram definidas 3 estações estimadoras que poderiam ser utilizadas no preenchimento dos dados faltantes de cada região. Isto pode ser observado na Tabela 3.

No entanto, essa escolha não foi adotada para todos os métodos. A RNA treinada com auxílio do AG definiu seu próprio conjunto de dados estimadores, optando por um número menor, igual ou superior de variáveis. Para tanto, as definições adotadas pelo modelo híbrido para variável de evaporação podem ser visualizadas na Tabela 6.

Tabela 6: Conjunto de variáveis estimadoras de evaporação definidas com o AG.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>	<i>4º estimador</i>	<i>5º estimador</i>
CAM	COR	ITA	PAT	X	X
COR	CAM	ITA	RES	RJ	X
ITA	COR	CAM	PAT	RES	RJ
RJ	CAM	PAT	X	X	X
PAT	CAM	ITA	RJ	X	X
RES	COR	ITA	RJ	X	X

Onde :

- CAM = Campos dos Goytacazes.
- COR = Cordeiro.
- ITA= Itaperuna.
- RJ= Rio de Janeiro.
- PAT= Paty do Alferes.
- RES= Resende.
- X= Não estão sendo utilizados.

Tomando por base o coeficiente de correlação obtido como critério de avaliação das variáveis adotadas para estimativa de evaporação, pode-se analisar e comparar a qualidade das escolhas definidas de forma manual com as escolhas adotadas pelo AG. Segue abaixo as matrizes de coeficiente de correlação (Tabelas 7 e 8).

Tabela 7: Índices de correlação dos conjuntos de estimadores definidos manualmente para variável de evaporação.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>
CAM	0,67	0,69	0,63
COR	0,54	0,79	0,67
ITA	0,56	0,79	0,69
RJ	0,63	0,39	0,54
PAT	0,63	0,73	0,71
RES	0,39	0,73	0,63

Tabela 8: Índices de correlação dos conjuntos de estimadores definidos pelo AG para variável de evaporação.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>	<i>4º estimador</i>	<i>5º estimador</i>
CAM	0,67	0,69	0,53	X	X
COR	0,67	0,79	0,63	0,54	X
ITA	0,79	0,69	0,81	0,70	0,56
RJ	0,63	0,63	X	X	X
PAT	0,53	0,81	0,63	X	X
RES	0,63	0,70	0,39	X	X

A partir destas tabelas, observa-se que as escolhas adotadas por cada metodologia encontram-se equilibradas. No entanto, foram poucos os conjuntos de estimadores definidos com os maiores graus de relação linear existentes, ou seja, as combinações ótimas. Algumas dessas combinações podem ser observadas na Tabela 7, onde os estimadores para Campos dos Goytacazes possuem os maiores coeficientes de correlação, e analisando a Tabela 8 observa-se que as combinações para Itaperuna e para Rio de Janeiro também possuem os maiores coeficientes de correlação.

6.2.2. Resultados e Preenchimento das Falhas de Evaporação para estação de Campos dos Goytacazes

Para geração dos resultados, o modelo MLP-AG optou por uma função (*satlin*) em sua primeira camada intermediária e uma (*logsig*) em sua segunda camada intermediária,

diferindo das funções fixadas para o modelo MLP que foram na primeira camada intermediária uma (*tansig*) e na segunda camada uma (*logsig*).

Outra diferença constatada foi os tipos de treinamentos em que o MLP-AG escolheu o (*traingd*), enquanto que para MLP foi fixado o (*trainbfg*). O modelo MLP-AG ainda utilizou o fator de auxílio (λ), aumentando seu conjunto de dados de treinamento (APÊNDICE A).

Os resultados obtidos por cada modelo aplicado na estimativa dos 27 meses pertencentes ao conjunto de validação da variável de evaporação, podem ser visualizados na Tabela 9.

Tabela 9: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Campos dos Goytacazes.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(ml)	EMP(%)
MED	0,83	19,1	0,87	0,72	Bom	16,35	16,12
RLM	0,82	16,1	0,87	0,71	Bom	12,4	11,9
MLP	0,85	14,2	0,92	0,78	Muito Bom	10,0	9,8
RBF	0,79	16,54	0,86	0,68	Bom	12,69	12,10
MLP-AG	0,88	12,45	0,93	0,82	Muito Bom	9,39	9,67

Como demonstrado na Tabela 9, os índices de confiança obtidos pelas predições originadas por cada modelo aplicado variaram entre Bom e Muito Bom. Contudo, é possível verificar que os melhores resultados foram obtidos pelos modelos de RNAs MLP e MLP-AG.

Entretanto, apesar de pequena, há uma diferença nos resultados dos modelos, onde é possível visualizar que o modelo MLP-AG apresentou o maior índice de (r), tendo ainda alcançado os menores índices de erro, variando entre 9,39 ml para EMA, 12,45 para REMQ, e 9,67% de EMP, o que caracteriza a qualidade da estimação dos dados em aproximadamente 90%.

Por isso, o modelo MLP-AG foi adotado para preenchimento das falhas reais ocorridas na série de dados de evaporação. Esse modelo permitiu o preenchimento de 100% das falhas ocorridas. Assim a estimação do conjunto de dados de validação pode ser visualizada na Figura 22, e o preenchimento é demonstrado pela cor vermelha na Figura 23. Mais informações sobre as falhas veja o ANEXO A.

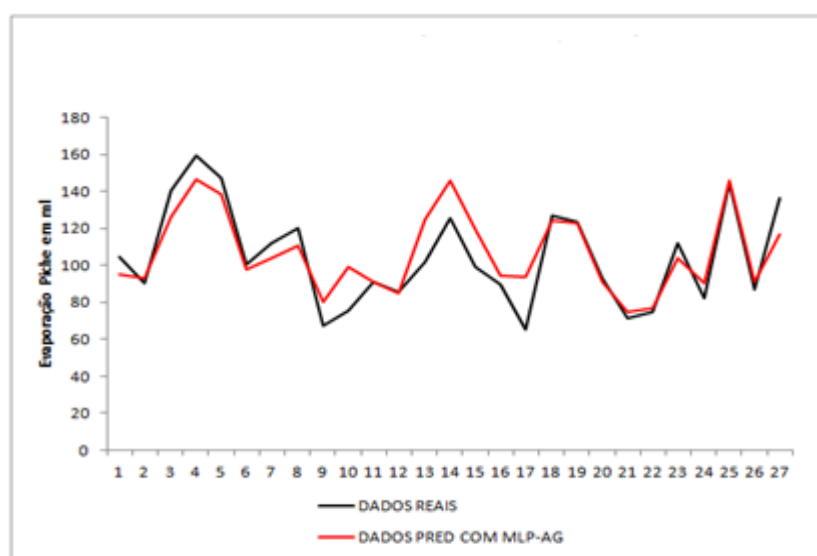


Figura 22: Dados estimados de evaporação pelo modelo MLP-AG para a estação de Campos dos Goytacazes.

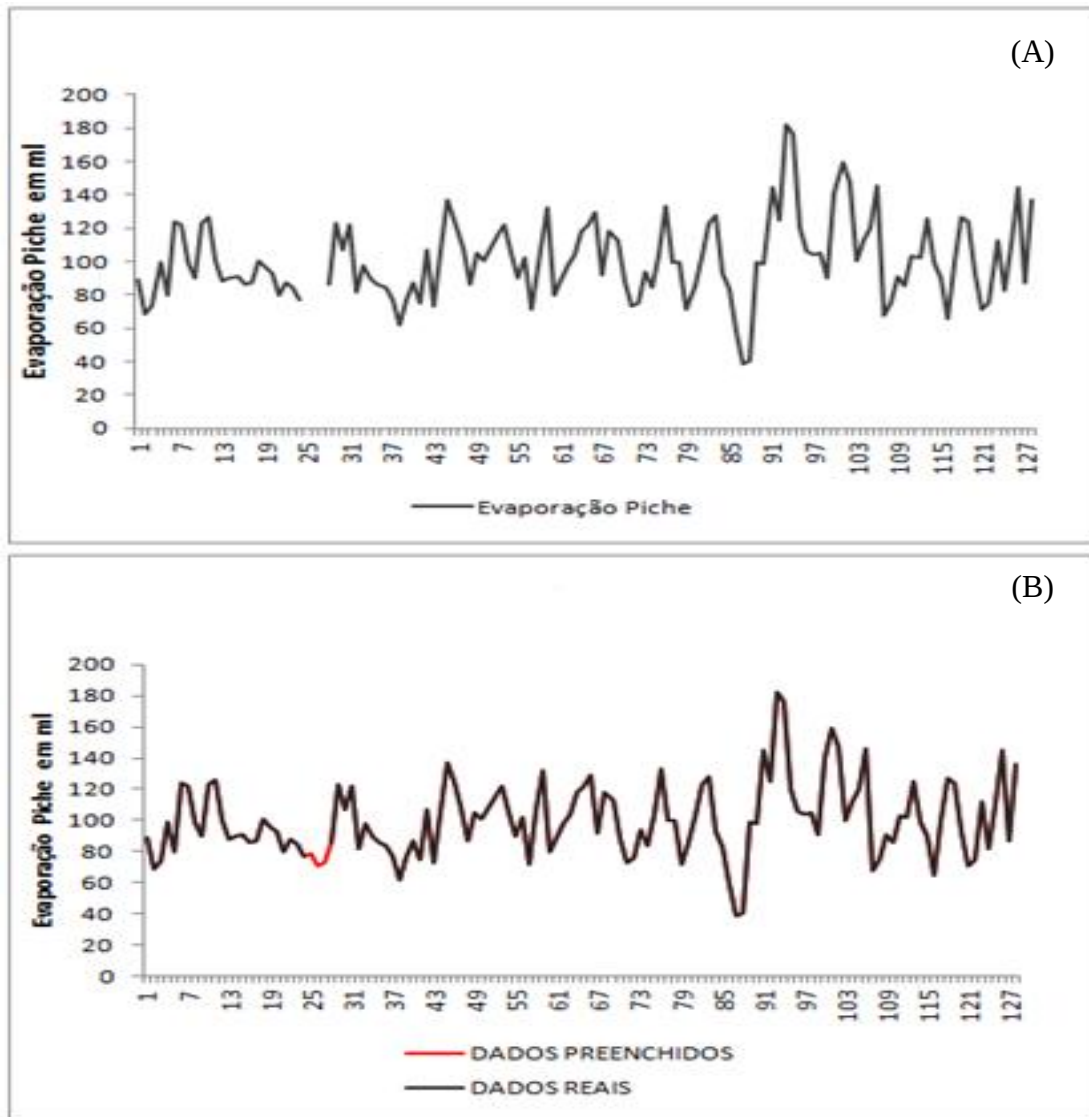


Figura 23: Estação de Campos dos Goytacazes , (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.

6.2.3. Resultados e Preenchimento das Falhas de Evaporação para estação de Cordeiro

Para gerar os resultados de Cordeiro, o modelo MLP-AG utilizou o fator de auxílio (λ) e optou por alguns parâmetros diferentes dos adotados pelo modelo MLP, como uso das funções *satlin* em suas camadas intermediárias. No entanto, o modelo MLP-AG também fez uso do algoritmo de treinamento *trainbfg* (APÊNDICE A).

Os resultados obtidos com a aplicação dos modelos para estimativa dos 27 meses apresentaram pequenos erros e um alto coeficiente de correlação. Esses resultados podem ser visualizados na Tabela 10.

Tabela 10: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Cordeiro.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(ml)	EMP(%)
MED	0,92	5,1	0,96	0,88	Ótimo	4,2	8,95
RLM	0,94	5,03	0,95	0,90	Ótimo	4,1	8,32
MLP	0,95	4,42	0,97	0,92	Ótimo	3,21	6,5
RBF	0,93	5,58	0,94	0,87	Ótimo	4,04	7,93
MLP-AG	0,96	3,86	0,97	0,93	Ótimo	3,07	6,65

Analisando a Tabela 10, observa-se que o modelo MLP obteve o menor EMP (6,5%). No entanto, o mesmo possui alguns índices inferiores aos obtidos pelo modelo MLP-AG, como menor valor de (r), maior (EMA), maior (REMQ) e menor índice (C). Por isso, o modelo MLP-AG foi considerado como solução. É possível verificar a diferença entre os dados preditos por cada modelo na Figura 24.

O conjunto de dados definidos pelo modelo MLP-AG permitiu o preenchimento de 100% das falhas ocorridas na série histórica de evaporação pertencente a região de cordeiro, ou seja, houve preenchimento dos 3 meses com falhas da série. Isso pode ser identificado na Figura 25 pela cor vermelha, outras informações estão apresentadas no ANEXO B.

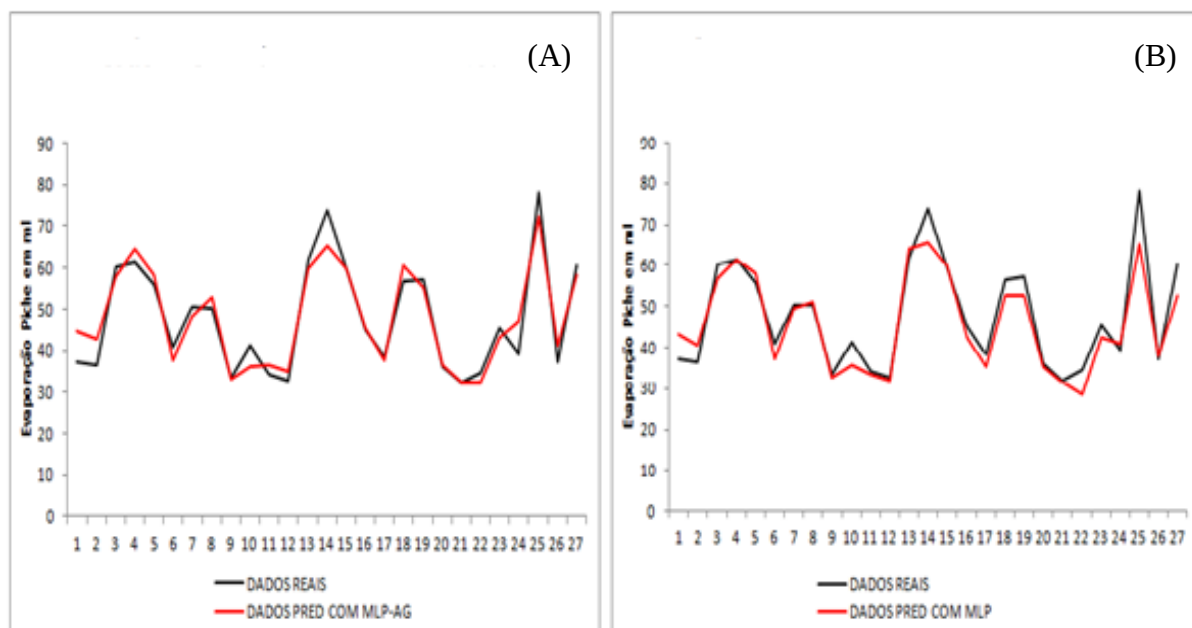


Figura 24: Estação de Cordeiro, (A) Dados reais e estimados de evaporação com o modelo MLP-AG, (B) Dados reais e estimados de evaporação com o modelo MLP.

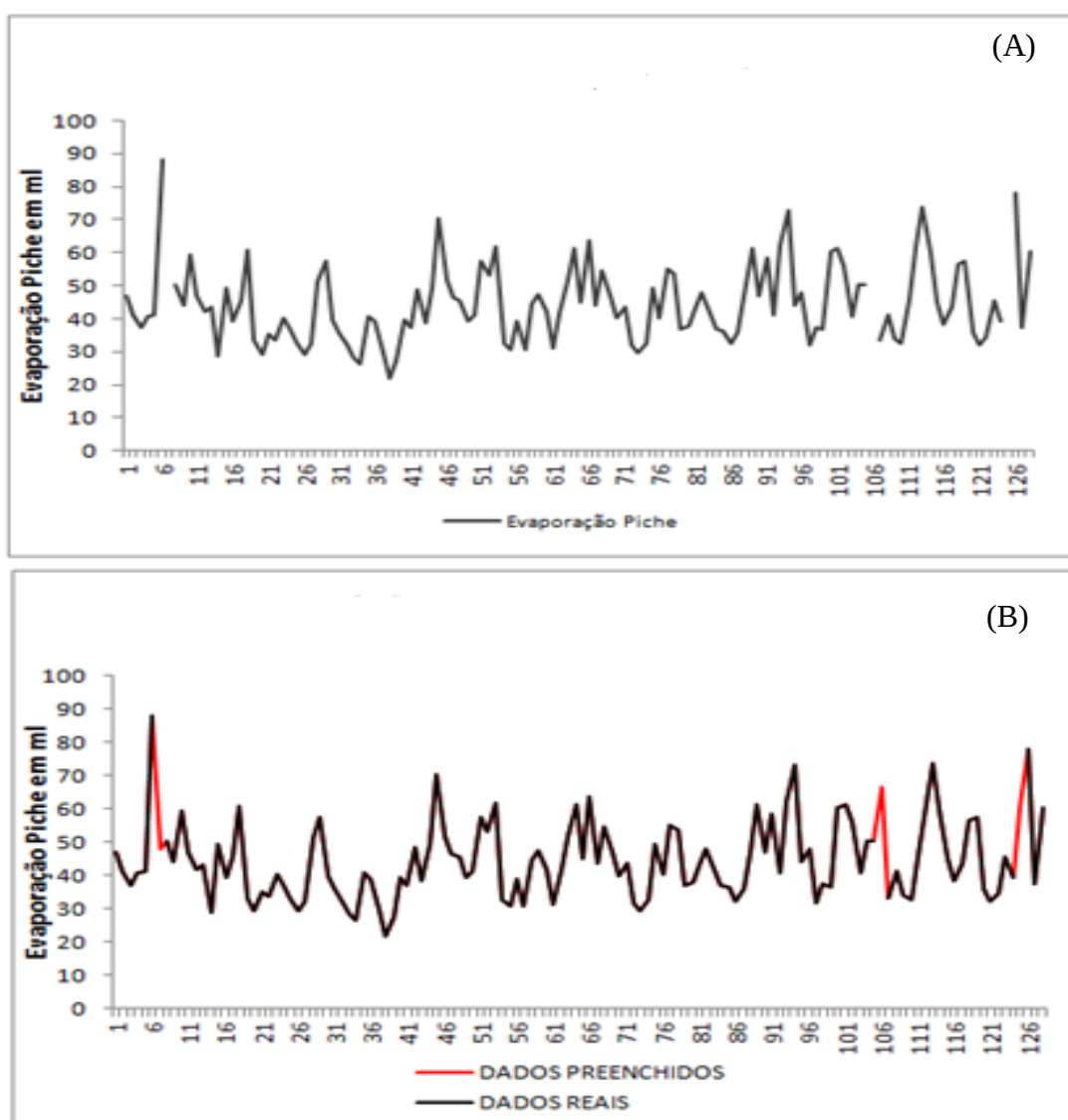


Figura 25: Estação de Cordeiro, (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.

6.2.4. Resultados e Preenchimento das Falhas de Evaporação para estação de Itaperuna

O modelo híbrido MLP-AG utilizou em suas camadas intermediárias a função *purelin*, o algoritmo de treinamento *trainrp* e fator de auxílio (λ), aumentando seu conjunto de dados. Para mais informações consulte o APÊNDICE A.

Os resultados obtidos com a aplicação dos modelos na estimativa os 27 meses de evaporação da estação de Itaperuna podem ser visualizados na Tabela 11.

Tabela 11: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Itaperuna.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(ml)	EMP(%)
MED	0,92	23,45	0,85	0,78	Muito Bom	18,22	14,11
RLM	0,95	13,11	0,95	0,90	Ótimo	11,10	13,11
MLP	0,95	11,38	0,97	0,92	Ótimo	9,24	9,42
RBF	0,94	12	0,97	0,91	Ótimo	10,35	10,01
MLP-AG	0,97	8,73	0,98	0,96	Ótimo	6,87	6,01

As análises mostraram que os modelos RLM, MLP, RBF e MLP-AG apresentaram os melhores desempenhos mantendo-se em ótimos. Contudo, através de uma observação mais apurada verifica-se que o modelo RLM, assim como o modelo MED também apresentou um EMP elevado, se mantendo em 13,11%. Porém, o modelo MLP-AG demonstrou resultados superiores como o maior coeficiente (r) entre os dados estimados e os observados, menor taxa de EMA (6,87 ml), menor REMQ (8,73) e EMP igual a 6,01%, que foi cerca de 3,41% menor que o EMP apresentado pelo modelo MLP, e 4% menor que o erro apresentado pelo modelo RBF. Isto caracteriza que cada dado estimado pelo modelo MLP-AG tem uma taxa de acerto de 94%.

Por esse motivo, o modelo MLP-AG foi adotado como solução. No entanto, como não ocorre falhas em evaporação no período utilizado, segue abaixo somente o gráfico dos dados reais registrados e dos dados estimados por esse modelo (Figura 26).

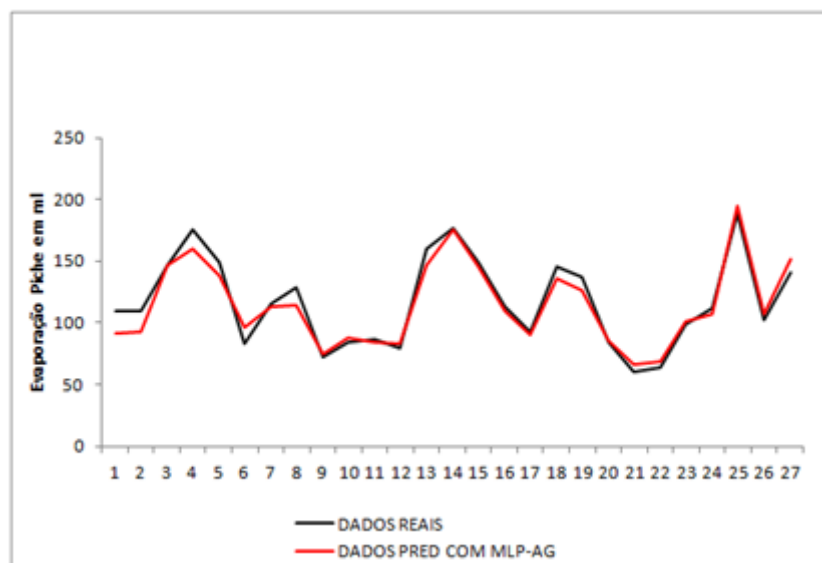


Figura 26: Dados estimados de evaporação com o modelo MLP-AG para a estação de Itaperuna.

6.2.5. Resultados e Preenchimento das Falhas de Evaporação para estação do Rio de Janeiro

Os resultados obtidos com o modelo MLP-AG foram conseguidos por meio das configurações, sem usar o fator de auxílio (λ), empregando-se em sua primeira camada intermediária uma função *logsig*, na segunda camada intermediária uma função *satlin* e para treinamento utilizou o algoritmo *trainbfg*. Mais informações são apresentadas no APÊNDICE A.

Os resultados obtidos com a aplicação dos modelos na estimativa dos 27 meses de evaporação da estação do Rio de Janeiro podem ser visualizados na Tabela 12.

Tabela 12: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação do Rio de Janeiro.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(ml)	EMP(%)
MED	0,75	17,87	0,78	0,58	Sofrível	14,18	14,45
RLM	0,71	19,71	0,65	0,46	Mau	17,83	19,12
MLP	0,84	9,96	0,91	0,76	Muito Bom	7,77	7,89
RBF	0,65	14,24	0,77	0,50	Mau	8,49	8,05
MLP-AG	0,82	13,43	0,87	0,72	Bom	8,84	9,11

Como pode ser visto na tabela anterior, os modelos MED, RLM e RBF não apresentaram bons resultados, e acabaram demonstrando uma certa dificuldade na estimativa dos dados de evaporação da região do Rio de Janeiro.

Sendo assim, apenas os modelos MLP e MLP-AG apresentaram desempenhos satisfatórios. No entanto, apesar de superestimar e subestimar os registros de alguns meses, o modelo MLP mostrou-se superior, obtendo o maior coeficiente (r) entre os dados estimados e os registrados, o menor EMA (7,77 ml), o menor REMQ (9,96), e o EMP igual a 7,89%, o que caracteriza que cada dado estimado pelo modelo MLP teve uma taxa de acerto de 92,11%. É possível verificar os dados preditos por esse modelo na Figura 27.

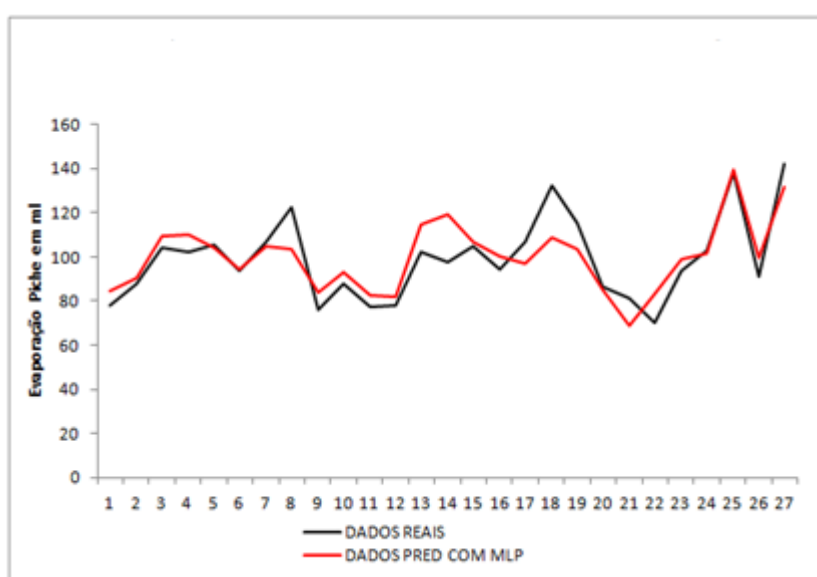


Figura 27: Resultados da estimação de evaporação obtidos com o modelo MLP para a estação do Rio de Janeiro.

O preenchimento das falhas reais por meio desse modelo pode ser visualizado na Figura 28. O conjunto de dados definidos para estimar os valores permitiu o preenchimento de 100% das falhas, ou seja, 3 meses falhados da série identificados pela cor vermelha. Outras informações podem ser obtidas no ANEXO F.

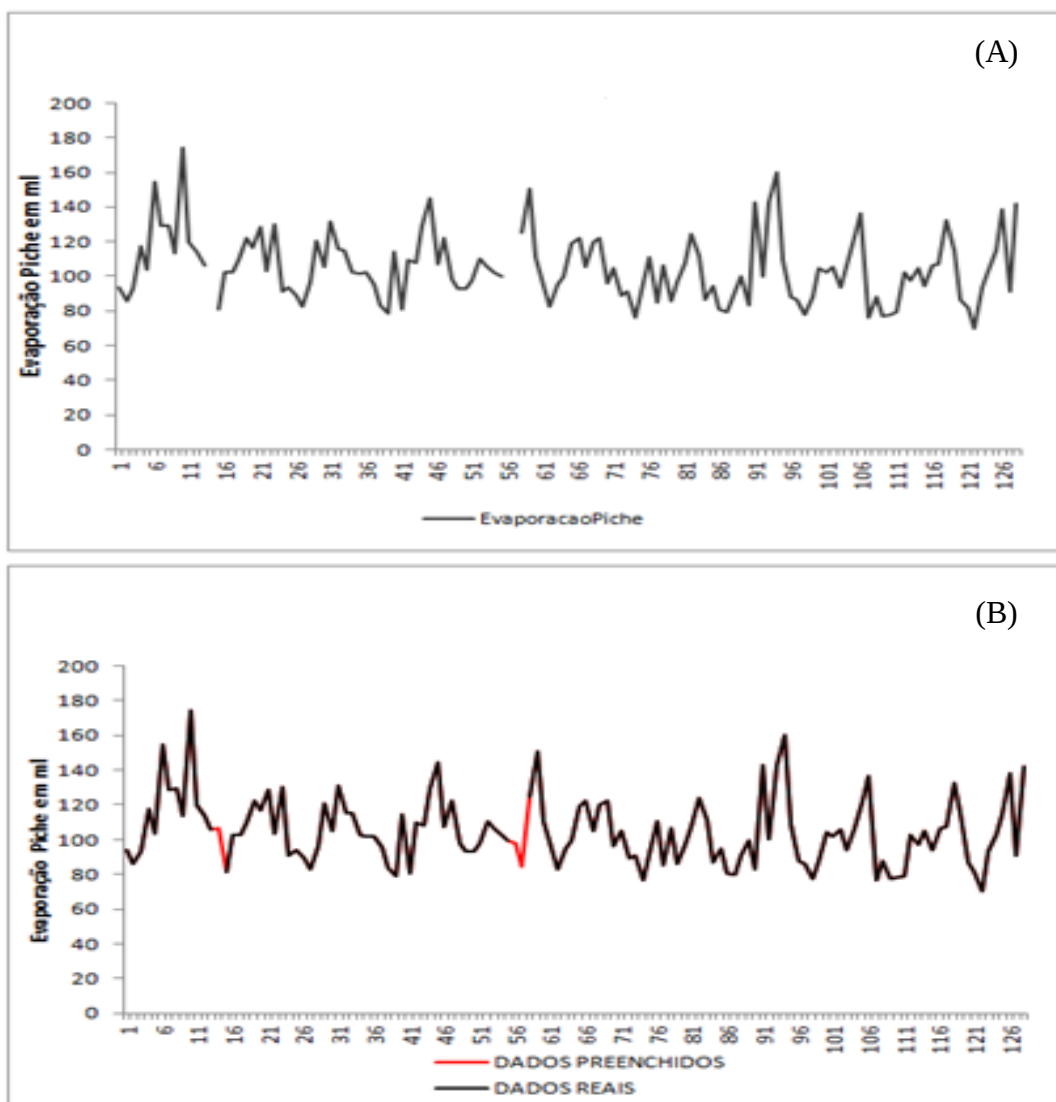


Figura 28: Estação do Rio de Janeiro, (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.

6.2.6. Resultados e Preenchimento das Falhas de Evaporação para estação de Paty do Alferes

Para geração dos resultados da região de Paty do Alferes, os parâmetros impostos pelo modelo MLP-AG se diferiram completamente dos fixados para o modelo MLP, em que a MLP-AG utiliza em suas camadas intermediárias funções *purelin* e o tipo de treinamento *trainrp* e não faz uso do fator de auxílio (λ) (APÊNDICE A).

Sendo assim, os resultados obtidos com a aplicação dos modelos para estimativa dos 27 meses de evaporação da estação de Paty do Alferes estão apresentados na Tabela 13.

Tabela 13: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Paty do Alferes.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(ml)	EMP(%)
MED	0,96	9,15	0,95	0,91	Ótimo	8,09	11,14
RLM	0,96	11,83	0,92	0,88	Ótimo	10,76	15,18
MLP	0,94	9,38	0,96	0,90	Ótimo	7,28	9,17
RBF	0,91	10,78	0,93	0,853	Ótimo	9,12	11,90
MLP-AG	0,95	6,97	0,97	0,93	Ótimo	5,34	6,54

Novamente é possível verificar que o desempenho dos modelos se mantiveram em Ótimo. Contudo, apesar de não ter obtido o maior (r), o modelo MLP-AG apresentou melhores resultados, como o menor EMA (5,34 ml), o menor REMQ (6,97), e o menor EMP (6,54 %), o que indica que este modelo conseguiu prever os dados mensais de evaporação com um índice de acerto de 93,46%. Os dados estimados podem ser vistos na Figura 29.

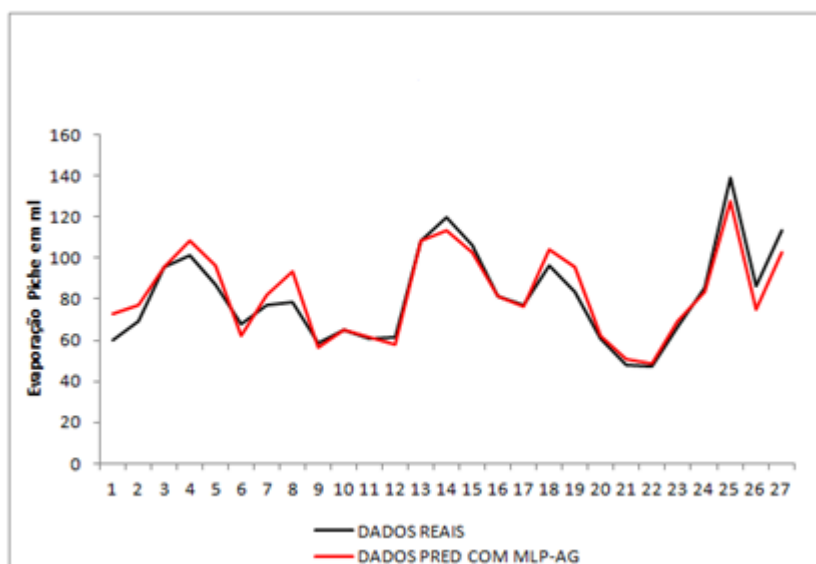


Figura 29: Resultados de estimativa de evaporação obtidos com o modelo MLP-AG para estação Paty do Alferes.

Deve ser ressaltado que a série histórica de evaporação pertencente a Paty do Alferes possuía apenas um mês com falha, e o preenchimento dessa falha pode ser identificado pela cor vermelha na Figura 30. O ANEXO D apresenta mais informações.

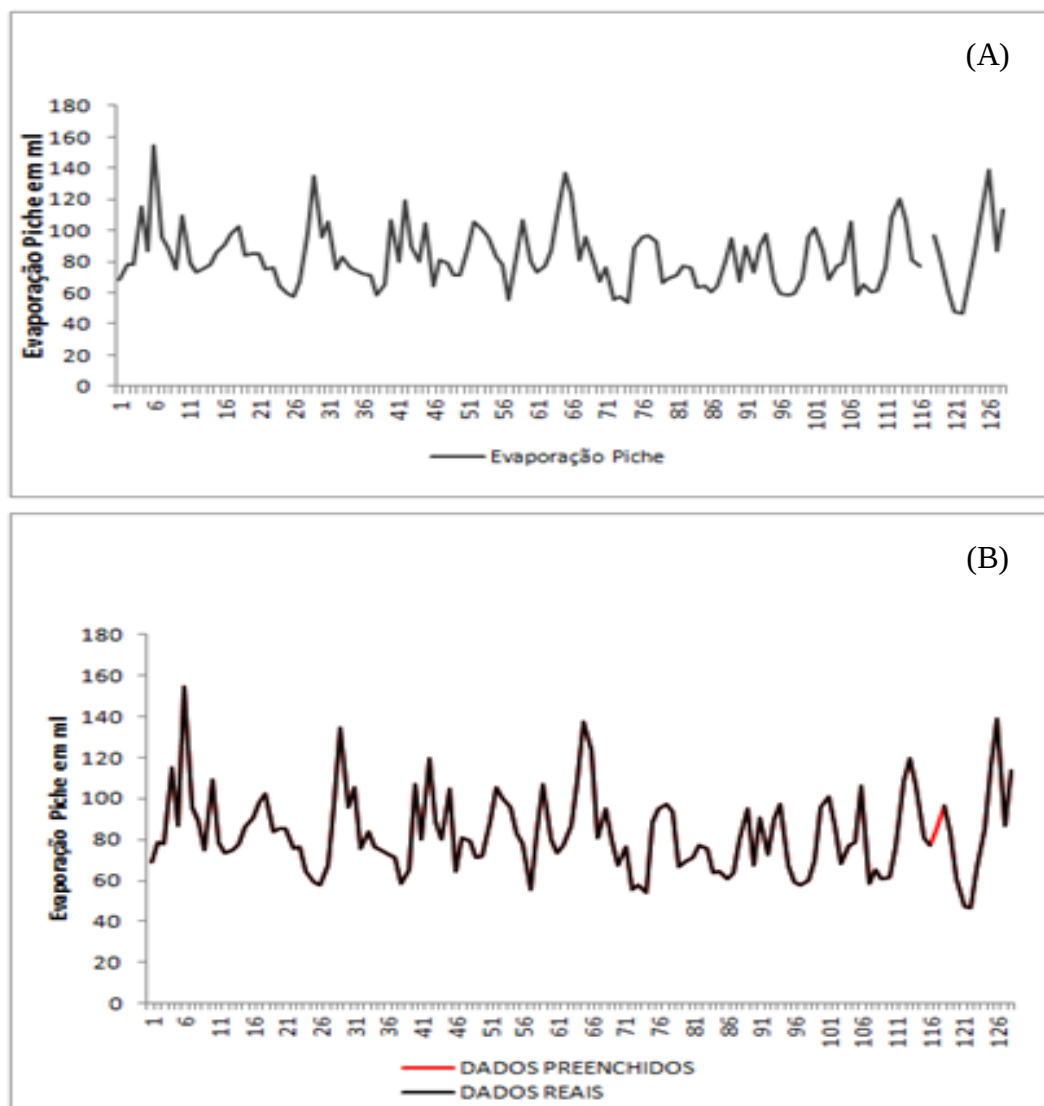


Figura 30: Estação de Paty do Alferes, (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.

6.2.7. Resultados e Preenchimento das Falhas de Evaporação para estação de Resende

O modelo MLP-AG utilizou em sua primeira camada intermediária a função *purelin*, e na segunda camada intermediária a função *logsig*, não fazendo uso do fator de auxílio (λ) e aplicando o algoritmo de treinamento *trainbfg*. Mais informações no APÊNDICE A.

Os resultados obtidos com a aplicação dos modelos na estimativa dos 27 meses de evaporação da estação de Resende podem ser visualizados na Tabela 14.

Tabela 14: Resultados obtidos com os modelos aplicados na estimativa de evaporação da estação de Resende.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(ml)	EMP(%)
MED	0,88	28,78	0,83	0,73	Bom	24,91	18,68
RLM	0,92	28,73	0,82	0,76	Bom	25,32	18,44
MLP	0,85	19,67	0,92	0,78	Muito Bom	14,46	11,13
RBF	0,85	25,98	0,85	0,72	Bom	21,39	15,54
MLP-AG	0,88	17,86	0,92	0,82	Muito Bom	12,22	9,25

Conforme pode ser constatado na Tabela 14, os modelos MLP, RBF e MLP-AG apresentaram os melhores resultados. No entanto, o modelo MLP-AG se destaca em relação às menores taxas de erros, como EMA (12,22 ml) que respectivamente é a metade do EMA apresentado pelos modelos MED e RLM, menor REMQ (17,86) e EMP igual 9,25 %. Isto é 9,43% menor que o EMP apresentado pelo modelo MED, 9,19% menor que o EMP apresentado pelo modelo RLM, 6,29 % menor que o EMP gerado pelo modelo RBF e 1,88% menor que o EMP apresentado pelo modelo MLP. Os dados estimados por esse modelo podem ser vistos na Figura 31.

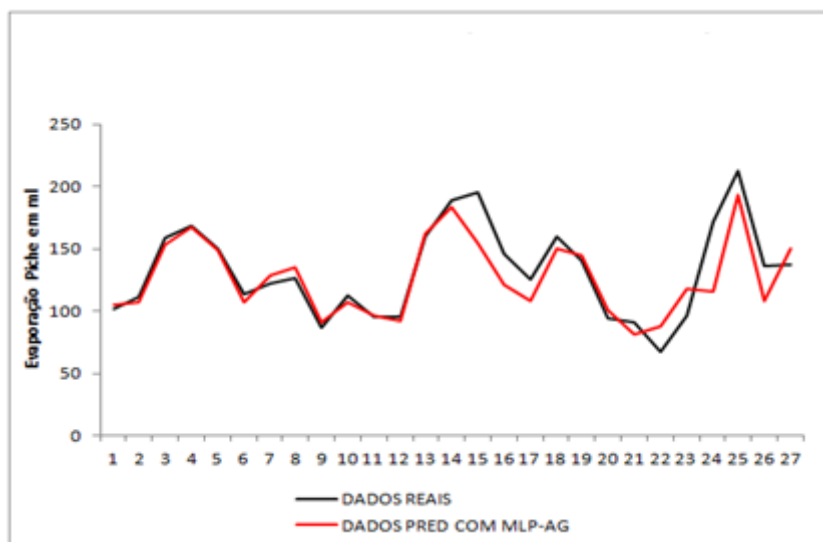


Figura 31: Dados de evaporação estimados com o modelo MLP-AG para a estação de Resende.

O conjunto de dados definidos pelo AG permitiu o preenchimento de 100% das falhas, completando as 10 lacunas ocorridas na série histórica de evaporação pertencente a região de Resende. Isto pode ser observado na Figura 32 pela cor vermelha (ANEXO E).

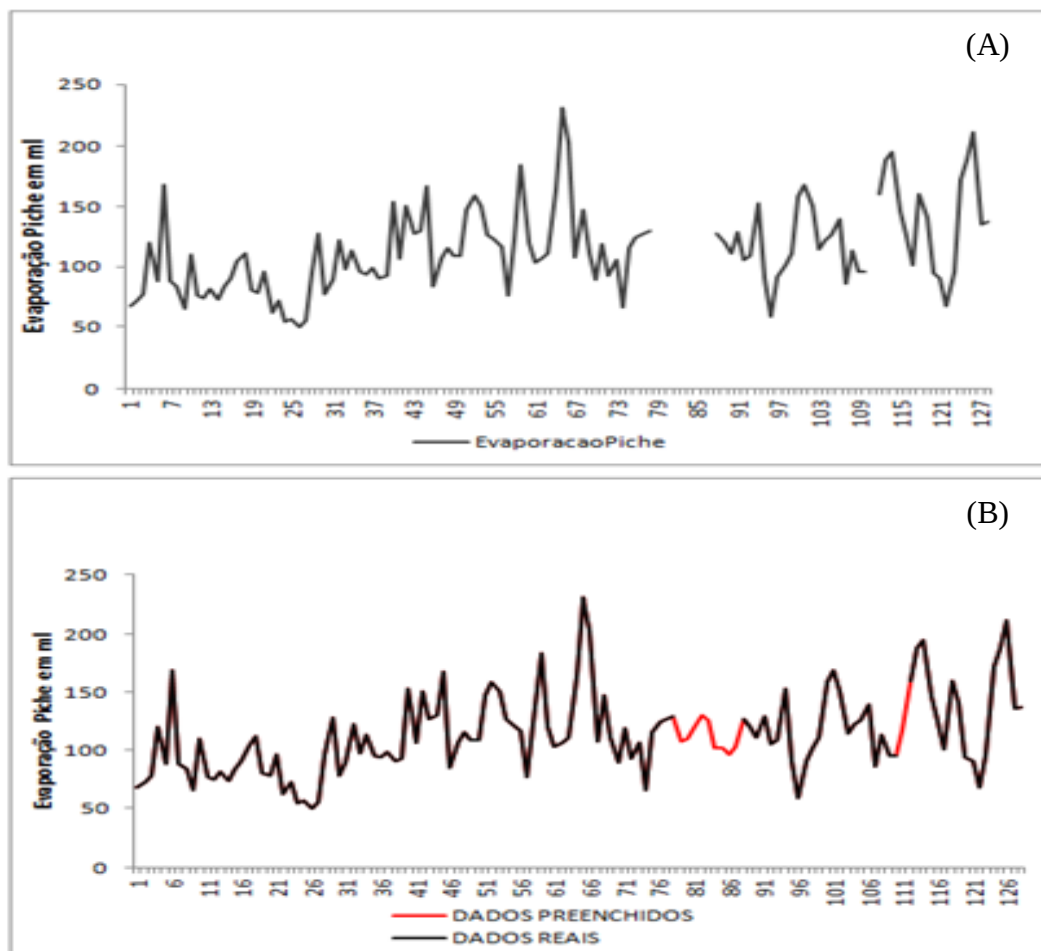


Figura 32: Estação de Resende, (A) Dados de evaporação com falhas, (B) Dados de evaporação preenchidos.

6.3. Resultados de Temperatura Máxima do Ar

Para comparar os conjuntos de estações estimadoras escolhidas de forma manual com os conjuntos definidos pelo modelo híbrido, foram realizadas as correlações entre todas as estações. A matriz de correlações apresentada para variável de temperatura máxima do ar é demonstrada na Tabela 15.

Tabela 15: Índice de correlação apresentado para as variáveis de temperatura máxima do ar.

Índice de correlação das variáveis de Temperatura Máxima do Ar						
<i>Estações</i>	<i>CAM</i>	<i>COR</i>	<i>ITA</i>	<i>PAT</i>	<i>RES</i>	<i>RJ</i>
Campos dos Goytacazes	1					
Cordeiro	0,95	1				
Itaperuna	0,93	0,98	1			
Paty do Alferes	0,94	0,99	0,97	1		
Resende	0,87	0,94	0,92	0,95	1	
Rio de Janeiro	0,96	0,94	0,92	0,94	0,91	1

Onde :

- CAM = Correlação entre as regiões e Campos dos Goytacazes.
- COR = Correlação entre as regiões e Cordeiro.
- ITA= Correlação entre as regiões e Itaperuna.
- PAT= Correlação entre as regiões e Paty do Alferes.
- RES= Correlação entre as regiões e Resende.
- RJ= Correlação entre as regiões e Rio de Janeiro.

Os resultados apresentados para temperatura máxima do ar foram gerados a partir do conjunto de dados de treinamento e validação, que foram divididos em 75% e 25% dos dados. Respectivamente, cada conjunto de dados pertencentes à variável de temperatura máxima do ar após o pré-processamento ficou com 90 dados, sendo 67 dados para treinamento dos modelos e 23 para validação.

6.3.1. Análise do Conjunto de Dados Estimadores para Variável de Temperatura Máxima do Ar

Esta análise ocorreu por meio da comparação das correlações apresentadas por cada conjunto de dados definidos para estimar os valores de temperatura máxima do ar, sejam eles definidos de forma manual ou pelo modelo híbrido. As definições adotadas pelo modelo híbrido para variável de temperatura máxima do ar podem ser visualizadas na Tabela 16.

Tabela 16: Conjunto de variáveis estimadoras de temperatura máxima do ar definidas com o AG.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>	<i>4º estimador</i>	<i>5º estimador</i>
CAM	COR	RJ	X	X	X
COR	CAM	ITA	PAT	X	X
ITA	CAM	COR	RJ	X	X
RJ	CAM	ITA	X	X	X
PAT	CAM	COR	ITA	RES	X
RES	PAT	X	X	X	X

Onde :

- CAM= Campos dos Goytacazes.
- COR= Cordeiro.
- ITA= Itaperuna.
- RJ= Rio de Janeiro.
- PAT= Paty do Alferes.
- RES= Resende.
- X= Não estão sendo utilizados.

A princípio, analisando a Tabela 16 nota-se uma vantagem da metodologia híbrida MLP-AG sobre o método manual de escolha dos estimadores. O método híbrido utilizou em certas predições, como em Campos dos Goytacazes, apenas dados de duas estações, ou como em Resende, onde utilizou apenas dados de Paty do Alferes. Isso demonstra uma otimização, diminuindo a quantidade de dados para se chegar em uma estimativa aceitável.

No entanto, as escolhas determinadas pelo modelo híbrido não podem ser consideradas como escolhas ótimas. Tomando por base o coeficiente de correlação obtido como critério de avaliação das variáveis adotadas para estimação de temperatura máxima do ar, pode-se analisar e comparar a qualidade das escolhas definidas de forma manual com as escolhas adotadas pelo AG. Segue abaixo as matrizes de coeficiente de correlação nas Tabelas 17 e 18.

Tabela 17: Índices de correlação dos conjuntos de estimadores definidos manualmente para variável de temperatura máxima do ar.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>
CAM	0,95	0,93	0,96
COR	0,94	0,98	0,95
ITA	0,92	0,98	0,93
RJ	0,94	0,91	0,94
PAT	0,94	0,95	0,99
RES	0,91	0,95	0,94

Tabela 18: Índices de correlação dos conjuntos de estimadores definidos pelo AG para variável de temperatura máxima do ar.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>	<i>4º estimador</i>	<i>5º estimador</i>
CAM	0,95	0,96	X	X	X
COR	0,95	0,98	0,99	X	X
ITA	0,93	0,98	0,92	X	X
RJ	0,96	0,92	X	X	X
PAT	0,94	0,99	0,97	0,95	X
RES	0,95	X	X	X	X

Na Tabela 17 foi visto que por meio da escolha manual não foi possível determinar os melhores conjuntos de estimadores, segundo o coeficiente de correlação. Entretanto, a Tabela 18 pertencente ao modelo híbrido de RNA com AG fez quatro combinações ótimas minimizando ou maximizando a quantidade de dados utilizados, como em Campos dos Goytacazes, Cordeiro, Paty do Alferes e Resende, onde optou pela utilização das regiões com maiores graus de relação linear (r).

6.3.2. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Campos dos Goytacazes

Alguns parâmetros adotados pelo modelo híbrido MLP-AG são semelhantes aos fixados manualmente para o modelo MLP. O modelo MLP-AG utilizou em sua primeira camada intermediária uma função *tansig*, na segunda camada intermediária foi empregada uma função *satlin*, e por último o modelo aplicou o algoritmo de treinamento *trainbfg*, junto ao fator de auxílio (λ). Mais informações no APÊNDICE B.

Os resultados obtidos com a aplicação dos modelos para estimativa dos 23 meses de temperatura máxima do ar da estação de Campos dos Goytacazes podem ser visualizados na Tabela 19.

Tabela 19: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da estação de Campos dos Goytacazes.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(°C)	EMP(%)
MED	0,97	0,56	0,98	0,96	Ótimo	0,47	1,57
RLM	0,98	0,48	0,99	0,97	Ótimo	0,37	1,24
MLP	0,98	0,46	0,99	0,97	Ótimo	0,36	1,19
RBF	0,98	0,46	0,99	0,97	Ótimo	0,37	1,25
MLP-AG	0,99	0,44	0,99	0,98	Ótimo	0,32	1,09

Analisando a Tabela 19, é possível verificar que todos os modelos apresentaram taxas de erro muito próximas. Entretanto, o modelo MLP-AG se destacou por possuir um (r) quase perfeito, chegando bem próximo de 1, e por demonstrar um ajuste apurado do modelo tendo erros baixos como, EMA igual 0,32°C, REMQ igual a 0,44 e EMP igual a 1,09%. Após todas as análises, foi possível determinar o modelo MLP-AG como solução para o problema. Os dados estimados por esse modelo podem ser vistos na Figura 33.

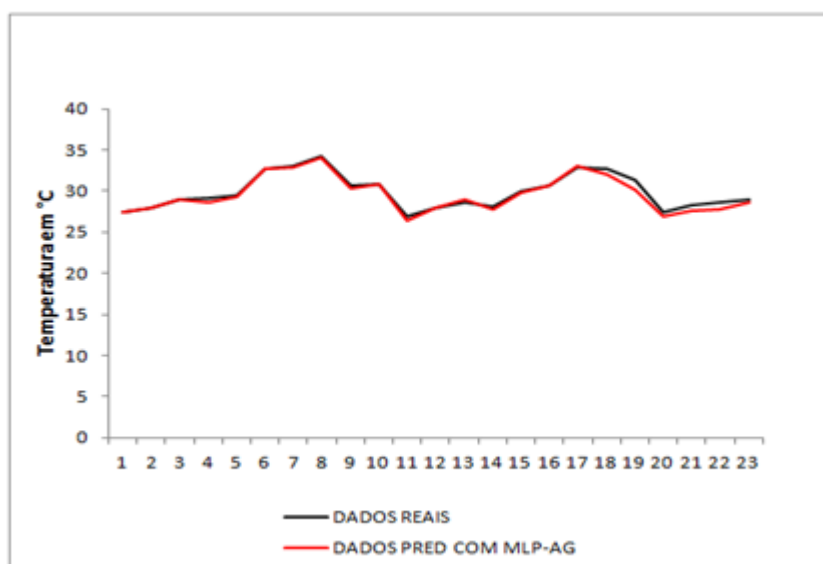


Figura 33: Dados estimados de temperatura máxima do ar obtidos com MLP-AG para estação de Campos dos Goytacazes.

Uma vantagem do modelo MLP-AG sobre os demais modelos foi que o mesmo permitiu o preenchimento de 100% das falhas ocorridas na série histórica de temperatura máxima do ar desta região (Figura 34), o que não aconteceria com a utilização dos outros modelos, pois os dados definidos de forma manual possuem uma falha na estação de Itaperuna na data 31/03/2009, que também deve ser preenchida na estação de Campos dos Goytacazes. Mais informações nos ANEXOS A e C.

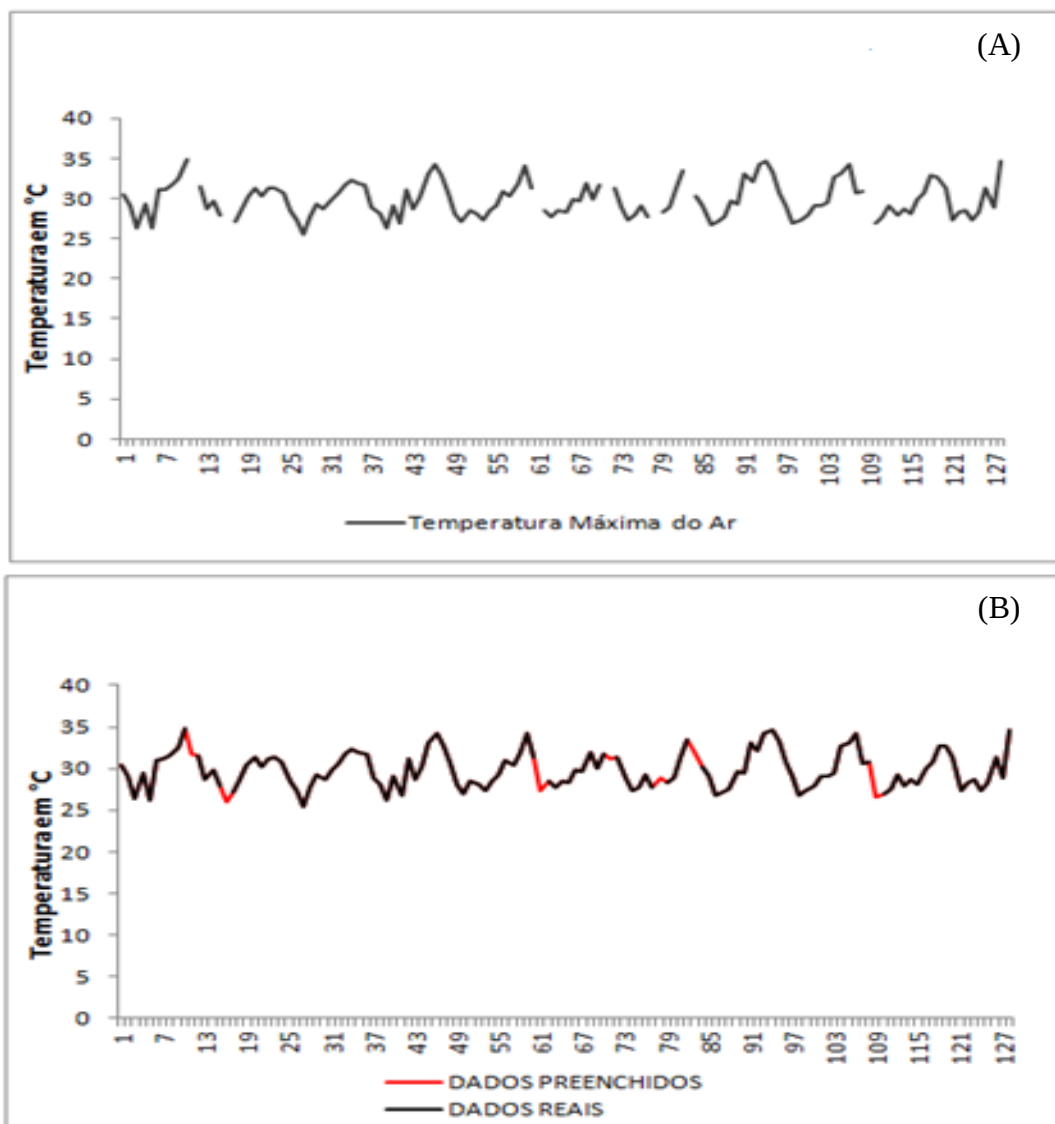


Figura 34: Estação de Campos dos Goytacazes, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.

6.3.3. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Cordeiro

Nesta estação o modelo MLP-AG maximizou seu conjunto de dados de treinamento utilizando (λ), aplicou as funções *satlin* em suas duas camadas intermediárias junto ao algoritmo de treinamento *trainrp*. Mais informações no APÊNDICE B. Os resultados obtidos com a aplicação dos modelos para estimativa dos 23 meses de temperatura máxima do ar da estação de Cordeiro podem ser visualizados na Tabela 20.

Tabela 20: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da região de Cordeiro.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(°C)	EMP(%)
MED	0,97	0,69	0,97	0,94	Ótimo	0,52	1,92
RLM	0,98	0,43	0,99	0,97	Ótimo	0,38	1,43
MLP	0,98	0,41	0,99	0,97	Ótimo	0,34	1,25
RBF	0,97	0,52	0,98	0,96	Ótimo	0,41	1,53
MLP-AG	0,99	0,38	0,99	0,98	Ótimo	0,31	1,15

Os resultados demonstraram que os modelos conseguiram prever os dados de temperatura máxima do ar de forma satisfatória. Apesar da diferença apresentada pelos modelos ser aparentemente pequena, o modelo MLP-AG apresentou a maior relação entre as variáveis preditas e esperadas com um (r) igual a 0,99 quase atingindo a perfeição. Junto a isso, ele também atingiu os menores coeficientes de erro como EMA (0,31°C), REMQ (0,38), e EMP (1,15%), caracterizando a estimativa de cada dado em 98,75% de acerto (Figura 35).

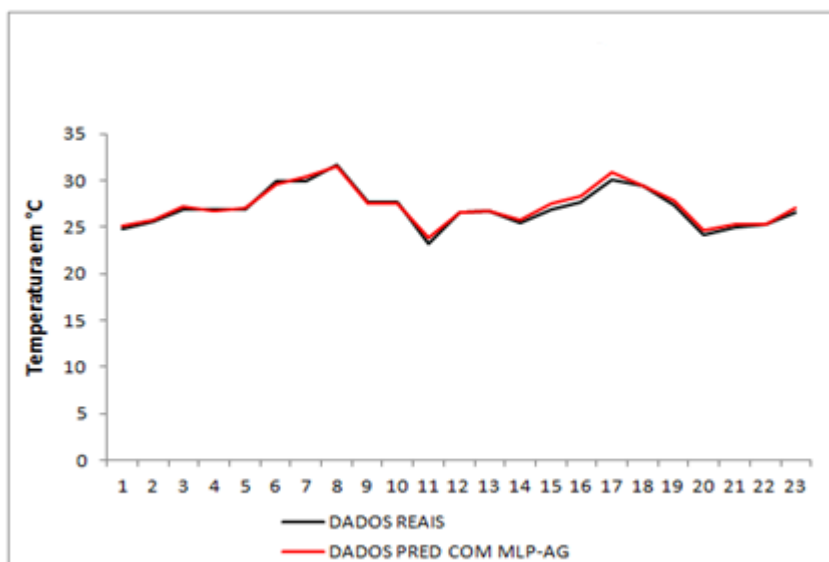


Figura 35: Resultados da estimativa de temperatura máxima do ar obtidos com o modelo MLP-AG para estação de Cordeiro.

O MLP-AG permitiu o preenchimento de 100% das falhas ocorridas na série histórica de temperatura máxima do ar desta estação (Figura 36). Isso não aconteceu com a utilização dos outros modelos, pois nos dados definidos de forma manual há uma falha na estação do Rio de Janeiro na data 31/10/2012 que também deve ser preenchida na estação de Cordeiro (ANEXOS B e F).

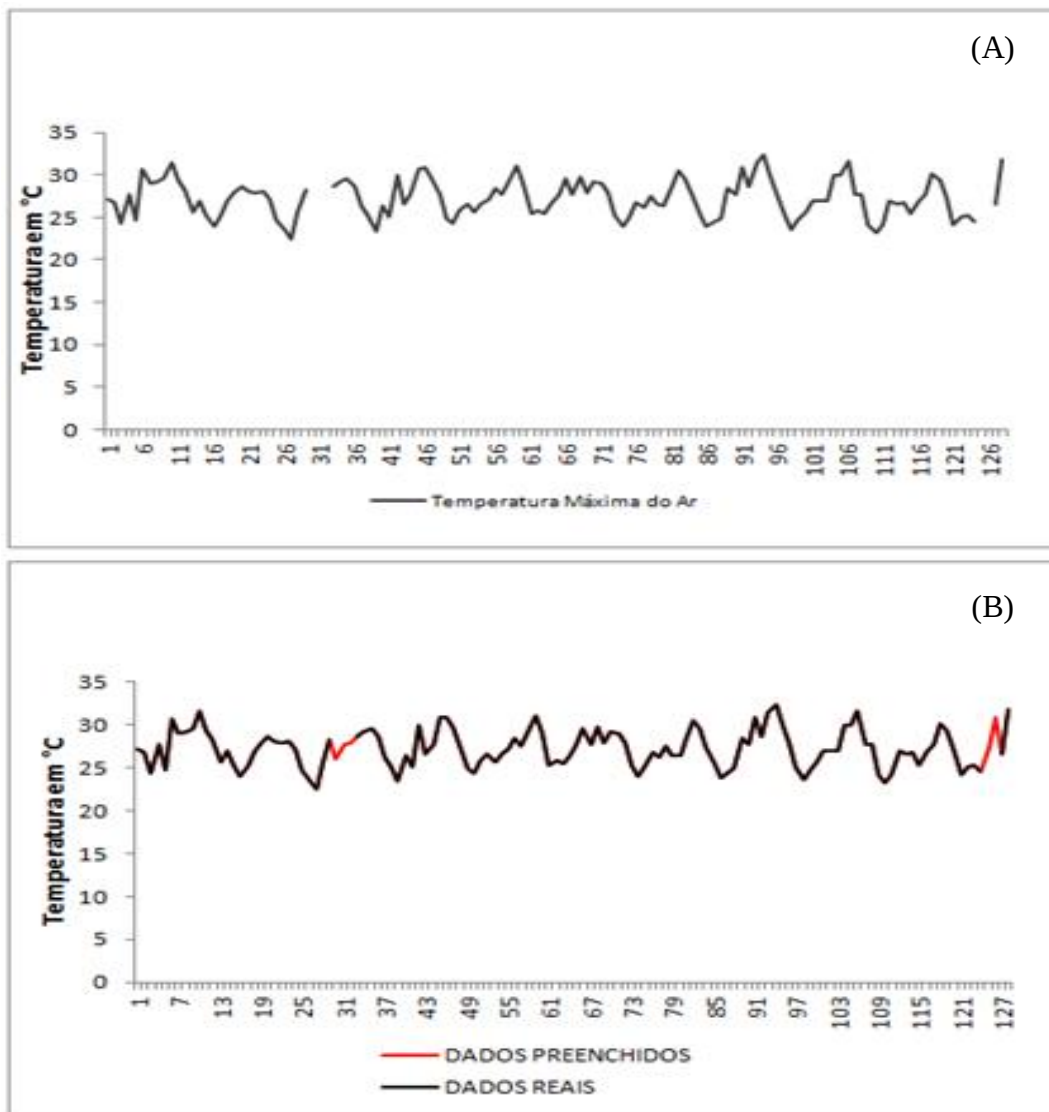


Figura 36: Estação de Cordeiro, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.

6.3.4. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Itaperuna

Para estimar os dados dos 23 meses pertencentes a Itaperuna, o modelo MLP-AG empregou em sua primeira camada intermediária uma função *logsig*, na segunda camada intermediária uma função *satlin* e o algoritmo de treinamento *trainbfg*. Mais informações no APÊNDICE B.

Os resultados obtidos com a aplicação dos modelos na estimativa de temperatura máxima do ar da estação de Itaperuna podem ser visualizados na Tabela 21.

Tabela 21: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da região de Itaperuna.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(°C)	EMP(%)
MED	0,96	0,66	0,98	0,94	Ótimo	0,54	1,80
RLM	0,97	0,49	0,99	0,96	Ótimo	0,41	1,35
MLP	0,98	0,47	0,99	0,97	Ótimo	0,36	1,21
RBF	0,97	0,49	0,98	0,95	Ótimo	0,40	1,32
MLP-AG	0,98	0,44	0,99	0,97	Ótimo	0,32	1,07

Os resultados apresentaram elevados índices de (r), (D) e (C), o que manteve o desempenho dos modelos em ótimo. Porém, nota-se que o modelo que apresentou as menores taxas de erro foi MLP-AG, com EMA (0,32), REMQ (0,44) e EMP (1,07%), o que representa cerca de 98,93% de acerto em suas estimações (Figura 37).

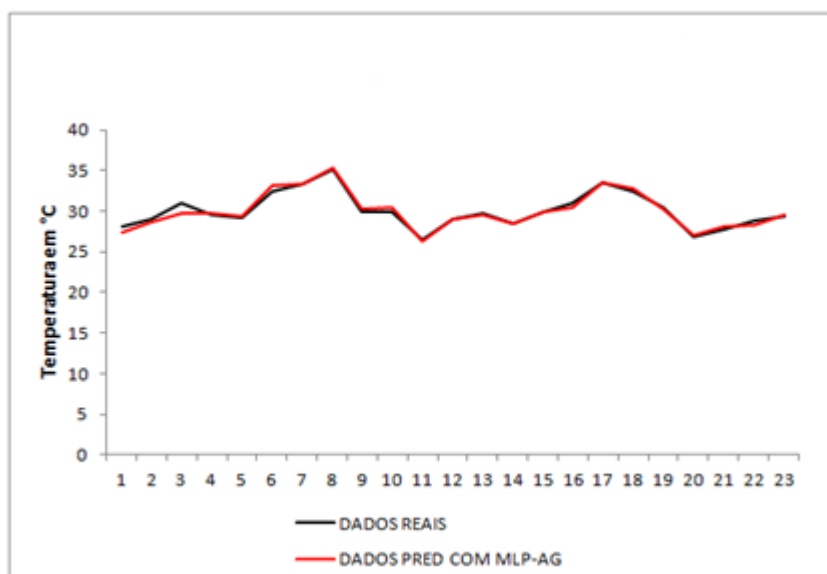


Figura 37: Resultados da estimativa de temperatura máxima do ar obtidos com o modelo MLP-AG para estação de Itaperuna.

O conjunto de dados definidos pelo modelo MLP-AG, assim como o conjunto definido manualmente, utilizou dados da estação de Campos dos Goytacazes para preenchimento das falhas de Itaperuna. Devido à ocorrência de uma falha em Campos dos Goytacazes na data de 31/03/2009, não foi possível realizar o preenchimento de todas as falhas, sendo possível apenas o preenchimento de 75% dos dados, ou seja, 3 das 4 falhas existentes (Figura 38) (ANEXOS A e C).

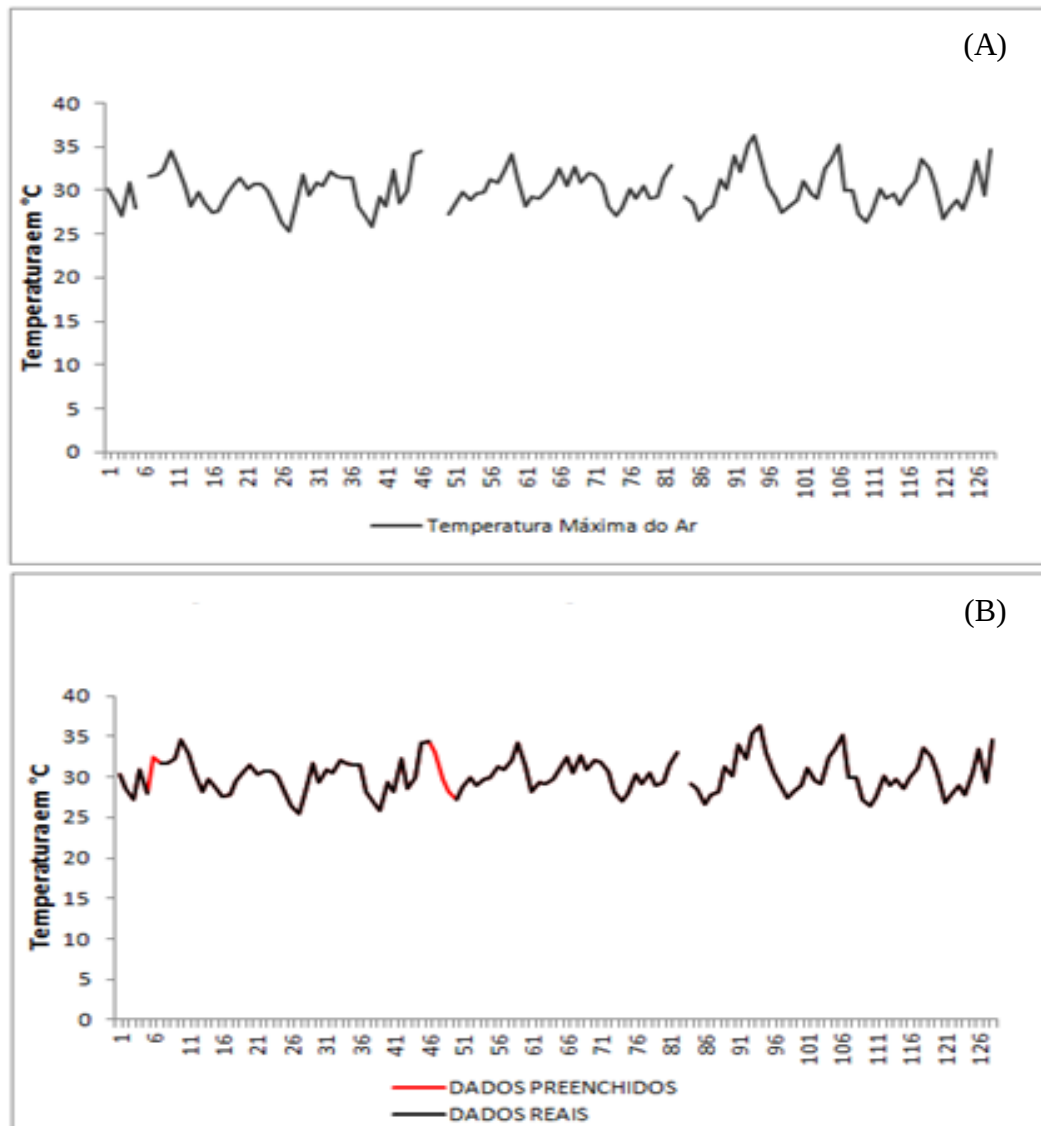


Figura 38: Estação de Itaperuna, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.

6.3.5. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação do Rio de Janeiro

O modelo MLP-AG aplicado na estimativa dos 23 meses de temperatura máxima do ar pertencentes a estação do Rio de Janeiro utilizou em sua primeira camada intermediária uma função *purelin*, na segunda camada intermediária uma função *logsig*, e o algoritmo de treinamento *trainlm* (APÊNDICE B).

Os resultados obtidos com a aplicação dos modelos para estimativa de temperatura máxima do ar da estação do Rio de Janeiro podem ser visualizados na Tabela 22.

Tabela 22: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da estação do Rio de Janeiro.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(°C)	EMP(%)
MED	0,91	1,91	0,85	0,77	Muito Bom	1,62	5,49
RLM	0,91	1,14	0,94	0,86	Ótimo	0,90	3,05
MLP	0,95	0,79	0,97	0,92	Ótimo	0,59	2,01
RBF	0,94	0,89	0,96	0,91	Ótimo	0,69	2,33
MLP-AG	0,97	0,58	0,99	0,96	Ótimo	0,47	1,57

Novamente é possível observar que todos os modelos aplicados geraram bons resultados. No entanto, as RNAs se destacaram obtendo os menores erros e apresentando os maiores (r) entre os dados estimados e os esperados. Contudo, comparando o modelo MLP-AG com as demais RNAs é possível afirmar que o modelo MLP-AG realizou uma estimativa mais apurada devido o mesmo ter obtido os menores erros, como EMA (0,47), REMQ (0,58) e EMP (1,57%), o que define que cada dado estimado por meio deste modelo teve 98,43% de acerto (Figura 39).

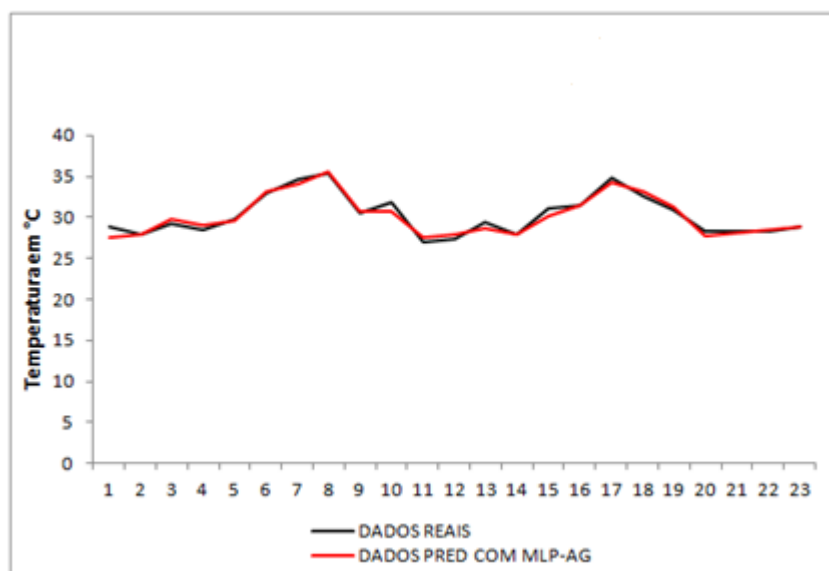


Figura 39: Dados estimados de temperatura máxima do ar com o modelo MLP-AG para estação do Rio de Janeiro.

O modelo MLP-AG apresentou uma grande vantagem no preenchimento dos dados de temperatura máxima do ar da estação do Rio de Janeiro, pois o mesmo conseguiu preencher 100% das falhas ocorridas na série de dados. Isto não se observa com o conjunto de dados definidos manualmente para os demais modelos devido à ocorrência de duas falhas, uma na estação de Paty do Alferes na data 31/10/2007, e outra nas estações de Cordeiro e Resende em 31/10/2012. Mais informações nos ANEXOS B, D, E e F. A Figura 40 demonstra o preenchimento realizado pelo modelo MLP-AG.

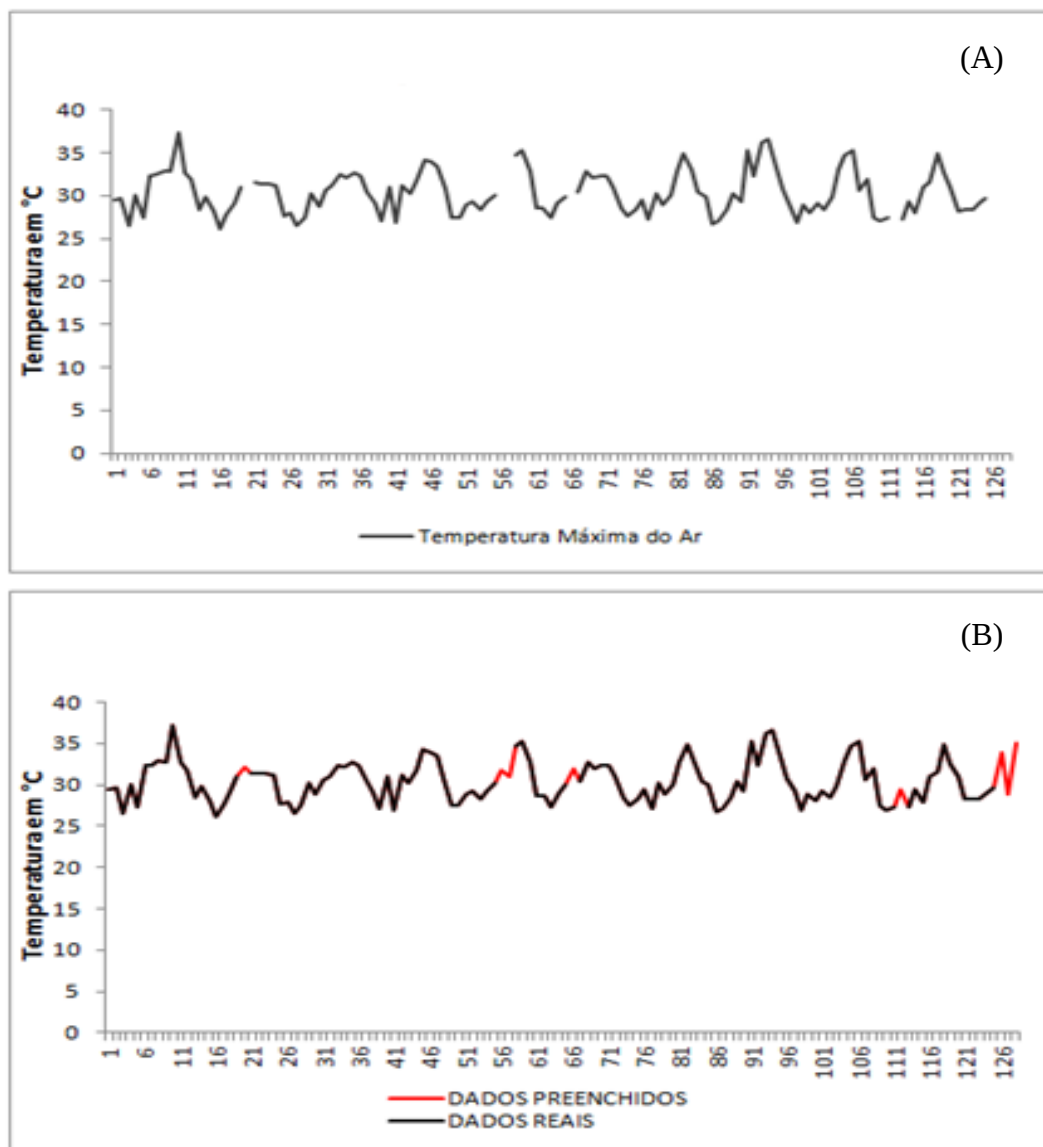


Figura 40: Estação do Rio de Janeiro, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.

6.3.6. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Paty do Alferes

Os resultados apresentados pelo modelo MLP-AG aplicado na estimativa dos 23 meses de temperatura máxima do ar pertencentes a estação de Paty do Alferes foram obtidos por meio das configurações adotadas. Este modelo aplicou em sua primeira camada intermediária uma função *purelin*, na segunda camada intermediária uma função *satlin*, e o algoritmo de treinamento *traingd* (APÊNDICE B).

Os resultados obtidos com a aplicação de todos modelos para estimativa de temperatura máxima do ar podem ser visualizados na Tabela 23.

Tabela 23: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da estação de Paty do Alferes.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(°C)	EMP(%)
MED	0,98	0,41	0,99	0,97	Ótimo	0,33	1,21
RLM	0,99	0,47	0,99	0,98	Ótimo	0,36	1,29
MLP	0,99	0,38	0,99	0,98	Ótimo	0,32	1,17
RBF	0,98	0,43	0,99	0,97	Ótimo	0,31	1,12
MLP-AG	0,99	0,38	0,99	0,98	Ótimo	0,27	0,95

Ao analisar a Tabela 23, verifica-se que todos os modelos apresentaram estimativas satisfatórias, com medidas de desempenho se mantendo em ótimos. No entanto, os modelos de RNAs novamente se destacaram obtendo estimativas mais apuradas que os demais modelos.

Os modelos RBF e MLP-AG alcançaram os menores erros, porém, o modelo MLP-AG apresentou dados mais próximos dos dados reais, tendo erros de EMA igual a 0,27°C, REMQ igual 0,38, e EMP igual a 0,95%. Isto define que cada dado determinado por este modelo obteve cerca de 99,05% de acerto (Figura 41).

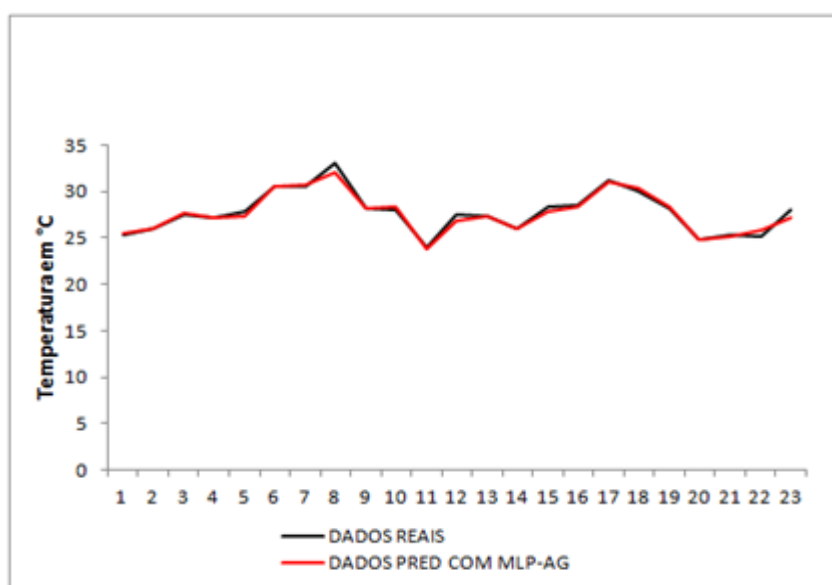


Figura 41: Resultados da estimativa de temperatura máxima do ar obtidos com o modelo MLP-AG para estação de Paty do Alferes.

Da mesma forma que o preenchimento realizado nas outras estações, o modelo MLP-AG apresentou uma vantagem comparado aos demais. As regiões estimadoras definidas por este modelo permitiram o preenchimento de 5 das 7 falhas ocorridas, o que não aconteceria com os demais modelos que só poderiam preencher 4 das 7 falhas, porque, além de fazer uso dos dados de Resende que possuem falhas em dois períodos (31/07/2009 e 31/08/2012), o mesmo também utilizaria os dados do Rio de Janeiro que possui uma falha em 31/10/2007. Portanto, pode-se inferir que o modelo MLP-AG conseguiu preencher cerca de 71% das falhas ocorridas (Figura 42) (ANEXOS D, E e F).

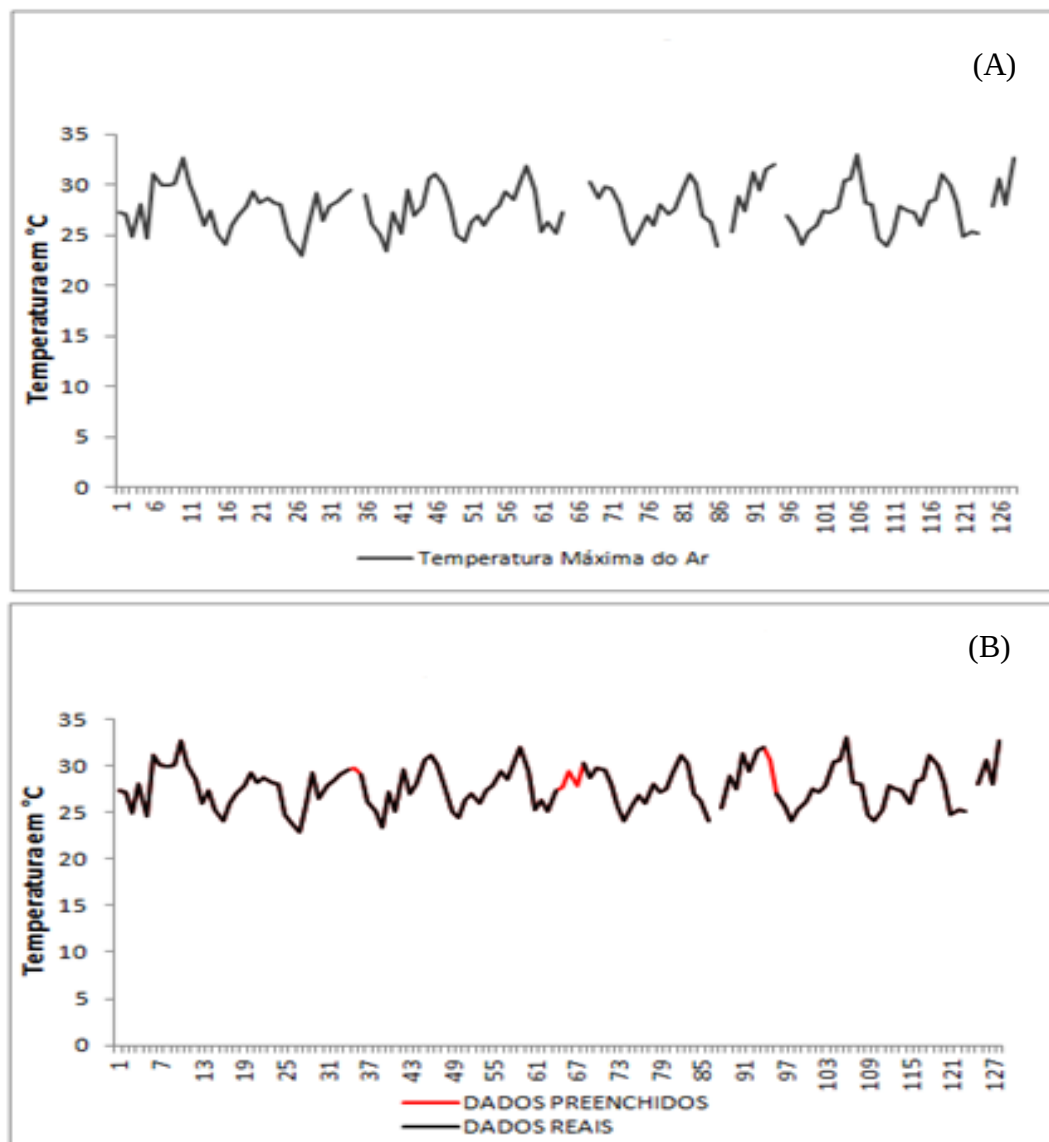


Figura 42: Estação de Paty do Alferes, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.

6.3.7. Resultados e Preenchimento das Falhas de Temperatura Máxima do Ar para estação de Resende

Na estação de Resende, o modelo MLP-AG aplicado na estimativa dos 23 meses de temperatura máxima do ar usou em suas duas camadas intermediárias a função *purelin*, e o algoritmo de treinamento *trainrp* (APÊNDICE B). Os resultados obtidos com a aplicação dos modelos na estimativa de dados de temperatura máxima do ar da estação de Resende podem ser vistos na Tabela 24.

Tabela 24: Resultados obtidos com os modelos aplicados na estimativa de temperatura máxima do ar da estação de Resende.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(°C)	EMP(%)
MED	0,95	2,09	0,83	0,79	Muito Bom	1,96	6,99
RLM	0,95	1,10	0,94	0,90	Ótimo	0,89	3,16
MLP	0,97	0,72	0,98	0,95	Ótimo	0,56	2,00
RBF	0,96	0,95	0,96	0,92	Ótimo	0,74	2,65
MLP-AG	0,97	0,57	0,98	0,95	Ótimo	0,43	1,55

Os resultados demonstraram que os modelos conseguiram estimar os dados de temperatura máxima do ar de forma satisfatória. Entretanto, os modelos de RNAs, MLP, RBF, MLP-AG, possuem uma estimativa mais apurada, apesar da diferença apresentada por esses modelos ser pequena. No entanto, o modelo MLP-AG obteve uma maior precisão em seus resultados, alcançando taxas de erro muito pequenas como EMA igual a 0,43°C, REMQ igual a 0,57 e EMP 1,55%, o que possibilitou afirmar que o modelo conseguiu estimar cada dado com uma taxa de acerto de 98,45% (Figura 43).

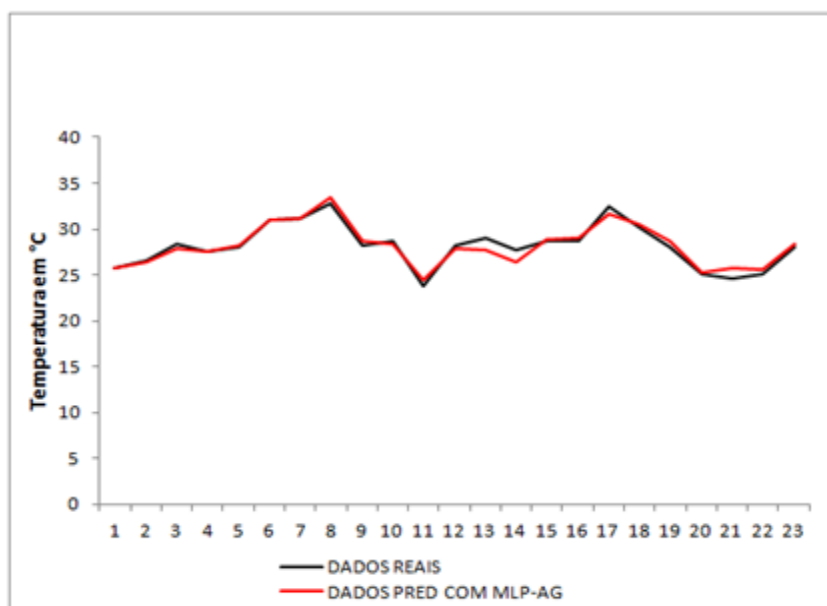


Figura 43: Resultados de estimativa de temperatura máxima do ar obtidos com o modelo MLP-AG para estação de Resende.

Assim como no preenchimento realizado nas outras estações, esse modelo também apresentou vantagens consideráveis em se tratando do conjunto de dados definidos pelo AG. Eles possibilitaram o preenchimento de 15 das 17 falhas ocorridas nesta estação, o que não seria possível a partir da utilização dos demais modelos que preencheriam cerca de 14 das 17 falhas ocorridas, pois os mesmos possuem falhas em 31/07/2009 e 31/12/2012 na estação de Paty do Alferes, e outra falha em 31/10/2012 na estação do Rio de Janeiro. Por isso, pode-se afirmar que o modelo MLP-AG permitiu o preenchimento de 89 % dos dados falhados da estação de Resende (Figura 44) (ANEXOS D, E e F).

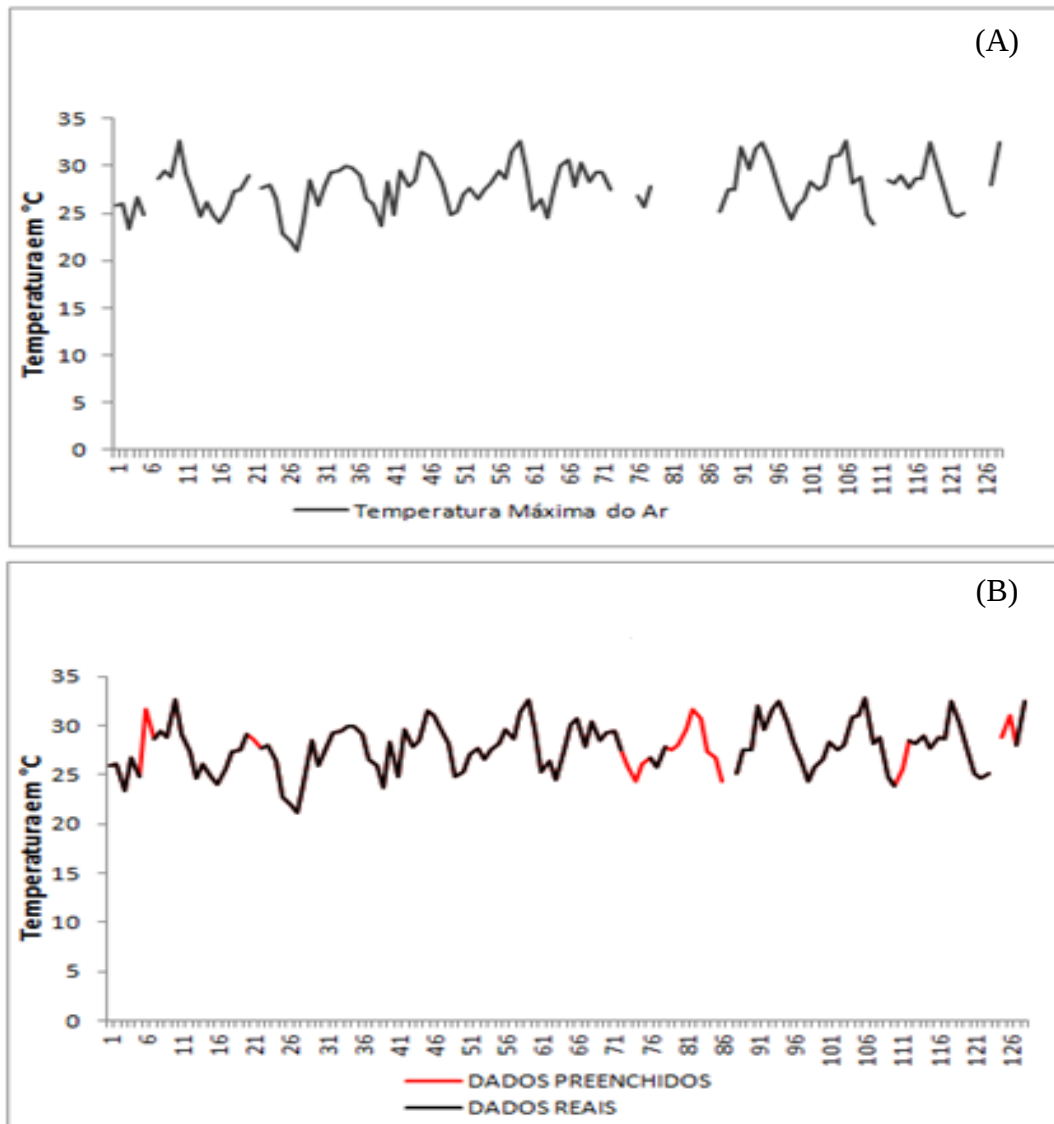


Figura 44: Estação de Resende, (A) Dados de temperatura máxima do ar com falhas, (B) Dados de temperatura máxima do ar preenchida.

6.4. Resultados de Umidade Relativa do Ar

Para comparar os conjuntos de estações estimadoras escolhidas de forma manual com os conjuntos definidos pelo modelo híbrido, foram feitas as correlações entre as variáveis de umidade relativa do ar. A matriz de correlações é demonstrada na Tabela 25.

Tabela 25: Índice de correlação apresentado para as variáveis de umidade relativa do ar.

Índice de correlação das variáveis de Umidade Relativa do Ar						
Estações	CAM	COR	ITA	PAT	RES	RJ
Campos dos Goytacazes	1					
Cordeiro	0,76	1				
Itaperuna	0,83	0,90	1			
Paty do Alferes	0,69	0,86	0,79	1		
Resende	0,61	0,76	0,75	0,69	1	
Rio de Janeiro	0,62	0,67	0,69	0,64	0,64	1

Onde :

- CAM = Correlação entre as regiões e Campos dos Goytacazes.
- COR = Correlação entre as regiões e Cordeiro.
- ITA= Correlação entre as regiões e Itaperuna.
- PAT= Correlação entre as regiões e Paty do Alferes.
- RES= Correlação entre as regiões e Resende.
- RJ= Correlação entre as regiões e Rio de Janeiro.

Os resultados apresentados neste tópico para variável de umidade relativa do ar foram gerados a partir dos dados do conjunto de treinamento e validação. Respectivamente, cada conjunto de dados pertencentes à variável de umidade relativa do ar, após o pré-processamento, ficou com 64 dados, sendo 75% para treinamento dos modelos cerca de 48 dados, e 25% para validação cerca de 16.

6.4.1. Análise do Conjunto de Dados Estimadores para Variável de Umidade Relativa do Ar

Por meio das correlações apresentadas, foi realizada uma análise de cada conjunto de estimadores a fim de comparar a qualidade dos conjuntos definidos de forma manual com os definidos pelo modelo híbrido. O conjunto de estimadores definidos pelo modelo híbrido para variável de umidade relativa do ar podem ser visualizados na Tabela 26.

Tabela 26: Conjunto de variáveis estimadoras de umidade relativa do ar definidas com o AG.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>	<i>4º estimador</i>	<i>5º estimador</i>
CAM	ITA	RJ	PAT	X	X
COR	ITA	RJ	X	X	X
ITA	CAM	COR	PAT	X	X
RJ	CAM	COR	ITA	PAT	X
PAT	COR	ITA	RES	X	X
RES	CAM	COR	PAT	X	X

Onde :

- CAM = Campos dos Goytacazes.
- COR = Cordeiro.
- ITA= Itaperuna.
- RJ= Rio de Janeiro.
- PAT= Paty do Alferes.
- RES= Resende.
- X= Não estão sendo utilizados.

Tendo como base o coeficiente de correlação obtido como critério de avaliação das variáveis adotadas para estimação de umidade relativa do ar, pode-se analisar e comparar a qualidade das escolhas definidas de forma manual com as escolhas adotadas pelo AG. Segue abaixo as matrizes de coeficiente de correlação nas Tabelas 27 e 28.

Tabela 27: Índices de correlação dos conjuntos de estimadores definidos manualmente para variável de umidade relativa do ar.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>
Campos dos Goytacazes	0,76	0,83	0,62
Cordeiro	0,67	0,90	0,76
Itaperuna	0,69	0,90	0,83
Rio de Janeiro	0,64	0,64	0,67
Paty do Alferes	0,64	0,69	0,86
Resende	0,64	0,69	0,76

Tabela 28: Índices de correlação dos conjuntos de estimadores definidos pelo AG para variável de umidade relativa do ar.

<i>Estações</i>	<i>1º estimador</i>	<i>2º estimador</i>	<i>3º estimador</i>	<i>4º estimador</i>	<i>5º estimador</i>
CAM	0,83	0,62	0,69	X	X
COR	0,90	0,67	X	X	X
ITA	0,83	0,90	0,79	X	X
RJ	0,62	0,67	0,69	0,64	X
PAT	0,86	0,79	0,69	X	X
RES	0,61	0,76	0,69	X	X

Analisando a Tabela 27, verifica-se que o conjunto de estimadores definidos de forma manual não conseguiu determinar as melhores variáveis para estimação dos dados, segundo o coeficiente de correlação. Entretanto, na Tabela 28 pertencente ao modelo híbrido de RNA com AG, identifica-se três combinações ótimas como em Itaperuna, Rio de Janeiro e Paty do Alferes, onde foram utilizadas as estações com os maiores coeficientes de correlação.

6.4.2. Resultados e Preenchimento das falhas de Umidade Relativa do Ar para estação de Campos dos Goytacazes

O modelo MLP-AG aplicado na estimativa dos 16 meses de umidade relativa do ar pertencentes a estação de Campos dos Goytacazes aumentou seu conjunto de dados de treinamento através da inserção do fator de auxílio (λ), utilizou em sua primeira camada intermediária uma função *purelin*, na segunda camada uma função *satlin*, e o algoritmo de treinamento *trainbfg*. Mais informações no APÊNDICE C.

Os resultados obtidos com a aplicação dos modelos para estimativa dos 16 meses de umidade relativa do ar da estação de Campos dos Goytacazes podem ser visualizados na Tabela 29.

Tabela 29: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação de Campos dos Goytacazes.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(%)	EMP(%)
MED	0,79	3,24	0,78	0,61	Mediano	2,64	3,45
RLM	0,80	1,94	0,88	0,71	Bom	1,46	1,92
MLP	0,87	1,56	0,92	0,80	Muito Bom	1,07	1,41
RBF	0,81	1,86	0,88	0,71	Bom	1,40	1,84
MLP-AG	0,89	1,39	0,94	0,84	Muito Bom	1,03	1,35

Através da Tabela 29, constatou-se que os modelos de RNAs MLP e MLP-AG apresentaram os melhores resultados. No entanto, o modelo MLP-AG conseguiu atingir erros menores que os do modelo MLP, como EMA (1,03%), REMQ (1,39%) e EMP (1,35%). Com isso, pode-se afirmar que o modelo MLP-AG conseguiu estimar cada um dos 16 meses com uma qualidade de 98,65% de acerto. A Figura 45 demonstra os dados reais e os estimados pelo modelo MLP-AG.

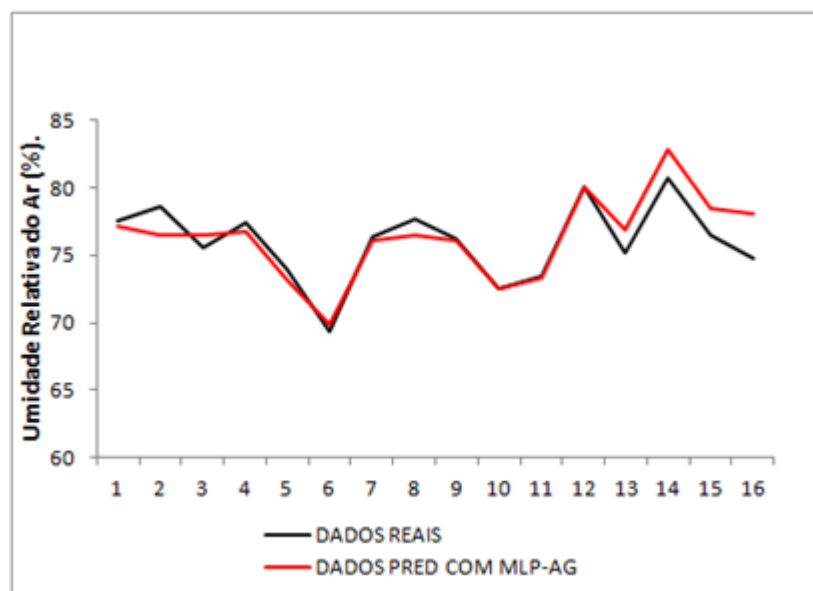


Figura 45: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Campos dos Goytacazes.

Similar ao preenchimento realizado nas outras estações, esse modelo também apresentou vantagens consideráveis em relação ao conjunto de estimadores definidos pelo AG. Os mesmos conseguiram preencher 4 das 5 falhas ocorridas nesta estação, e isso não seria possível utilizando os demais modelos onde só poriam ser preenchidas 3 das 5 falhas existentes, pois os mesmos possuem falhas em 31/08/2006 em Itaperuna e outra falha em 31/05/2009 em Cordeiro (ANEXOS A, B e C). Por isso, pode-se afirmar que o modelo MLP-AG permitiu o preenchimento de 80 % das falhas de umidade relativa do ar desta estação (Figura 46).

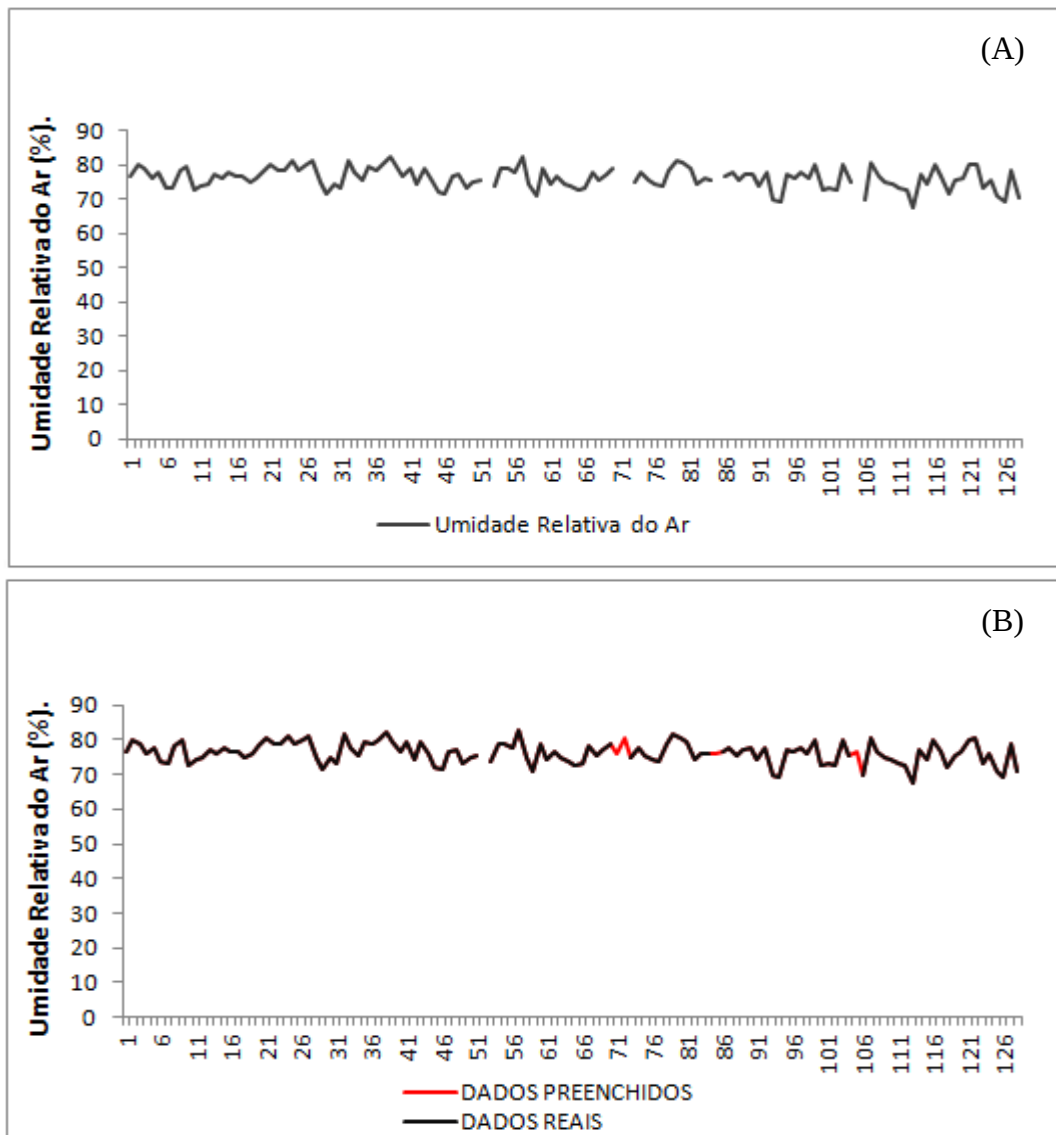


Figura 46: Estação de Campos dos Goytacazes, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.

6.4.3. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação de Cordeiro

Para realizar a estimativa dos dados de umidade relativa do ar pertencente a estação de Cordeiro, o modelo MLP-AG utilizou em sua primeira camada intermediária uma função *purelin*, na segunda camada uma função *logsig* e o algoritmo de treinamento *trainrp* (APÊNDICE C). Os resultados obtidos com a aplicação dos modelos podem ser visualizados na Tabela 30.

Tabela 30: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação de Cordeiro.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(%)	EMP(%)
MED	0,93	2,89	0,90	0,83	Muito Bom	2,48	3,08
RLM	0,95	1,75	0,97	0,92	Ótimo	1,42	1,78
MLP	0,94	1,74	0,97	0,91	Ótimo	1,24	1,54
RBF	0,93	1,85	0,96	0,90	Ótimo	1,55	1,94
MLP-AG	0,97	1,31	0,98	0,95	Ótimo	1,03	1,28

Os erros demonstram que os modelos apresentaram resultados satisfatórios. No entanto, apesar de haver uma proximidade entre os resultados, o modelo MLP-AG demonstrou um melhor ajuste na estimação dos dados, conseguindo erros menores como EMA igual a 1,03%, REMQ igual a 1,31 e EMP 1,28%. Desta forma, é possível afirmar que o modelo MLP-AG teve 98,72% de acerto em suas estimações (Figura 47).

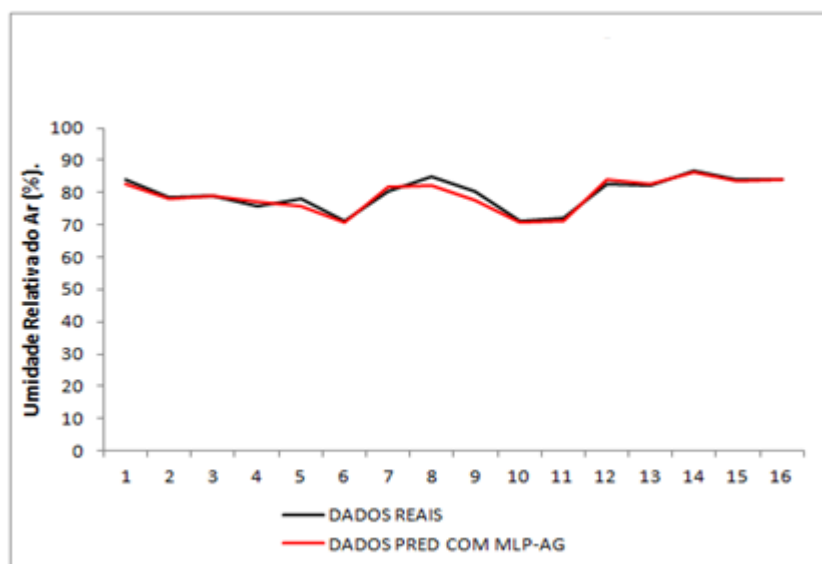


Figura 47: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Cordeiro.

Por meio da combinação de estimadores definidos pelo modelo MLP-AG, fez-se o preenchimento de 22 das 23 falhas ocorridas na série histórica de umidade relativa do ar pertencente a região de Cordeiro. A utilização deste modelo se mostrou mais eficiente do que os demais, pois eles conseguiriam preencher somente 21 das 23 falhas ocorridas. Isso ocorre porque, além de possuir a mesma falha que o conjunto de estimadores utilizados pelo modelo MLP-AG em 31/05/2012 na estação do Rio de Janeiro, o conjunto utilizado pelos demais modelos possui outra falha em 31/05/2009 em Campos dos Goytacazes (ANEXOS A e F). A partir destes fatores, pode-se afirmar que o modelo MLP-AG possibilitou o preenchimento de aproximadamente 96% das falhas ocorridas (Figura 48).

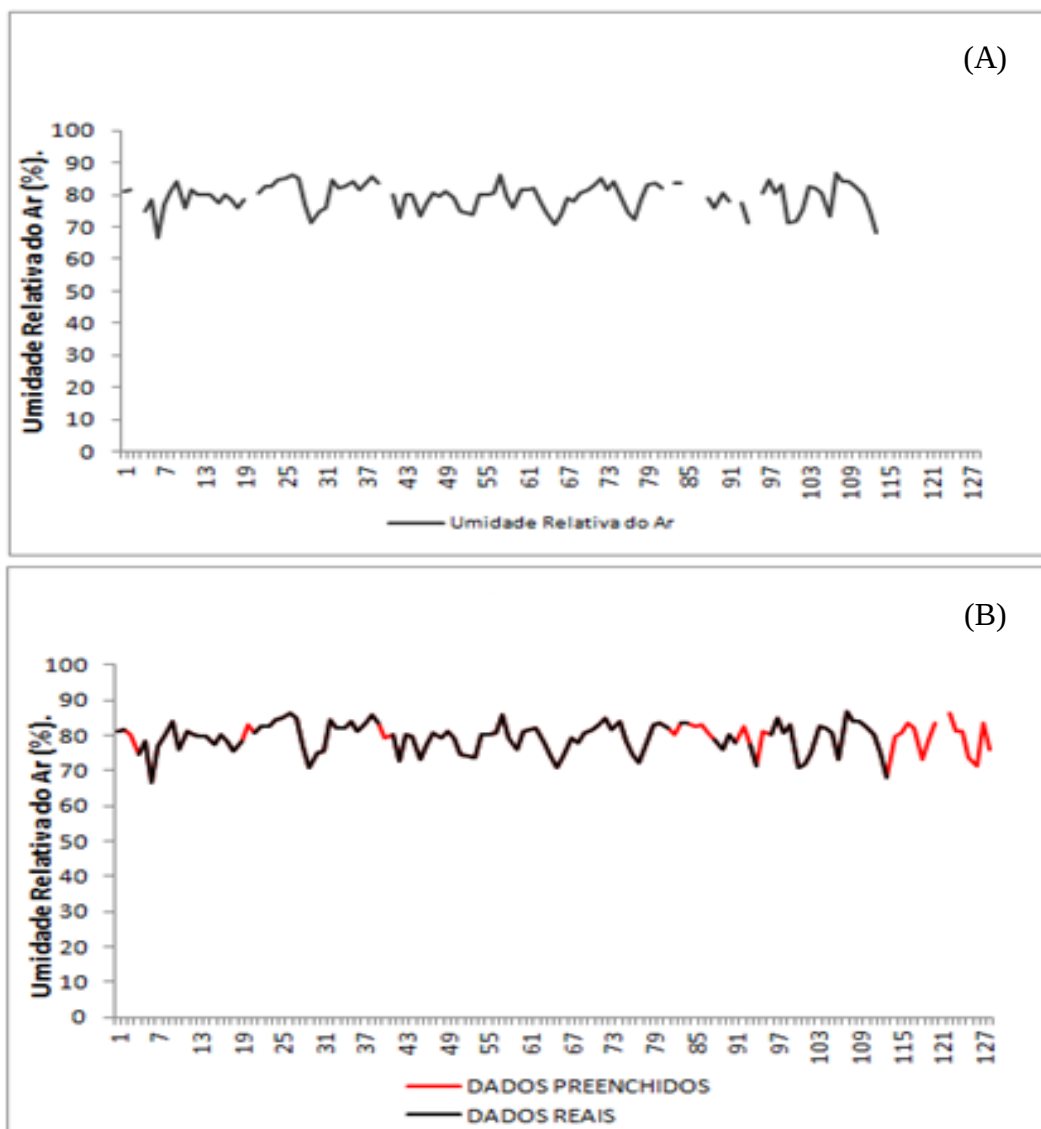


Figura 48: Estação de Cordeiro, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida .

6.4.4. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação de Itaperuna

Para geração dos resultados de Itaperuna, o modelo MLP-AG fez uso do fator de auxílio (λ), empregando em suas camadas intermediárias a função *purelin*, e o algoritmo de treinamento *trainbfg* (APÊNDICE C). Os resultados obtidos com a aplicação dos modelos para estimativa dos 16 meses de umidade relativa do ar da estação de Itaperuna podem ser visualizados na Tabela 31.

Tabela 31: Resultados obtidos com os modelos aplicados na estimação de umidade relativa do ar da estação de Itaperuna.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(%)	EMP(%)
MED	0,95	2,71	0,93	0,88	Ótimo	2,14	3,09
RLM	0,96	2,19	0,96	0,92	Ótimo	1,83	2,61
MLP	0,95	2,06	0,96	0,92	Ótimo	1,64	2,29
RBF	0,94	2,45	0,94	0,89	Ótimo	2,06	2,91
MLP-AG	0,96	1,73	0,98	0,94	Ótimo	1,37	1,89

Os resultados demonstraram que todos os modelos alcançaram ótimos desempenhos, com altos coeficientes (r) chegando perto da perfeição entre os dados preditos e observados. Entretanto, o modelo MLP-AG apresentou uma estimativa com menores taxas de erro, como EMA igual 1,37, REMQ 1,73 e EMP igual a 189%. Com isso, constata-se que o modelo que menos teve variação entre os dados estimados e observados foi o MLP-AG, conseguindo alcançar a taxa de acerto de 98,11% (Figura 49).

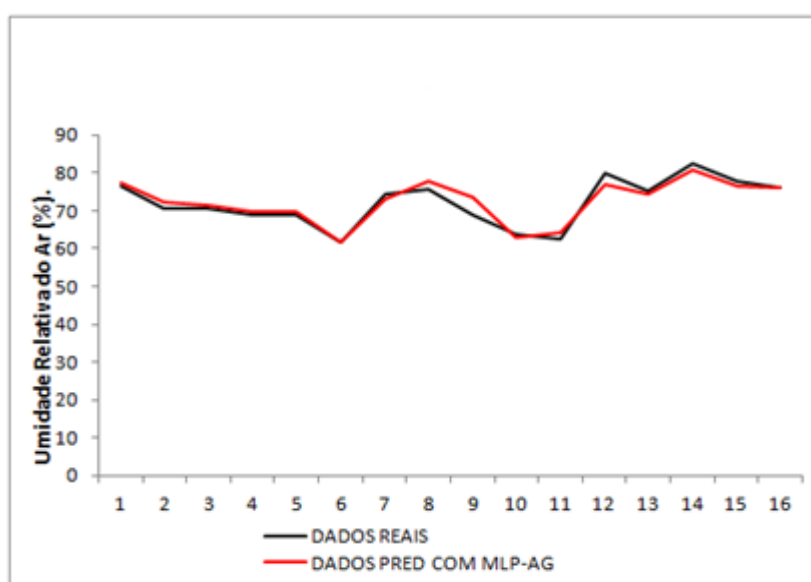


Figura 49: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Itaperuna.

O modelo MLP-AG conseguiu preencher 86% das falhas, ou seja, 6 das 7 falhas ocorridas na série de dados de umidade relativa do ar pertencente a estação de Itaperuna. Contudo, os outros modelos também conseguiriam preencher o mesmo número de falhas, pois a única falha existente no conjunto de estimadores utilizados nos modelos é em 31/08/2006 em Campos dos Goytacazes (ANEXOS A e C). A Figura 50 demonstra o preenchimento realizado pelo modelo MLP-AG.

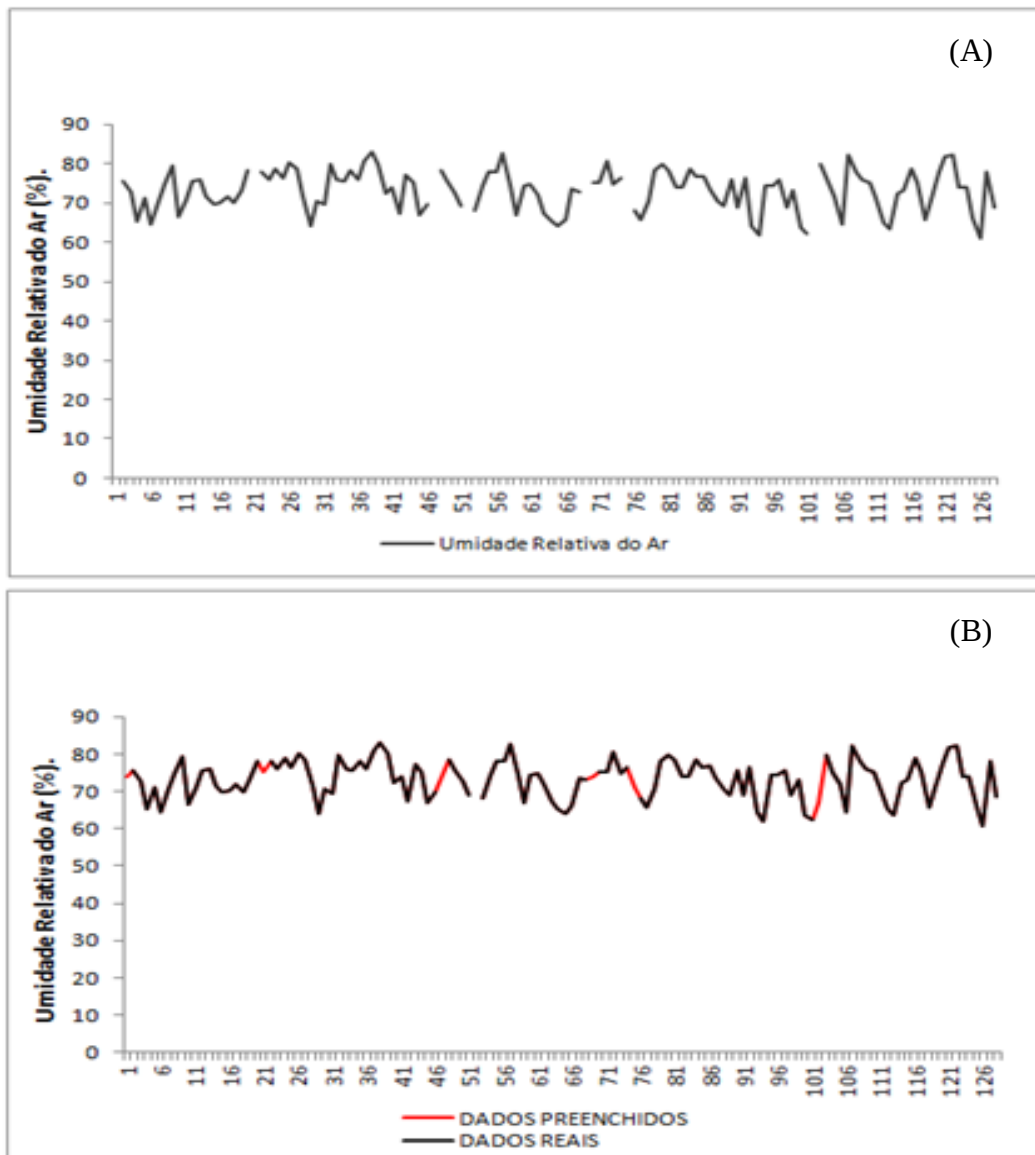


Figura 50: Estação de Itaperuna, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.

6.4.5. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação do Rio de Janeiro

Nesta estação, o modelo MLP-AG adotou em sua primeira camada intermediária uma função *purelin*, na segunda camada uma função *tansig*, e o algoritmo de treinamento *traingd*. As configurações adotadas por todos os modelos estão apresentadas no APÊNDICE C.

Sendo assim, os resultados obtidos com a aplicação dos modelos para estimativa dos 16 meses de umidade relativa do ar da estação do Rio de Janeiro podem ser visualizados na Tabela 32.

Tabela 32: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação do Rio de Janeiro.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(%)	EMP(%)
MED	0,83	2,46	0,86	0,72	Bom	1,72	2,47
RLM	0,83	2,88	0,69	0,57	Sufrível	2,13	2,98
MLP	0,84	2,14	0,89	0,75	Bom	1,61	2,28
RBF	0,81	2,77	0,73	0,59	Sufrível	2,03	2,86
MLP-AG	0,90	1,79	0,93	0,83	Muito Bom	1,36	1,90

A partir da Tabela 32 é possível constatar que o modelo MLP-AG apresentou os melhores resultados, alcançando os maiores valores de (r), e menores erros como EMA (1,36), REMQ (1,79) e EMP (1,90%) (Figura 51).

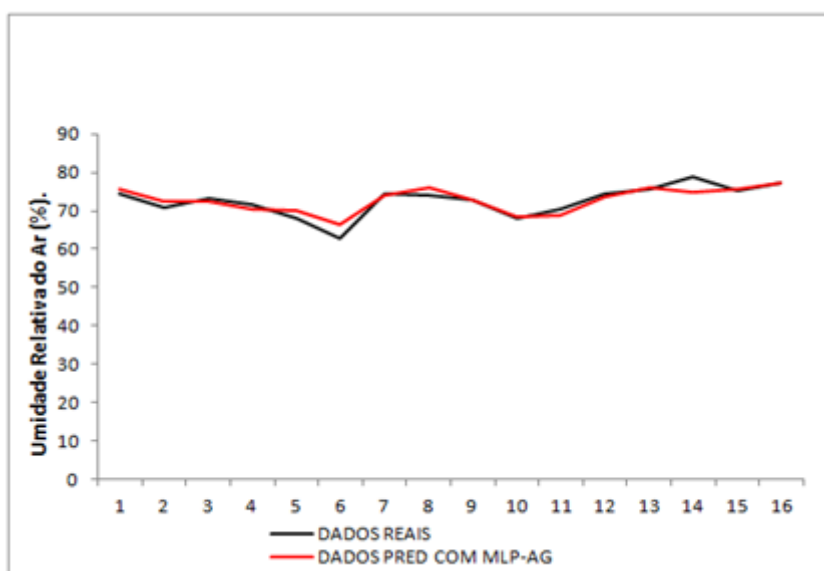


Figura 51: Resultados de estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação do Rio de Janeiro.

A partir do modelo MLP-AG foi realizado o preenchimento de 72% das falhas da variável de umidade relativa do ar ocorridas na estação do Rio de Janeiro, ou seja, 5 das 7 falhas. Esse modelo demonstrou uma vantagem em comparação com outros modelos, pois o conjunto de estimadores definidos de forma manual utilizados pelos modelos MED, RLM, MLP e RBF conseguiriam preencher apenas 4 das 7 falhas ocorridas, devido os mesmos possuírem falhas em 31/01/2007 na estação de Resende, 31/01/2009 em Paty do Alferes e Resende e em 31/05/2012 na estação de Cordeiro. Mais informações nos ANEXOS B, D e E.

Os dados preenchidos pelo modelo MLP-AG podem ser visualizados na Figura 52 pela cor vermelha.

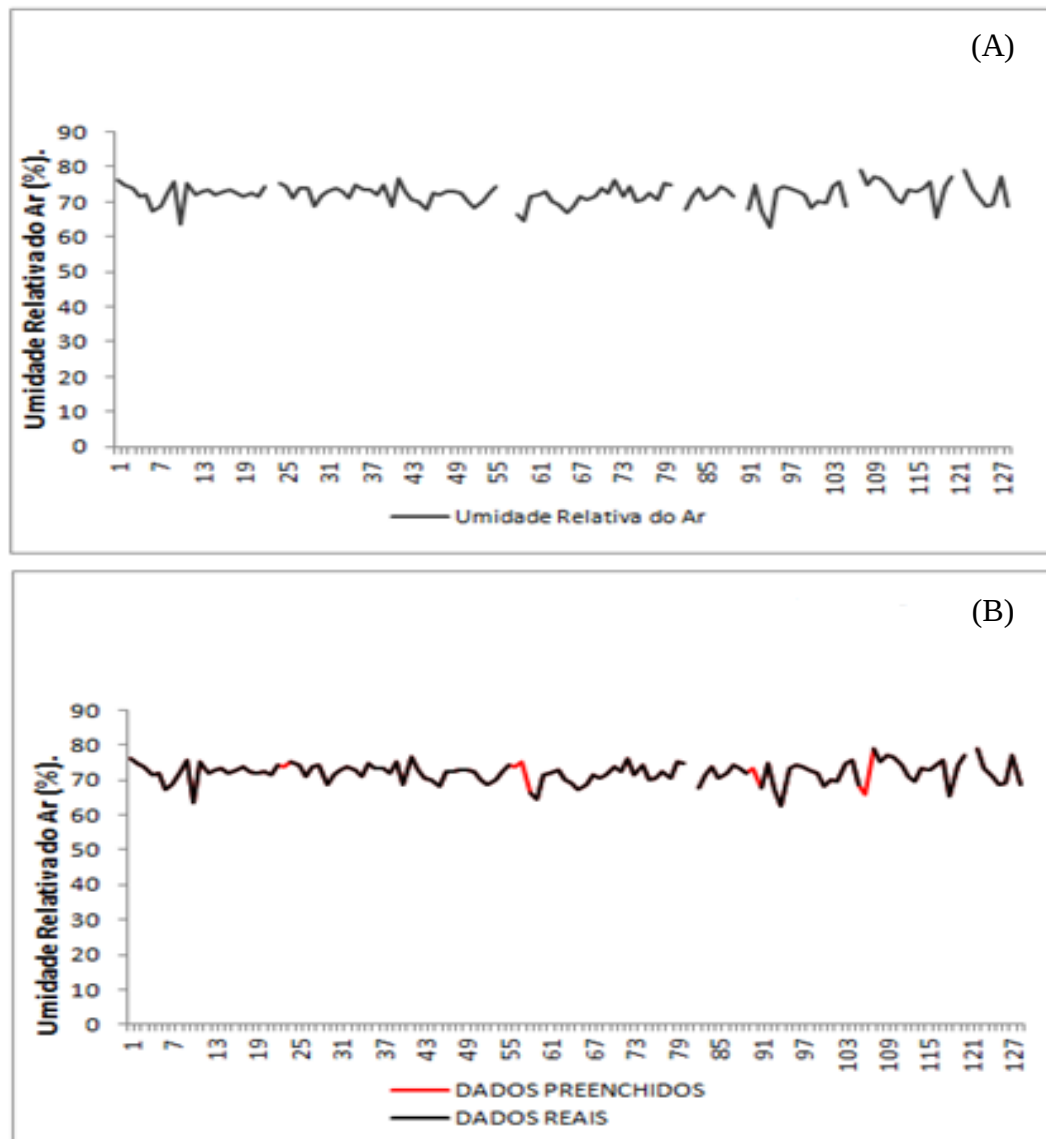


Figura 52: Estação do Rio de Janeiro, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.

6.4.6. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação de Paty do Alferes

O modelo MLP-AG aplicado na estimação dos dados de umidade relativa do ar da estação do Rio de Janeiro utilizou em sua primeira camada intermediária uma função *purelin*, na segunda camada uma função *satlin*, e o algoritmo de treinamento *traingd*. As configurações podem ser visualizadas no APÊNDICE C.

Os resultados obtidos com a aplicação dos modelos para estimativa dos 16 meses de umidade relativa do ar podem ser visualizados na Tabela 33.

Tabela 33: Resultados obtidos com os modelos aplicados na estimativa de umidade relativa do ar da estação de Paty do Alferes.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(%)	EMP(%)
MED	0,91	1,92	0,94	0,86	Ótimo	1,60	2,05
RLM	0,93	2,67	0,91	0,84	Muito Bom	2,20	2,80
MLP	0,91	2,12	0,93	0,84	Muito Bom	1,67	2,12
RBF	0,91	1,89	0,94	0,86	Ótimo	1,54	1,95
MLP-AG	0,94	1,17	0,97	0,92	Ótimo	0,90	1,15

Analisando os resultados obtidos foi possível verificar que os modelos apresentaram bons desempenhos na estimativa dos dados de umidade relativa do ar pertencente a estação de Paty do Alferes. No entanto, os modelos MED, RBF e MLP-AG se destacaram por apresentar resultados mais apurados que os modelos RLM e MLP. Contudo, o modelo MLP-AG mostrou maior aptidão em suas estimações onde o mesmo conseguiu um EMA igual a 0,90, REMQ igual a 1,17 e um EMP igual a 1,15%. Isso garante que o modelo MLP-AG apresentou 98,85% de acerto em suas predições (Figura 53).

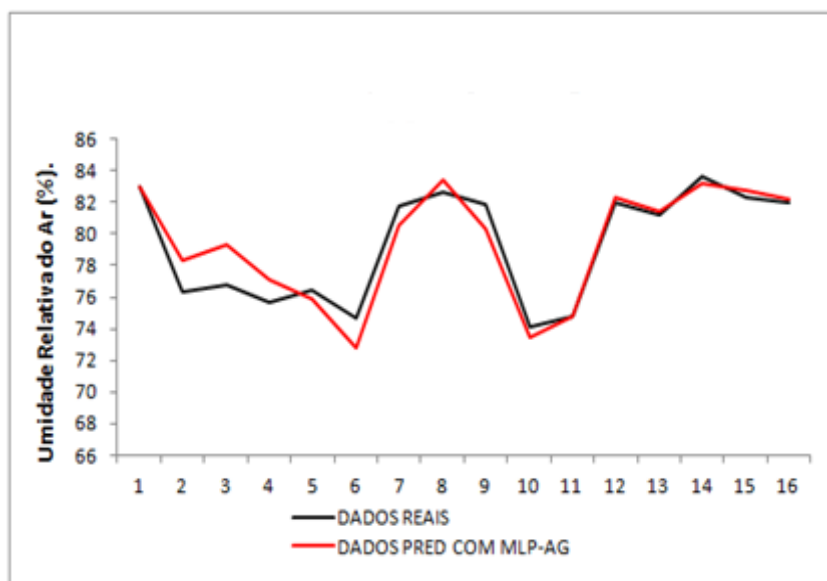


Figura 53: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Paty do Alferes.

O modelo MLP-AG permitiu o preenchimento de 54,55% das falhas, ou seja, 6 das 11 falhas ocorridas na série de umidade relativa do ar. Novamente, o modelo MLP-AG demonstrou uma vantagem comparado aos demais modelos empregados neste estudo, pois o conjunto de estimadores definidos de forma manual utilizados pelos modelos MED, RLM, MLP e RBF conseguiriam preencher apenas 5 das 11 falhas existentes, devido os mesmos possuírem falhas em 31/12/2008, 31/01/2009, 31/02/2009, 31/01/2012 na estação de Resende, e em 29/02/2012, 31/07/2012 na estação de Cordeiro (ANEXOS D, B e E). A Figura 54 demonstra o preenchimento realizado pelo modelo MLP-AG.

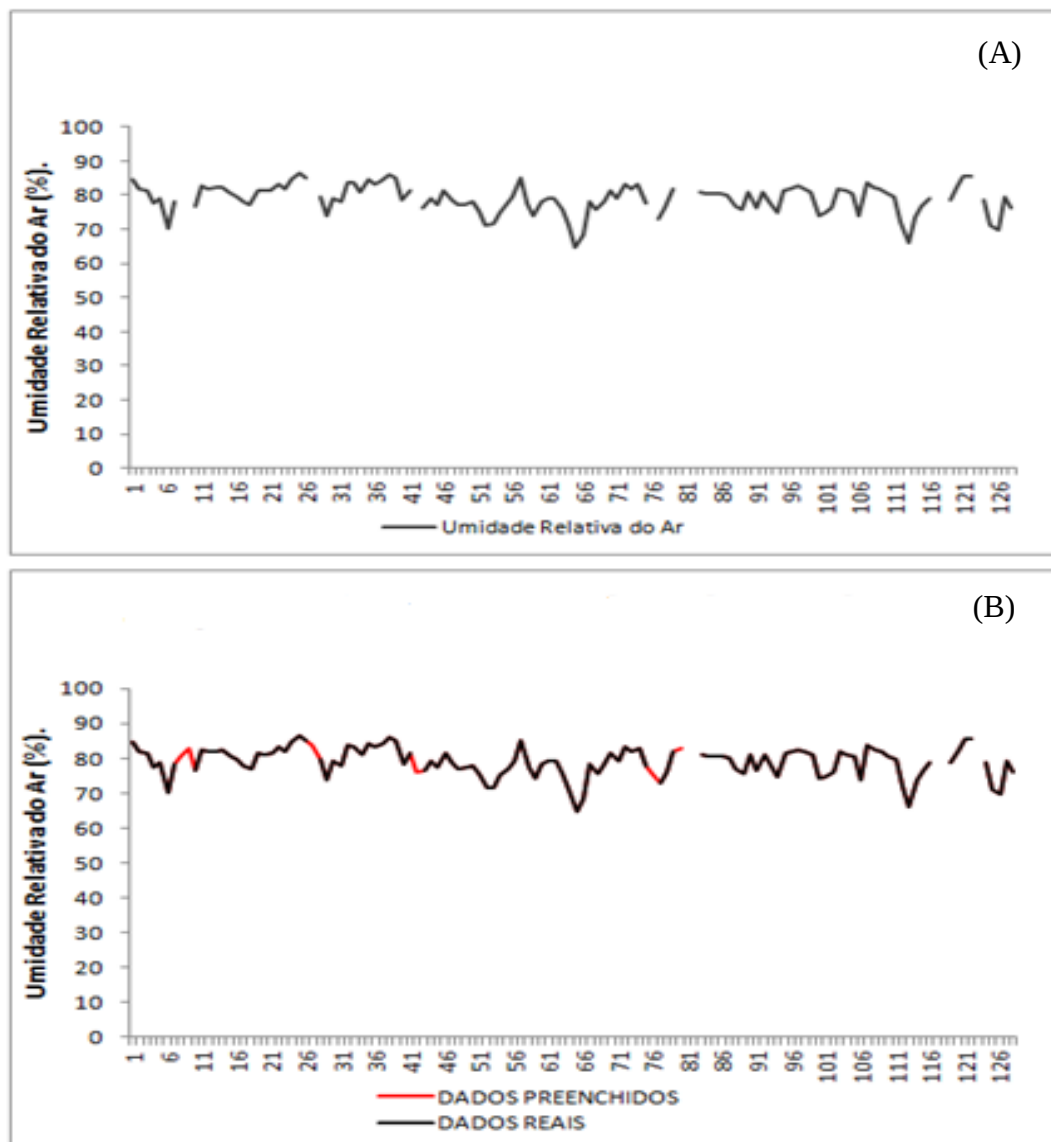


Figura 54: Estação Paty do Alferes, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.

6.4.7. Resultados e Preenchimento das Falhas de Umidade Relativa do Ar para estação de Resende

As configurações adotadas pelo modelo MLP-AG são similares com as utilizadas pelo modelo MLP, como o emprego da função *tansig* na primeira camada intermediária, a função *logsig* na segunda camada intermediária e algoritmo *trainbfg* para treinamento. As configurações podem ser visualizadas no APÊNDICE C.

Os resultados obtidos com a aplicação dos modelos para estimativa dos 16 meses de umidade relativa do ar da estação de Resende podem ser visualizados na Tabela 34.

Tabela 34: Resultados obtidos com os modelos aplicados na estimação de umidade relativa do ar da estação de Resende.

Modelos	(r)	REMQ	D	C	Desempenho	EMA(%)	EMP(%)
MED	0,81	3,60	0,85	0,69	Bom	2,53	3,27
RLM	0,85	3,33	0,88	0,75	Bom	2,50	3,24
MLP	0,90	2,34	0,92	0,83	Muito Bom	1,82	2,30
RBF	0,89	2,31	0,93	0,83	Muito Bom	1,79	2,30
MLP-AG	0,90	2,09	0,93	0,84	Muito Bom	1,49	1,87

A avaliação dos resultados permitiu verificar o ganho de cada modelo no aumento dos termos (r), (D), (C), e na minimização dos erros EMA, REMQ, e EMP. Por meio dessa avaliação foi identificado que os modelos de RNAs proporcionaram os melhores resultados.

Contudo, o modelo MLP-AG apresentou estimações mais próximas da realidade, obtendo altos índices de (r), (D), (C), e erros como EMA igual 1,49, REMQ igual a 2,09 e EMP igual a 1,87%, indicando que este modelo possuiu 98,13% de acerto em suas estimações.

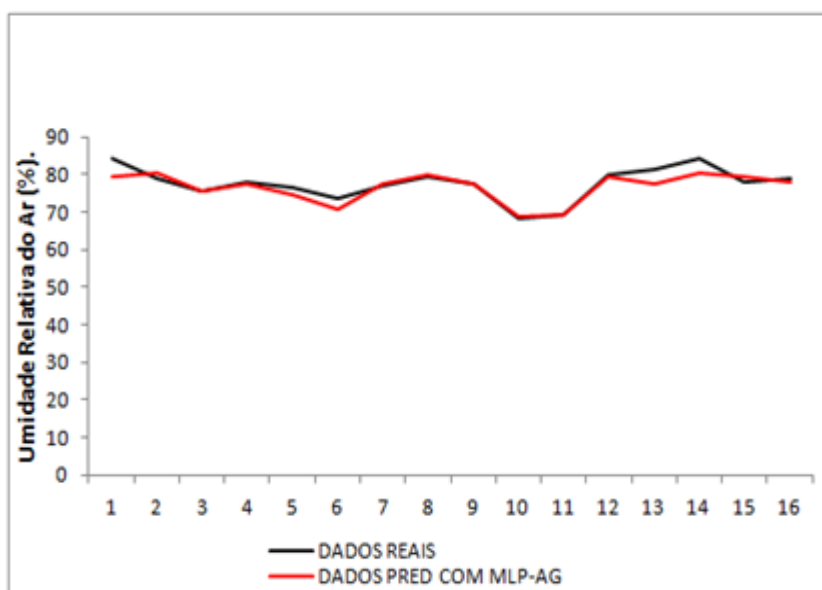


Figura 55: Resultados da estimativa de umidade relativa do ar obtidos com o modelo MLP-AG para a estação de Resende.

A partir do modelo MLP-AG foi possível preencher cerca 63% das falhas de umidade relativa do ar ocorridas na estação de Resende, ou seja, 17 das 27 falhas. No entanto, com a aplicação dos outros modelos, também seria possível preencher o mesmo número de falhas (ANEXOS A, B, D e E). O preenchimento dos dados a partir do modelo MLP-AG pode ser visto na Figura 56.

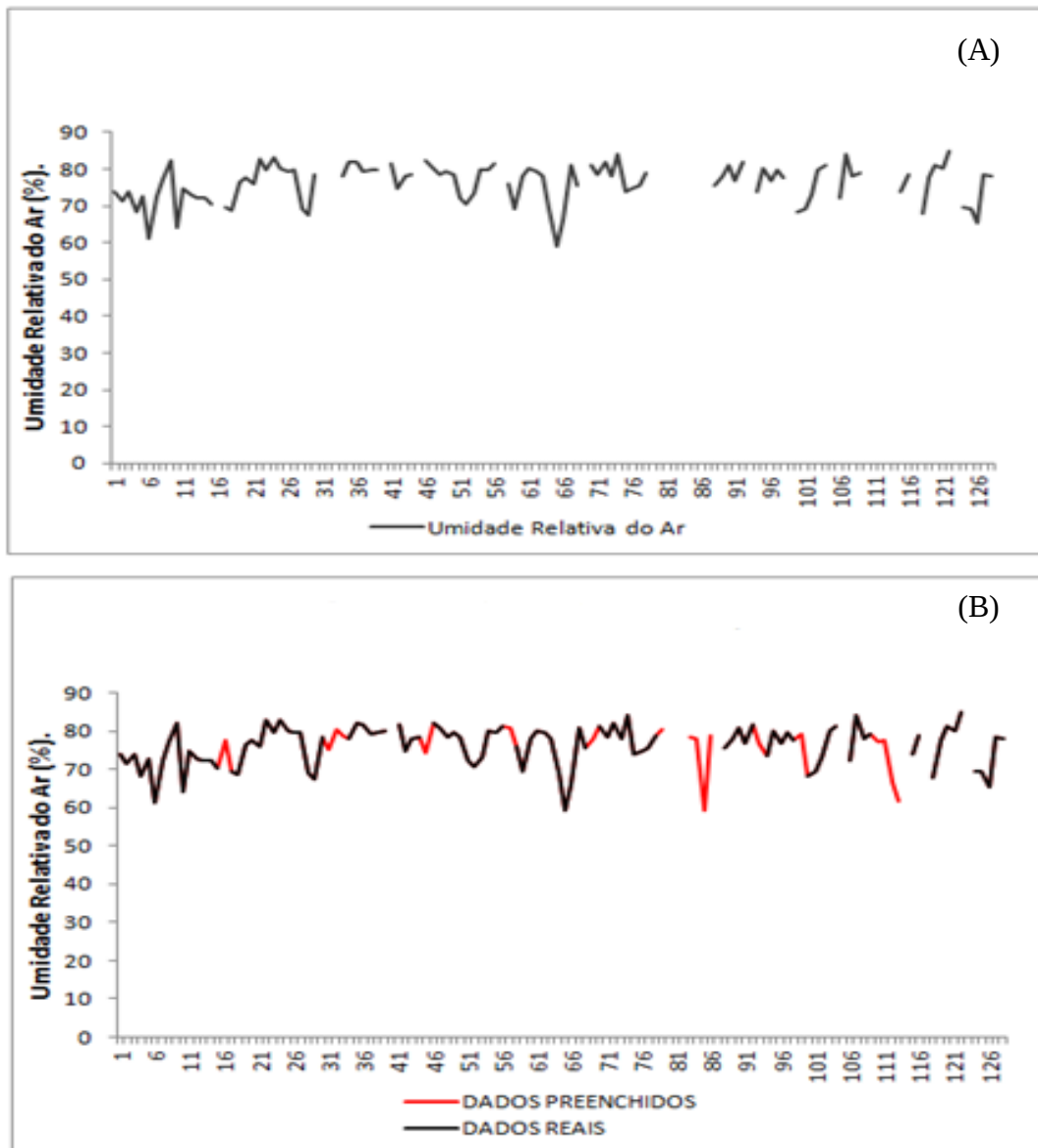


Figura 56: Estação de Resende, (A) Dados de umidade relativa do ar com falhas, (B) Dados de umidade relativa do ar preenchida.

7. DISCUSSÃO

Determinar o conjunto de estações estimadoras de forma manual, utilizando em parte a proximidade entre as localidades, se mostrou inviável na obtenção das variáveis com maior índice de correlação. Os índices encontrados por esta técnica divergiram bastante dos encontrados pelo modelo MLP-AG, que, apesar de não ter consigo obter em todas as combinações os melhores estimadores, o mesmo apresentou mais vezes as combinações de estimadores com maiores índices de correlação entre as variáveis a serem preenchidas. Isto pode ser verificado nas Tabelas 7, 8, 17, 18, 27 e 28.

Uma possível explicação para esse fato é a existência da diversidade climática entre as regiões, onde não necessariamente uma região próxima de outra pode possuir características climáticas similares. No entanto, apesar do modelo MLP-AG ter demonstrado uma vantagem na definição das variáveis, ele não pode ser considerado como a melhor escolha para o problema, pois a realização da correlação identificaria de forma mais rápida as estações com maior relação. Porém, o método de escolha definido pelo AG, se justificaria com o aumento do número de estações e variáveis, onde o mesmo poderia poupar tempo definindo combinações aceitáveis para geração de bons resultados.

Mesmo assim, o uso das informações meteorológicas definidas por estes métodos para estimativa de dados faltantes de uma determinada estação apresentou bons resultados. Com isso, foi possível observar que em quase todos os casos os modelos aplicados mostraram desempenhos satisfatórios. Comparando de modo geral as estimativas de evaporação, temperatura máxima do ar e umidade relativa do ar geradas por todos os modelos, com os dados observados, foi possível constatar que o modelo MLP-AG apresentou na maioria dos casos os melhores resultados.

Obtendo para evaporação índices de (r) entre os dados estimados e esperados igual a 0,82 para a pior estimativa apresentada e 0,97 para a melhor, (D) igual a 0,87 para a pior estimativa e 0,98 para o melhor caso, (C) igual a 0,72 para o pior caso e 0,96 para o melhor caso, o que manteve o desempenho do método nas estimativas de evaporação entre Bom, Muito Bom e Ótimo. Resultados similares foram encontrados por Leitão, M., Oliveira e Leitão, T., (2007), na estimativa de evaporação da região da Paraíba na cidade de Boqueirão, por meio da aplicação dos métodos de Penman, Linacre, Snyder, e outros, onde foram obtidos índices de (C) iguais a 0,71 para o pior caso e 0,91 para o melhor caso, mantendo o desempenho entre Bom, Muito Bom e Ótimo.

Para temperatura máxima do ar, o modelo MLP-AG apresentou erros de EMA que variaram entre 0,47°C para a pior estimativa e 0,27°C para a melhor estimativa, EMP igual a 1,57% para pior estimativa e 0,95% para melhor estimativa. Resultados inferiores, porém próximos, foram encontrados por Fernandez (2007), Chou et al (2007) e Ventura (2012), onde seus erros variaram entre 2°C a 0,57°C na estimativa de dados de temperatura para as regiões do Rio Grande do Sul, Nordeste e Mato Grosso, respectivamente.

Já para umidade relativa do ar, o modelo MLP-AG teve erros de EMA que oscilaram entre 1,49 para a pior estimativa e 0,90 para a melhor estimativa, EMP igual a 1,90% para a pior estimativa e 1,15% para a melhor estimativa. Resultados apresentados por Viecheneski (2012) demonstraram um EMP de 11,02% de acerto na pior estimativa e de 4,30% para a melhor estimativa.

Dentre as várias explicações para superioridade dos resultados apresentados pelo modelo MLP-AG, uma delas foi a escolha de estações que apresentam altos índices de (r), causando uma alta relação entre os dados utilizados para estimação e os dados a serem estimados. Outro fato é o teste de várias metodologias e parâmetros para treinamento das RNAs, onde é possível observar nos resultados que para cada dado estimado, o AG definiu

diferentes configurações para RNAs. Consequentemente, uma explicação empírica pode ser tomada através de uma das características do AG, que seria aumentar o espaço de busca de um ponto ótimo global. Desta forma, pode-se inferir que o AG aumentou a capacidade das RNAs aplicadas na estimativa das variáveis, chegando bem próximas dos dados reais.

No entanto, apesar do modelo MLP-AG apresentar um alto desempenho na estimativa dos dados meteorológicos, alguns problemas foram encontrados na realização do preenchimento das falhas reais. Isso ocorreu devido a existência de falhas no conjunto de variáveis estimadoras no mesmo período a ser preenchido por uma estação, fazendo com que não fosse possível realizar o preenchimento completo das falhas. Contudo, apesar do modelo MLP-AG não ter conseguido preencher completamente todas as falhas das variáveis estudadas, o mesmo apresentou resultados satisfatórios, preenchendo 100% dos dados referentes a evaporação, 89,2% da temperatura máxima do ar e 75,3% para umidade relativa do ar.

Outro resultado relevante foi que o modelo MLP-AG conseguiu gerar dados aceitáveis de forma rápida levando em média de 30 minutos a 4 horas para cada geração de resultados. Isso pode ter ocorrido por causa do pré-processamento realizado nas variáveis que gerou uma série padronizada e consistente.

8. CONCLUSÕES

O objetivo principal do presente estudo foi aplicar e analisar uma metodologia híbrida composta por RNAs e AG no preenchimento de falhas de dados meteorológicos oriundos de seis estações pertencentes ao estado do Rio de Janeiro. Para tanto, o modelo aplicado utilizou para estimativa de cada dado um conjunto de dados pertencente as demais estações.

De acordo com as análises dos resultados obtidos foi possível verificar que o modelo proposto MLP-AG obteve na maioria dos resultados baixos erros e altos índices de correlação entre os dados estimados e os esperados. Comparando este modelo aos outros modelos aplicados, o mesmo conseguiu na maioria das ocasiões um desempenho superior na estimativa dos dados de evaporação, temperatura máxima do ar e umidade relativa do ar com maior qualidade.

Outro fato importante foi que a utilização de AGs para auxiliar a escolha de treinamento e parâmetros das RNAs, pode maximizar a capacidade da RNAs, fazendo com que o modelo proposto apresenta-se resultados satisfatórios.

No entanto, a ocorrência de falhas nos dados utilizados para estimação prejudicou parte do preenchimento das séries de temperatura máxima do ar e umidade relativa do ar, que não foram realizados em 100%.

Apesar do modelo proposto não ter conseguido preencher completamente todas as falhas, o mesmo teve resultados bem satisfatórios, sendo ainda superior ao que os outros modelos aplicados conseguiriam preencher. Por isso, pode-se concluir que o modelo MLP-AG é uma alternativa viável no preenchimento de falhas dos dados meteorológicos utilizados e posteriormente poderá ser aperfeiçoado e aplicado em outras áreas científicas.

9. TRABALHOS FUTUROS

Este estudo apresentou um modelo híbrido composto RNA e AG para o preenchimento de dados de evaporação, temperatura máxima do ar e umidade relativa do ar. Tal metodologia foi comparada com outras técnicas tradicionais e RNAs. Como trabalhos futuros pretende-se:

Expandir a aplicação deste modelo na estimativa de diferentes parâmetros meteorológicos, como precipitação, radiação solar, velocidade do vento e outros, a fim de avaliar com maior precisão a capacidade do modelo, neste campo de aplicação.

Aplicar o AG junto a um fator de ponderação baseado na qualidade da série histórica mais estável para determinar conjuntos de dados com menor número de falhas, e assim evitar ou diminuir o problema de não preenchimento de uma falha.

Outra pretensão é a automatização da organização das informações e do processo de preenchimento de falhas através da criação de banco de dados junto a um software, que possa realizar o pré-processamento das variáveis de forma rápida e as estimativas dos dados com falhas com qualidade.

10. REFERÊNCIAS BIBLIOGRÁFICAS

- AGGARWAL, R.; SONG, Y. **Artificial Neural Networks in Power Systems**. Power Engineering Journal. v.12, p. 279-287, 1998.
- ALBARAKATI, N.; KECMAN, V. **Fast Neural Network Algorithm For Solving Classification Tasks**. IEEE-Southeastcon Proceedings. p.1 - 8, 2013.
- ALVES, M. R. **Construção de Base de Regras Fuzzy para Identificação de Malignidades Humanas a partir de Experimentos de Microarranjos de DNA**. 214 p. Tese (Programa de Pós-graduação em Engenharia - Doutor em Engenharia Biomédica). Universidade Federal do Rio de Janeiro, 2006. Rio de Janeiro- RJ.
- ALVES, A. C.; STEINER, M. T. A.; MUSSI, N. H.; ZANELATTO, G. **Algoritmos Genéticos Aplicados ao Planejamento da Distribuição de Energia Elétrica em Curitiba e Região Metropolitana**. In Computação Evolucionária em Problemas de Engenharia. 2011. p.17.
- ASADUZZAMAN, MD.; AHMED, S. U.; KHAN, F. E.; SHAHJAHAN, MD.; MURASE, K. **Making Use of Damped Noisy Gradient in Training Neural Network**. IEEE - Neural Networks (IJCNN), The 2010 International Joint Conference on. p. 1- 5, 2010.
- BARCELLOS, J. C. H. **Algoritmos Genéticos Adaptativos: Um estudo comparativo**. 143 p. Dissertação (Mestre em Engenharia). Escola Politécnica da Universidade de São Paulo, 2000. São Paulo - SP.
- BECKAMANN, M. **Algoritmos Genéticos como Estratégia de Pré-processamento em Conjuntos de Dados Desbalanceados**. 112 p. Dissertação (Programa de Pós-graduação em Engenharia Civil - Mestre em Engenharia Civil) - Universidade Federal do Rio de Janeiro – UFRJ, 2010. Rio de Janeiro-RJ.
- BISHOP, C. M. 1997. **Neural Networks for Pattern Recognition**. Oxford University Press.
- BOUGHRARA, H.; CHTOUROU, M.; AMAR, C. B. **MLP Neural Network Based Face Recognition System Using Constructive Training Algorithm**. IEEE - Multimedia Computing and Systems (ICMCS), International Conference on. p.233-238, 2012.
- BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. **Redes Neurais Artificiais - Teoria e Aplicações**. Ed. 2, p. 225, 2012.
- CASTRO, M. C. F. **Predição Não - Linear de Séries Temporais Usando Redes Neurais RBF por Decomposição em Componentes Principais**. 192 p. Tese (Doutor em Engenharia Elétrica). Universidade Estadual de Campinas - UNICAMP, 2001. Campinas -SP.
- CAUDILL, M. **Neural Network Primer**, Miller Freeman Publications, 1990.
- CAVALCANTE JUNIOR, J. A. C.; **Ferramenta Computacional Inteligente Aplicada à Otimização do Uso de Recursos Hídricos e Aumento da Receita Líquida na Agricultura - INTELIAGRI**. 105 p. Tese (Programa de Pós-graduação em Ciência, Tecnologia e

Inovação em Agropecuária- Doutor). Universidade Federal Rural do Rio de Janeiro - UFRRJ, 2013. Seropédica - RJ.

CEPEA. **PIB Agro CEPEA-USP/CNA - PIB do Agronegócio - Dados de 1994 a 2011.** Disponível em: <http://www.cepea.esalq.usp.br/pib/>. Acesso em agosto de 2013.

CERQUEIRA, E. O.; ANDRADE, J. C.; POPPI, R. J. **Redes Neurais e suas Aplicações em Calibração Multivariada.** Quim Nova. vol. 24, n. 6. p.864-873, 2001.

CHIBANA, E. Y.; FLUMIGNAN, D.; MOTA, R. G.; VIEIRA, A. S.; FARIA, R. T. **Estimativa de Falhas em Dados Meteorológicos.** V Congresso Brasileiro de Agroinformática- SBI-AGRO. 2005. 8 p. Londrina-PR.

CHOU, S. C.; SOUZA, C. R.; GOMES, J. L.; EVANGELISTA, E. F. D.; OSÓRIO, C.; CATALDI, M. **Refinamento Estatístico das Previsões Horárias de Temperatura a 2M do Modelo ETA em Estações do Nordeste do Brasil.** Revista Brasileira de Meteorologia, v. 22, n.3. p. 287-296, 2007.

COHOON, J. P.; PARIS, W. D. **Genetic Placement.** IEEE - Transactions on Computer AIDED Design, vol. 6, n. 6. p. 956-964, 1987.

COUTINHO, E. R.; VALE, I. G. **Sistema de Gerenciamento Meteorológico de Paracambi.** Instituto Superior de Tecnologia de Paracambi, 2011. 81 p. (Monografia de Término de Graduação).

COVER, T. M. **Geometrical and Statistical Properties of Systems of Linear Inequalities With Applications in Pattern Recognition.** IEEE - Electronic Computers, Transactions on, v. 14, n. 3. p. 326-334, 1965.

DESHMUKH, R. P.; GHATOL, A. A. **Comparative Study of Temporal Neural Networks for Short Term Flood Forecasting.** International Journal of Computer Applications, v. 5, n. 12. p. 24-28, 2010.

EBEID, H. M. **Using MLP and RBF Neural Networks for Face Recognition: an Insightful Comparative Case Study.** IEEE- Computer Engineering & Systems (ICCES), International Conference on. p. 123-128, 2011.

FERNANDEZ, M. N. **Preenchimento de Falhas em Séries Temporais.** 106 p. Dissertação (Programa de Pós-graduação em Engenharia Oceânica - Mestre em Engenharia Oceânica) Universidade Federal do Rio Grande, 2007. Rio Grande.

FERRAZ, R. C. **Desenvolvimento de um Sistema Web para Estimativa Numérica de Dados Meteorológicos do Rio Grande do Sul.** 87 p. Dissertação (Mestre em Geomática). Universidade Federal de Santa Maria, Santa Maria – RS. 2010.

FONSECA, J. S.; MARTINS, G. A.; TOLEDO, G. L. **Estatística Aplicada.** Ed. 2, p. 267, 2012.

GOLDBERG, D. E. **Genetic Algorithms in Search, Optimization and Machine Learning.** New York: Addison-wesley. Ed.1, p. 432, 1989.

GONÇALVES, R. M.; COELHO, L. S.; KRUEGER, C. P.; HECK, B. **Modelagem Preditiva de Linha de Costa Utilizando Redes Neurais Artificiais**. Bol. Ciênc. Geod., sec. Artigos, Curitiba, v. 16, n. 3, p.420-444, 2010.

GUARNIERI, R. A. **Emprego de Redes Neurais Artificiais e Regressão Linear Múltipla no Refinamento das Previsões de Radiação Solar do Modelo ETA**. 171 p. Dissertação (Programa de Pós-graduação em Meteorologia). Instituto Nacional de Pesquisas Espaciais-INPE, 2006. São José dos Campos-SP.

HAGAN, M. T.; MENHAJ, M. B. **Training Feedforward Networks with the Marquardt Algorithm**. IEEE Transactions on Neural Networks, v. 5, p. 989-993, 1994.

HAYKIN, S. **Redes Neurais Princípios e Prática**. Ed. 2, p. 902, 2001.

HEATON, J. **Programming Neural Networks with Encog 2 in Java**. p. 481, 2010.

HELDWEIN, A. B.; CARLESSO, R. ; PETRY, M. T.; ROSA, G. M. **Usos e Benefícios da Coleta Automática de Dados Meteorológicos na Agricultura**. Ed.1, p. 170,2007.

HU, Y. H.; HWANG, J. **Handbook of Neural Network Signal Processing**. CRC Press, New York. p. 384, 2002.

HUTH, R.; NEMESOVÁ, I.; **Estimation of Missing Daily Temperatures: Can a Weather Categorization Improve its Accuracy?**. Journal of Climate, v.8, p. 1901-1916, 1995.

INMET. **Estação Meteorológica de Observação de Superfície Convencional**. Disponível em: <http://www.inmet.gov.br/html/observacoes.php>. Acesso em novembro de 2013.

LEENAERS, H.; OHX, J. P.; BURROUGH, P. A. **Comparison of Spatial Prediction Methods for Mapping Floodplain Soil Pollution**. Catena, v.17, p.535-550, 1990.

LEITÃO, M. M. V. B. R.; OLIVEIRA, G. M.; LEITÃO, T. J. V. **Avaliação do Desempenho de Diferentes Métodos de Estimativa da Evaporação para duas Regiões da Paraíba**. Revista Brasileira de Engenharia Agrícola e Ambiental - Agriambi, v. 11, n. 6. p.585-593, 2007.

LINDEN, R. **Algoritmos Genéticos**. Ed. 3, p. 474, 2012.

LIPPMANN, R. P. **An Introduction to Computing with Neural Nets**. IEEE ASSP MAGAZINE. p.4- 22, 1987.

LONGOBARDI, A; VILLANI, P. **Trend Analysis of Annual and Seasonal Rainfall time Series in the Mediterranean Area**. International Journal of Climatology. p. 1538-1546, 2009.

LYRA, G. B.; SOUZA, M. O.; VIOLA, D. N. **Modelos Lineares Aplicados à Estimativa da Concentração do Material Particulado (PM₁₀) na Cidade do RIO DE JANEIRO, RJ**. Revista Brasileira de Meteorologia-RBMET, v. 26, n.3, p. 392 - 400, 2011.

MACHADO, F. W. **Modelagem Chuva-Vazão Mensal Utilizando Redes Neurais Artificiais**. 174 p. Dissertação (Programa de Pós-graduação em Engenharia de Recursos Hídricos e Ambiental - Mestre em Engenharia Hidrológica). Universidade Federal do Paraná, 2005. Curitiba-PR.

MATHWORKS. **Radial Basis Neural Networks**. Disponível em: <http://www.mathworks.com/help/nnet/ug/radial-basis-neural-networks.html>. Acesso em fevereiro de 2014.

MOREIRA, G. J. P.; WANNER, E. F., GUIMARÃES, F. G.; DUCZMAL, L. H.; TAKAHASHI, R. H. C. **LMI Formulation for Multi objective Learning in Radial Basis Function Neural Networks**. IEEE - Neural Networks (IJCNN), International Joint Conference on. p.1-7, 2010.

MULLUHI, Q. M.; BAYOUMI, M. A.; RAO, T.R.N. **An Efficient Mapping of Multilayer Perceptron with Backpropagation ANNs on Hypercubes**. IEEE - Parallel and Distributed Processing, Proceedings of the Fifth Symposium on. p.368-375, 1993.

NETO, P.; PELLI, A. **Avaliação de Imóveis Urbanos com Utilização de Sistemas Nebulosos (Redes Neuro-Fuzzy) e Redes Neurais Artificiais**. XXI Congresso Panamericano de Valuación Cartagena – Colombia. 2004. p.19.

PERES, F.; MOREIRA, J. C. **Saúde e Ambiente em sua Relação com o Consumo de Agrotóxicos em um Pólo Agrícola do Estado do Rio de Janeiro, Brasil**. Cad. Saúde Pública, Rio de Janeiro. p. 612-621, 2007.

PEZZOPANE, J. E. M.; CASTRO, F. S.; PEZZOPANE, J. R. M.; CECÍLIO, R. A. **Agrometeorologia Aplicações para o Espírito Santo**. p.163, 2012.

RIEDMILLER, M.; BRAUN, H. **A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm**. IEEE - Neural Networks, International Conference, v. 1, p. 586-591, 1993.

ROBINSON, T.P.; METTERNICHT, G. **Testing the Performance of Spatial Interpolation Techniques for Mapping Soil Properties**. Computers and Electronics in Agriculture, v. 50, p. 97-108, 2006.

ROCHA, M.; CORTEZ, P.; NEVES, J. M. **Análise Inteligente de Dados Algoritmo e Implementação em Java**. p. 177, 2008.

RODRIGUES, F. L.; LEITE, H. G.; SANTOS, H. N.; SOUZA, A. L.; SILVA, G. F. **Metaheurística Algoritmo Genético para Solução de Problemas de Planejamento Florestal com Restrições de Integridade**. Sociedade de Investigações Florestais - SIF, V. 28, n.2, p. 233-245, 2004.

RODRIGUES, G. C. **Utilização de Redes Neurais para Previsão de Ventos no Horizonte de 24 Horas**. 119 p. Dissertação (Programa de Pós-graduação em Engenharia - Mestre em Ciências em Engenharia Elétrica). Universidade Federal do Rio de Janeiro- UFRJ, 2007. Rio de Janeiro-RJ.

ROLIM, J.; CATALÃO, J.; TEIXEIRA, J. **Desenvolvimento de um Interpolador Espacial de Dados Agrometeorológicos. Aplicação à Região do Alentejo.** III Congresso Nacional de Rega e Drenagem. 2010. 10 p. Beja – Portugal.

SANTOS R. S. **Homogeneidade e Reconstrução de Séries Climatológicas para Localidades no Estado de Minas Gerais.** 100 p. Tese (Programa de Pós-graduação em Meteorologia Agrícola- Doctor Scientiae). Universidade Federal de Viçosa – UFV, 2012. Viçosa-MG.

SBMET. **Boletim de Sociedade Brasileira de Meteorologia – Dados Ambientais: qual Deve ser a Política de Dados?.** v. 32, n. 1, p. 120, 2008.

SERRANO, S. M. V.; BEGUERÍA, S.; MORENO, J. I. L.; VERA, M. A. G.; STEPANEK, P. **A Complete Daily Precipitation Database for Northeast Spain: Reconstruction, Quality Control, and Homogeneity.** International Journal of Climatology. v. 30, p. 1146-1163, 2010.

SHAH, H.; GHAZALI, R. **Prediction of Earthquake Magnitude by an Improved ABC-MLP.** IEEE- Developments in E-Systems Engineering. p. 312 - 317, 2011.

SKODZIK, J.; ALTMANN, V.; WAGNER, B.; DANIELIS, P.; TIMMERMAN, D. **A Highly Integrable FPGA-Based Runtime-Configurable Multilayer Perceptron.** IEEE- 27th International Conference on Advanced Information Networking and Applications. p. 429-436, 2013.

SOMASUNDARAM, R. S.; NEDUNCHEZHIAN, R. **Radial Basis Function Network Dependent Exclusive Mutual Interpolation for Missing Value Imputation.** Journal of Computer Science.v. 9, p. 327-335, 2013.

SONG, Q.; KASABOV, N. **WDN-RBF: Weighted Data Normalization for Radial Basic Function Type Neural Networks.** IEEE - Neural Networks, Proceedings. International Joint Conference on. p. 349 - 356, 2004. .

SOUSA, N. M. N.; DANTAS, R. T.; LIMEIRA, R. C. **Influência de Variáveis Meteorológicas Sobre a Incidência do Dengue, Meningite e Pneumônia em JOÃO PESSOA-PB.** Revista Brasileira de Meteorologia-RBMET, v. 22, n.2, p. 183-192, 2007.

SRINIVAS, M.; PATNAIK, L.M. **Genetic Algorithms: A Survey.** IEEE- Computer. v. 27, p.17-26, 1994.

TASPINAR, N.; SEYMAN, M. N. **Back Propagation Neural Network Approach for Channel Estimation in OFDM System.** IEEE - Wireless Communications, Networking and Information Security (WCNIS), International Conference on. p. 265-268, 2010.

TEIXEIRA, L.; TAVARES, P.; ARAÚJO, M. **Princípios de Meteorologia e Meio Ambiente.** Disponível em: <http://www.cptec.inpe.br/glossario.shtml>. Acesso em janeiro de 2013.

THIRUGNANAM, K.; JOY, T. P. E R.; SINGH, M.; KUMAR, P. **Mathematical Modeling of Li-Ion Battery Using Genetic Algorithm Approach for V2G Applications.** IEEE - Transactions on Ennergy Conversion. p.1 - 12, 2014.

TORRES JUNIOR, R. G.; MACHADO, M. A. S.; SOUZA, R. C. **Previsão de Séries Temporais de Falhas em Manutenção Industrial Usando Redes Neurais**. ENGEVISTA, v. 7, n.2, p. 4-18, 2005.

VENTURA T. M. **Preenchimento de Falhas de Dados Micrometeorológicos Utilizando Técnicas de Inteligência Artificial**. 84 p. Dissertação (Programa de Pós-graduação em Física Ambiental- Mestre em Física)- Universidade Federal do Mato Grosso -UFMT, 2012. Cuiabá – MT.

VIECHENESKI, R. **Aplicação de Redes Neurais Artificiais no Tratamento de Dados Agrometeorológicos Visando a Correção de Séries Temporais**. 86 p. Dissertação (Programa de Pós-graduação em Computação Aplicada - Mestre em Computação Aplicada). Universidade Estadual de Ponta Grossa-UEPG, 2012. Ponta Grossa-PR.

WANKHADE, K.; RANE, D.; THOOL, R. **A New Feature Selection Algorithm for Stream Data Classification**. IEEE - Advances in Computing, Communications and Informatics (ICACCI), International Conference on. p.1843-1848, 2013.

WMO-World Meteorological Organization. **Manual on the Global Observing System**. Volume I, WMO-No. p. 544, Geneva, 2003.

XIAN, Z.; WU, H.; SIQING, S.; SHAOQUAN, Z. **Application of Genetic Algorithm-Neural Network for the Correction of Bad Data in Power System**. IEEE- Electronics, Communications and Control (ICECC), International Conference on. p.1894 -1897, 2011.

XUE, S.; YANG, M.; LI, C.; NIE, J. **Meteorological Prediction Using Support Vector Regression with Genetic Algorithms**. IEEE - The 1st International Conference on Information Science and Engineering(ICISE2009). p. 4931-4935, 2009.

YI, Q.; YORK, B. W. **Some Results for Chaotic Times Series Prediction Using Clifford Neural Networks**. IEEE- Signal Processing, 2004. Proceedings ICSP'04 7th International Conference on. p. 61 - 64, 2004.

ANEXOS

ANEXO A- Dados Meteorológicos da Estação de Campos dos Goytacazes

Série de dados utilizados da estação de Campos dos Goytacazes			
Data	Evaporação	Temperatura Máxima do Ar	Umidade Relativa do Ar
31/05/2002	89,3	30,487097	76,596774
30/06/2002	68,7	29,116667	79,908333
31/07/2002	73,4	26,322581	78,887097
31/08/2002	99,2	29,390323	76,008065
30/09/2002	79,6	26,266667	77,85
31/10/2002	124,2	31,051613	73,516129
30/11/2002	122,4	31,18	73,108333
31/12/2002	99,2	31,609677	78,403226
31/01/2003	90,2	32,532258	79,822581
28/02/2003	122,5	34,814286	72,625
31/03/2003	126,5		74,112903
30/04/2003	100,1	31,57	74,608333
31/05/2003	88,3	28,629032	77,209677
30/06/2003	90,4	29,77	75,916667
31/07/2003	90,8	27,854839	77,629032
31/08/2003	86,1		76,419355
30/09/2003	86,8	27,153333	76,55
31/10/2003	100,5	28,725806	74,830645
30/11/2003	96,2	30,343333	75,85
31/12/2003	93,3	31,345161	78,379032
31/01/2004	79,7	30,316129	80,346774
29/02/2004	87,7	31,244828	78,586207
31/03/2004	84,7	31,370968	78,653226
30/04/2004	77,1	30,773333	81,366667
31/05/2004		28,483871	78,556452
30/06/2004		27,343333	80,016667
31/07/2004		25,506452	81,362903
31/08/2004	86,2	27,653333	75,008621
30/09/2004	123,1	29,263333	71,625
31/10/2004	106,9	28,696774	74,629032
30/11/2004	122,3	29,913333	73,258333
31/12/2004	81,8	30,622581	81,516129
31/01/2005	97,6	31,670968	77,943548
28/02/2005	89,6	32,267857	75,5
31/03/2005	86	31,874194	79,370968
30/04/2005	84,3	31,68	78,658333
31/05/2005	76,6	28,922581	79,951613
30/06/2005	62	28,133333	82,391667
31/07/2005	76,7	26,264516	79,274194
31/08/2005	86,9	29,051613	76,733871
30/09/2005	74,8	26,86	79,15
31/10/2005	106,7	31,129032	74,258065
30/11/2005	73,2	28,64	79,2
31/12/2005	105,8	30,129032	76,193548
31/01/2006	137	33,180645	71,887097
28/02/2006	125,1	34,282143	71,705357
31/03/2006	107,5	32,793548	76,419355
30/04/2006	86,5	30,696667	77,4

31/05/2006	104,8	28,109677	73,233871
30/06/2006	100,7	27,02	75,033333
31/07/2006	107,9	28,574194	75,362903
31/08/2006	115,1	28,2	
30/09/2006	122,2	27,36	73,883333
31/10/2006	104,6	28,412903	78,959677
30/11/2006	90,3	29,176667	78,816667
31/12/2006	101,8	30,922581	77,75
31/01/2007	71,7	30,335484	82,620968
28/02/2007	107,1	31,896429	74,660714
31/03/2007	132	34,167742	70,822581
30/04/2007	80	31,286667	78,841667
31/05/2007	88,9		74,120968
30/06/2007	97,6	28,48	76,791667
31/07/2007	104	27,703226	74,629032
31/08/2007	118,1	28,593548	73,927419
30/09/2007	121,6	28,313333	72,7
31/10/2007	129,1	29,877419	73,024194
30/11/2007	92,2	29,696667	78,083333
31/12/2007	117,8	31,870968	75,467742
31/01/2008	112,6	29,967742	77,193548
29/02/2008	88,1	31,717241	78,956897
31/03/2008	73,2		
30/04/2008	75,5	31,353333	
31/05/2008	94	28,948387	75,145161
30/06/2008	84,4	27,273333	77,616667
31/07/2008	103,7	27,832258	75,266129
31/08/2008	132,9	29,180645	74,201613
30/09/2008	99,8	27,746667	73,966667
31/10/2008	99,5		78,620968
30/11/2008	71,5	28,313333	81,533333
31/12/2008	85,8	28,974194	80,806452
31/01/2009	103,4	31,290323	79,153226
28/02/2009	122,7	33,507143	74,303571
31/03/2009	127,9		76,072581
30/04/2009	93,2	30,2	75,775
31/05/2009	83,2	29,109677	
30/06/2009	58,4	26,76	76,583333
31/07/2009	38,5	27,212903	77,758065
31/08/2009	40,5	27,625806	75,637097
30/09/2009	99,1	29,7	77,425
31/10/2009	98,1	29,387097	77,524194
30/11/2009	144,9	33,14	74,041667
31/12/2009	125,1	32,103226	77,846774
31/01/2010	182	34,283871	69,806452
28/02/2010	176	34,592857	69,375
31/03/2010	119,7	33,2	77,306452
30/04/2010	106,1	30,76	76,375
31/05/2010	104,1	29,122581	77,758065
30/06/2010	104,8	26,846667	76,225
31/07/2010	90,5	27,380645	80,129032
31/08/2010	140,5	27,864516	72,556452
30/09/2010	159,2	29,043333	73,441667
31/10/2010	147,4	29,083871	72,596774

30/11/2010	100,2	29,553333	80,025
31/12/2010	112,1	32,735484	75,217742
31/01/2011	120,2	33,025806	
28/02/2011	145,8	34,321429	69,633929
31/03/2011	67,5	30,658065	80,701613
30/04/2011	75,3	30,873333	76,533333
31/05/2011	91		74,790323
30/06/2011	85,6	26,966667	74,591667
31/07/2011	102,8	27,619355	73,072581
31/08/2011	102,1	29,187097	72,629032
30/09/2011	125,3	27,973333	67,808333
31/10/2011	99,2	28,709677	77,225806
30/11/2011	90	28,046667	74,533333
31/12/2011	65,3	29,948387	79,975806
31/01/2012	99,7	30,729032	76,653226
29/02/2012	126,8	32,806897	71,793103
31/03/2012	123,6	32,709677	75,709677
30/04/2012	93,2	31,346667	76,316667
31/05/2012	71,3	27,393548	80,145161
30/06/2012	75,1	28,24	80,333333
31/07/2012	112,1	28,6	73,298387
31/08/2012	82,4	27,335484	75,806452
30/09/2012	110,7	28,406667	71
31/10/2012	144,8	31,367742	69,225806
30/11/2012	87,1	28,933333	78,566667
31/12/2012	136,5	34,580645	70,653226

ANEXO B- Dados Meteorológicos da Estação de Cordeiro

Série de dados utilizados da estação de Cordeiro			
Data	Evaporação	Temperatura Máxima do Ar	Umidade Relativa do Ar
31/05/2002	47	27,183871	81,040323
30/06/2002	41,2	26,72	81,791667
31/07/2002	37,2	24,325806	
31/08/2002	40,6	27,651613	74,846774
30/09/2002	41,1	24,77	78,233333
31/10/2002	88	30,722581	66,822581
30/11/2002		29,02	76,991667
31/12/2002	50,3	29,148387	80,806452
31/01/2003	44,1	29,545161	83,991935
28/02/2003	59,4	31,514286	75,839286
31/03/2003	47	29,432258	81,443548
30/04/2003	41,9	28,203333	80,258333
31/05/2003	43,4	25,593548	79,806452
30/06/2003	28,9	26,893333	79,766667
31/07/2003	49,5	25,322581	77,629032
31/08/2003	39,4	24	80,104839
30/09/2003	45,7	25,236667	78,591667
31/10/2003	60,9	27,003226	75,798387
30/11/2003	33,4	27,796667	78,475
31/12/2003	29,2	28,625806	
31/01/2004	35,3	28,029032	80,516129
29/02/2004	33,7	27,868966	82,818966
31/03/2004	40,3	28,045161	82,419355
30/04/2004	37	27,186667	84,516667
31/05/2004	33	24,716129	84,943548
30/06/2004	29,2	23,513333	86,208333
31/07/2004	32,4	22,445161	85,080645
31/08/2004	51,3	25,593548	76,991935
30/09/2004	57,4	28,296667	71,033333
31/10/2004	39,9		75,016129
30/11/2004	36	27,616667	75,691667
31/12/2004	32		84,620968
31/01/2005	28,4	28,6	82,153226
28/02/2005	26,5	29,125	82,357143
31/03/2005	40,8	29,512903	83,887097
30/04/2005	39,1	28,7	81,35
31/05/2005	29,8	26,296774	83,532258
30/06/2005	22	25,173333	85,816667
31/07/2005	27,3	23,393548	83,5
31/08/2005	39,5	26,387097	

30/09/2005	37,1	25,066667	80,075
31/10/2005	48,6	29,970968	72,774194
30/11/2005	38,6	26,496667	80,116667
31/12/2005	49,5	27,709677	79,854839
31/01/2006	70,3	30,780645	73,403226
28/02/2006	51,6	30,921429	77,223214
31/03/2006	46,4	29,654839	80,532258
30/04/2006	45,7	27,743333	79,466667
31/05/2006	39,3	24,912903	81,024194
30/06/2006	41,2	24,353333	79,191667
31/07/2006	57,6	25,835484	74,870968
31/08/2006	53,1	26,564516	74,153226
30/09/2006	61,8	25,673333	73,691667
31/10/2006	32,7	26,622581	80,048387
30/11/2006	30,6	27,143333	80,041667
31/12/2006	39	28,422581	80,524194
31/01/2007	30,7	27,606452	86,056452
28/02/2007	44,7	29,607143	79,455357
31/03/2007	47,3	31,087097	76,048387
30/04/2007	41,9	28,916667	81,333333
31/05/2007	31,2	25,351613	81,717742
30/06/2007	41,6	25,826667	82,041667
31/07/2007	51,8	25,406452	78,016129
31/08/2007	61,3	26,493548	74,225806
30/09/2007	45,2	27,69	70,95
31/10/2007	63,9	29,483871	73,991935
30/11/2007	43,8	27,73	79,191667
31/12/2007	54,6	29,790323	78
31/01/2008	46,9	27,919355	80,766129
29/02/2008	40,1	29,275862	81,577586
31/03/2008	43,5	28,925806	82,870968
30/04/2008	32	27,85	85,15
31/05/2008	29,5	25,345161	81,709677
30/06/2008	32,6	23,94	84,05
31/07/2008	49,3	25,096774	78,354839
31/08/2008	40,2	26,725806	74,596774
30/09/2008	55,2	26,263333	72,091667
31/10/2008	53,8	27,490323	78,5
30/11/2008	36,9	26,486667	82,941667
31/12/2008	37,8	26,409677	83,379032
31/01/2009	43,1	28,558065	82,016129
28/02/2009	47,8	30,546429	
31/03/2009	41,7	29,625806	83,459677
30/04/2009	36,9	27,27	83,583333

31/05/2009	36,5	25,516129	
30/06/2009	32,3	23,88	83,133333
31/07/2009	36,1	24,445161	
31/08/2009	47	24,858065	78,790323
30/09/2009	61,4	28,386667	75,941667
31/10/2009	46,9	27,619355	80,370968
30/11/2009	58,2	30,846667	77,841667
31/12/2009	41	28,722581	
31/01/2010	62,4	31,409677	77,346774
28/02/2010	73	32,317857	71,232143
31/03/2010	44	30,003226	
30/04/2010	47,9	27,81	80,325
31/05/2010	32	25,190323	84,758065
30/06/2010	37,3	23,646667	80,525
31/07/2010	36,6	24,735484	82,870968
31/08/2010	60,2	25,619355	71,120968
30/09/2010	61,3	26,953333	71,9
31/10/2010	55,9	26,877419	75,282258
30/11/2010	40,8	26,886667	82,758333
31/12/2010	50,4	29,954839	82,153226
31/01/2011	50,2	29,851613	80,540323
28/02/2011		31,592593	73,173077
31/03/2011	33,4	27,712903	86,733871
30/04/2011	41,2	27,66	83,875
31/05/2011	34,1	24,196774	84,056452
30/06/2011	32,6	23,213333	82,783333
31/07/2011	45,8	24,151613	80,233871
31/08/2011	61,8	26,903226	74,991935
30/09/2011	73,8	26,563333	67,982759
31/10/2011	59,7	26,751613	
30/11/2011	45,2	25,373333	
31/12/2011	38,4	26,906452	
31/01/2012	43,5	27,651613	
29/02/2012	56,6	30,151724	
31/03/2012	57,3	29,464516	
30/04/2012	36,1	27,323333	
31/05/2012	32,1	24,109677	
30/06/2012	34,5	25,02	
31/07/2012	45,4	25,319355	
31/08/2012	39,4	24,564516	
31/09/2012			
31/10/2012	78,1		
30/11/2012	37,3	26,64	
31/12/2012	60,5	31,864516	

ANEXO C- Dados Meteorológicos da Estação de Itaperuna

Série de dados utilizados da estação de Itaperuna			
Data	Evaporação	Temperatura Máxima do Ar	Umidade Relativa do Ar
31/05/2002	113,9	30,235484	
30/06/2002	82,8	28,533333	75,691667
31/07/2002	108,9	27,180645	72,758065
31/08/2002	159,5	30,835484	65,217742
30/09/2002	120	27,993333	71,158333
31/10/2002	168,3		64,467742
30/11/2002	117,4	31,646667	70,458333
31/12/2002	114,5	31,748387	74,822581
31/01/2003	94,7	32,248387	79,274194
28/02/2003	144,1	34,503571	66,464286
31/03/2003	138	32,832258	70,975
30/04/2003	95,8	30,47	75,508333
31/05/2003	86	28,170968	76,064516
30/06/2003	113,9	29,8	71,666667
31/07/2003	111,8	28,467742	69,637097
31/08/2003	120,9	27,487097	70,096774
30/09/2003	112,5	27,703333	71,708333
31/10/2003	135,7	29,509677	69,903226
30/11/2003	113,3	30,7	73,25
31/12/2003	86,9	31,512903	78,217742
31/01/2004	91,5	30,245161	
29/02/2004	80,4	30,734483	77,964286
31/03/2004	96,1	30,729032	76,016129
30/04/2004	71,1	30,08	78,841667
31/05/2004	70,9	28,025806	76,379032
30/06/2004	58,1	26,443333	80,058333
31/07/2004	67,2	25,351613	78,548387
31/08/2004	107,2	27,987097	71,556452
30/09/2004	167,1	31,713333	64,033333
31/10/2004	122	29,416129	70,5
30/11/2004	140,4	30,836667	69,55
31/12/2004	76,2	30,570968	79,7
31/01/2005	90,7	32,135484	76,120968
28/02/2005	80,5	31,692857	75,696429
31/03/2005	82,4	31,441935	78,169355
30/04/2005	87,4	31,406897	75,939655
31/05/2005	70,6	28,206452	80,790323
30/06/2005	47,4	26,833333	82,941667
31/07/2005	59,5	25,870968	79,991935
31/08/2005	100,5	29,354839	72,346774

30/09/2005	106,9	28,226667	73,841667
31/10/2005	150,7	32,319355	67,346774
30/11/2005	80,4	28,546667	77,225
31/12/2005	93,4	29,954839	75,091667
31/01/2006	146,8	34,087097	66,830645
28/02/2006	103	34,421429	69,740741
31/03/2006	84,8		
30/04/2006	66,7	29,803333	78,425
31/05/2006	82,3		75,604839
30/06/2006	87,3	27,25	72,566667
31/07/2006	108,2	28,806452	69,169355
31/08/2006	135,5	29,893548	
30/09/2006	133,3	28,973333	68,2
31/10/2006	113,3	29,651613	74,266129
30/11/2006	98,1	29,826667	78,025
31/12/2006	99	31,187097	77,975806
31/01/2007	78	30,883871	82,483871
28/02/2007	99,8	32,107143	74,401786
31/03/2007	150	34,193548	67,072581
30/04/2007	106,7	31,33	74,316667
31/05/2007	87,6	28,183871	74,75
30/06/2007	95,6	29,356667	71,991667
31/07/2007	125,1	29,125806	67,209677
31/08/2007	152,8	29,8	65,427419
30/09/2007	173,3	30,91	64,083333
31/10/2007	180,3	32,522581	65,919355
30/11/2007	109,5	30,54	73,6
31/12/2007	122,9	32,619355	72,975806
31/01/2008	112,4	30,845161	
29/02/2008	103,3	32,048276	74,991379
31/03/2008	109,4	31,809677	75,193548
30/04/2008	75,6	30,683333	80,533333
31/05/2008	98	28,190323	74,758065
30/06/2008	85,6	27,083333	76,383333
31/07/2008	118,2	27,974194	
31/08/2008	154,3	30,258065	68,040323
30/09/2008	143,4	29,093103	65,818966
31/10/2008	155,7	30,474194	70,540323
30/11/2008	110,2	29,006667	78,258333
31/12/2008	104,1	29,241935	79,790323
31/01/2009	105,7	31,445161	78,443548
28/02/2009	121,9	32,957143	73,964286
31/03/2009	133,4		74,094828
30/04/2009	85,7	29,22	78,575

31/05/2009	85,5	28,558065	76,55
30/06/2009	96,1	26,533333	76,666667
31/07/2009	109,9	27,758065	73,395161
31/08/2009	143,3	28,167742	70,443548
30/09/2009	143,7	31,21	69,075
31/10/2009	122,8	30,116129	75,782258
30/11/2009	142,9	33,92	68,791667
31/12/2009	99,2	32,167742	76,5
31/01/2010	159,6	35,254839	64,322581
28/02/2010	178,5	36,371429	61,883929
31/03/2010	108,7	33,029032	74,241935
30/04/2010	102,4	30,48	74,3
31/05/2010	87,4	29,170968	75,806452
30/06/2010	109,7	27,48	68,991667
31/07/2010	109,1	28,125806	73,008065
31/08/2010	146,6	28,945161	63,733871
30/09/2010	175,2	31,093333	62,333333
31/10/2010	148,9	29,590323	
30/11/2010	82,9	29,163333	79,85
31/12/2010	115,3	32,487097	75,290323
31/01/2011	129,3	33,335484	71,701613
28/02/2011	178,6	35,217857	64,517857
31/03/2011	72,3	30,003226	82,403226
30/04/2011	83,9	29,95	77,991667
31/05/2011	86,2	27,225806	75,903226
30/06/2011	79	26,45	75,333333
31/07/2011	111,5	27,606452	71,072581
31/08/2011	159,5	30,141935	65,137097
30/09/2011	176,4	29,033333	63,441667
31/10/2011	148,6	29,703226	72,274194
30/11/2011	113,3	28,453333	73,291667
31/12/2011	92,7	29,996774	78,798387
31/01/2012	115,3	31,058065	75,008065
29/02/2012	145,4	33,513793	65,715517
31/03/2012	137,1	32,493548	70,862903
30/04/2012	84,3	30,45	77,758333
31/05/2012	60,6	26,825806	81,798387
30/06/2012	63,3	27,783333	82,191667
31/07/2012	99	28,922581	74,024194
31/08/2012	111,8	27,767742	73,991935
31/09/2012	143,3	30,423333	65,941667
31/10/2012	188,4	33,341935	60,879032
30/11/2012	102,6	29,41	77,991667
31/12/2012	141,1	34,603226	68,75

ANEXO D- Dados Meteorológicos da Estação de Paty do Alferes

Série de dados utilizados da estação de Paty do Alferes			
Data	Evaporação	Temperatura Máxima do Ar	Umidade Relativa do Ar
31/05/2002	68,7	27,280645	84,524194
30/06/2002	78,6	27,183333	81,958333
31/07/2002	77,7	24,851613	81,596774
31/08/2002	115,3	28	77,475806
30/09/2002	86,8	24,646667	79
31/10/2002	154,3	31,119355	70,258065
30/11/2002	95,6	30	78,141667
31/12/2002	89,0	29,909677	
31/01/2003	75,1	30,103226	
28/02/2003	109,1	32,653571	76,517857
31/03/2003	78,9	30,119355	82,620968
30/04/2003	73,3	28,496667	82,058333
31/05/2003	75,0	26,041935	82,137097
30/06/2003	78,2	27,343333	82,233333
31/07/2003	86,2	25,267742	80,806452
31/08/2003	90,5	24,032258	79,548387
30/09/2003	97,8	26	78,041667
31/10/2003	102,6	27	76,975806
30/11/2003	84,1	27,886667	81,425
31/12/2003	85,1	29,296774	81,177419
31/01/2004	85,1	28,264516	81,548387
29/02/2004	75,6	28,696552	83,422414
31/03/2004	76,6	28,154839	82,056452
30/04/2004	64,5	28,003333	84,808333
31/05/2004	59,2	24,767742	86,379032
30/06/2004	57,8	23,956667	84,916667
31/07/2004	67,5	22,912903	
31/08/2004	98,3	25,829032	79,693548
30/09/2004	134,5	29,146667	74
31/10/2004	96,0	26,441935	79,266129
30/11/2004	105,4	27,893333	78
31/12/2004	75,4	28,351613	83,637097
31/01/2005	83,5	28,990323	83,483871
28/02/2005	76,1	29,560714	80,919643
31/03/2005	74,6		84,403226
30/04/2005	72,6	28,973333	83,425
31/05/2005	71,3	26,141935	84,193548
30/06/2005	58,4	25,066667	86
31/07/2005	65,4	23,409677	85,306452
31/08/2005	107,0	27,245161	78,491935

30/09/2005	80,3	25,153333	81,508333
31/10/2005	119,7	29,545161	
30/11/2005	89,4	27	76,358333
31/12/2005	80,4	27,951613	79,120968
31/01/2006	104,4	30,680645	77,274194
28/02/2006	64,5	31,057143	81,482143
31/03/2006	80,8	30,016129	78,677419
30/04/2006	79,4	28	77
31/05/2006	71,2	25,070968	77,362903
30/06/2006	71,7	24,446667	78,058333
31/07/2006	89,3	26,306452	75,475806
31/08/2006	105,3	26,974194	71,403226
30/09/2006	100,2	25,966667	71,658333
31/10/2006	95,9	27,403226	75,096774
30/11/2006	83,0	27,826667	77,075
31/12/2006	78,2	29,367742	79,314516
31/01/2007	55,4	28,490323	85,193548
28/02/2007	83,6	30,275	77,571429
31/03/2007	106,9	31,925806	74,104839
30/04/2007	80,4	29,453333	78,175
31/05/2007	73,2	25,306452	79,024194
30/06/2007	76,8	26,353333	79,033333
31/07/2007	86,8	25,135484	75,798387
31/08/2007	109,4	27,287097	71,217742
30/09/2007	137,2		64,775
31/10/2007	123,0		68,330645
30/11/2007	80,8		78,308333
31/12/2007	95,3	30,267742	75,733871
31/01/2008	79,6	28,635484	78,201613
29/02/2008	67,6	29,817241	81,525862
31/03/2008	76,4	29,580645	79,096774
30/04/2008	55,8	28	83,266667
31/05/2008	58,0	25,577419	81,741935
30/06/2008	54,1	24,016667	83
31/07/2008	88,6	25,751613	77,741935
31/08/2008	95,3	26,903226	
30/09/2008	96,9	26	72,875
31/10/2008	93,2	28,080645	76,282258
30/11/2008	66,5	27,086667	81,825
31/12/2008	69,5	27,567742	
31/01/2009	71,4	29,148387	
28/02/2009	77,3	31,121429	
31/03/2009	75,7	30,177419	81,112903
30/04/2009	63,4	27	80,383333

31/05/2009	64,7	26,309677	80,612903
30/06/2009	60,6	24	80,433333
31/07/2009	64,0		80,040323
31/08/2009	79,8	25,422581	76,806452
30/09/2009	95,0	28,826667	75,683333
31/10/2009	67,2	27,464516	80,895161
30/11/2009	90,3	31,266667	76,466667
31/12/2009	72,8	29,429032	81,008065
31/01/2010	90,0	31,625806	77,991935
28/02/2010	97,5	31,989286	74,669643
31/03/2010	67,2		81,439655
30/04/2010	59,1	27	81,808333
31/05/2010	58,1	25,674194	82,637097
30/06/2010	60,0	24,026667	81,875
31/07/2010	69,5	25,354839	80,967742
31/08/2010	95,8	26,067742	74,129032
30/09/2010	101,3	27,423333	74,791667
31/10/2010	86,7	27,183871	76,258065
30/11/2010	67,9	27,766667	82
31/12/2010	77,1	30,506452	81,225806
31/01/2011	78,8	30,593548	80,516129
28/02/2011	105,9	32,964286	74
31/03/2011	58,6	28,206452	83,620968
30/04/2011	65,0	28	82,358333
31/05/2011	60,5	24,767742	81,967742
30/06/2011	61,4	24	80,675
31/07/2011	76,1	25,245161	79,717742
31/08/2011	108,1	27,841935	72,072581
30/09/2011	119,9	27,466667	66
31/10/2011	106,4	27,341935	73,620968
30/11/2011	81,1	26	76,616667
31/12/2011	77,0	28,393548	79,016129
31/01/2012		28,487097	
29/02/2012	96,5	31,151724	
31/03/2012	83,3	30,035484	78,721154
30/04/2012	60,7	28,236667	82,266667
31/05/2012	47,6	24,819355	85,416667
30/06/2012	46,9	25,336667	85,383333
31/07/2012	66,2	25,148387	
31/08/2012	85,5		78,825
30/09/2012	116,2	27,943333	71,016667
31/10/2012	138,9	30,580645	69,651786
30/11/2012	86,5	28	79,325
31/12/2012	113,3	32,658065	76,291667

ANEXO E- Dados Meteorológicos da Estação de Resende

Série de dados utilizados da estação de Resende			
Data	Evaporação	Temperatura Máxima do Ar	Umidade Relativa do Ar
31/05/2002	67,2	25,906452	74,056452
30/06/2002	71,7	26,026667	71,5
31/07/2002	77,7	23,335484	74,016129
31/08/2002	120,4	26,735484	68,258065
30/09/2002	88	24,786667	72,558333
31/10/2002	168,4		61,274194
30/11/2002	88,4	28,686667	72,415
31/12/2002	83,5	29,419355	77,298387
31/01/2003	65,2	28,754839	82,201613
28/02/2003	110,3	32,678571	64,026786
31/03/2003	76,4	29,090323	74,548387
30/04/2003	74,4	27,393333	72,858333
31/05/2003	81,5	24,722581	72,217742
30/06/2003	73,1	26,153333	72,216667
31/07/2003	83,7	24,832258	70,451613
31/08/2003	90,1	23,974194	
30/09/2003	104,3	25,54	69,5
31/10/2003	111,6	27,406452	68,758065
30/11/2003	80,5	27,513333	76,232759
31/12/2003	78,1	29,045161	77,572581
31/01/2004	95,7		76,072581
29/02/2004	61,9	27,637931	82,706897
31/03/2004	72	28,048387	79,806452
30/04/2004	54	26,436667	83,008333
31/05/2004	56,2	22,793548	80,137097
30/06/2004	49,8	21,996667	79,516667
31/07/2004	55	21,064516	79,629032
31/08/2004	95,8	24,167742	69,225806
30/09/2004	128	28,49	67,416667
31/10/2004	77,5	25,9	78,459677
30/11/2004	89,3	27,885714	
31/12/2004	122,5	29,293548	
31/01/2005	97,5	29,445161	
28/02/2005	113,7	29,971429	78,169643
31/03/2005	95,4	29,870968	82,080645
30/04/2005	93,3	29,073333	81,783333
31/05/2005	98,6	26,519355	79,322581
30/06/2005	90,5	25,953333	79,775
31/07/2005	92,3	23,706452	79,967742
31/08/2005	153,4	28,309677	

30/09/2005	106,6	24,82	81,683333
31/10/2005	150,6	29,525806	74,693548
30/11/2005	127,1	27,896667	78,175
31/12/2005	130	28,445161	78,516129
31/01/2006	167,4	31,535484	
28/02/2006	84,3	31,007143	82,223214
31/03/2006	106,8	29,709677	80,725806
30/04/2006	115,5	28,173333	78,383333
31/05/2006	109,1	24,790323	79,443548
30/06/2006	109,2	25,226667	78,483333
31/07/2006	147	27,067742	72,258065
31/08/2006	159,1	27,712903	70,629032
30/09/2006	150,5	26,56	73,233333
31/10/2006	126,6	27,480645	79,959677
30/11/2006	122,7	28,163333	79,708333
31/12/2006	116,3	29,574194	81,354839
31/01/2007	76,6	28,619355	
28/02/2007	123,5	31,414286	75,883929
31/03/2007	184,3	32,593548	69,395161
30/04/2007	119,5	29,566667	78,15
31/05/2007	103,5	25,358065	80,096774
30/06/2007	106,8	26,46	79,516667
31/07/2007	110,9	24,551613	78,169355
31/08/2007	164,3	27,496774	69,112903
30/09/2007	230,2	30,01	59,183333
31/10/2007	202,9	30,732258	66,314516
30/11/2007	107,8	27,84	80,891667
31/12/2007	146,9	30,387097	75,435484
31/01/2008	111,3	28,406452	
29/02/2008	89,6	29,351724	81,206897
31/03/2008	118,8	29,387097	78,451613
30/04/2008	93	27,553333	82,058333
31/05/2008	106,2		77,903226
30/06/2008	65,9		84,075
31/07/2008	115,9		74,056452
31/08/2008	124,8	26,767742	74,666667
30/09/2008	127,4	25,703333	75,525
31/10/2008	129,6	27,819355	78,774194
30/11/2008			
31/12/2008			
31/01/2009			
28/02/2009			
31/03/2009			
30/04/2009			

31/05/2009			
30/06/2009			
31/07/2009			
31/08/2009	127,1	25,135484	75,516129
30/09/2009	120	27,586667	77,983333
31/10/2009	111,1	27,464516	80,854839
30/11/2009	128,9	31,953333	76,675
31/12/2009	105,3	29,6	81,798387
31/01/2010	108,9	31,741935	
28/02/2010	152,8	32,525	73,660714
31/03/2010	90,7	30,445161	80,193548
30/04/2010	58,2	28,266667	76,9
31/05/2010	91	26,296774	79,596774
30/06/2010	101,6	24,346667	77,458333
31/07/2010	111,1	25,780645	
31/08/2010	158,3	26,516129	68,169355
30/09/2010	168,5	28,336667	69,383333
31/10/2010	150,5	27,574194	73,064516
30/11/2010	114	28,03	79,933333
31/12/2010	122	30,935484	81,201613
31/01/2011	126,1	31,096774	
28/02/2011	139,6	32,707143	72,294643
31/03/2011	86,4	28,167742	84,080645
30/04/2011	113	28,766667	78,108333
31/05/2011	95,5	24,8	79,129032
30/06/2011	95,3	23,806667	
31/07/2011			
31/08/2011	160,2	28,441935	
30/09/2011	188,5	28,2	
31/10/2011	195,5	29,035484	
30/11/2011	146,4	27,633333	73,825
31/12/2011	125	28,722581	78,701613
31/01/2012	101,1	28,709677	
29/02/2012	160,2	32,475862	67,724138
31/03/2012	140,3	30,219355	77,653846
30/04/2012	94,2	28,06	81,191667
31/05/2012	90,9	25,096774	80,080645
30/06/2012	67,6	24,66	84,8
31/07/2012	96,2	25,083871	
31/08/2012	171,9		69,637097
30/09/2012	188,4	28,773333	69,338235
31/10/2012	212,3		65,225806
30/11/2012	135,7	27,983333	78,341667
31/12/2012	136,9	32,490323	77,983871

ANEXO F- Dados Meteorológicos da Estação do Rio de Janeiro

Série de dados utilizados da estação do Rio de Janeiro			
Data	Evaporação	Temperatura Máxima do Ar	Umidade Relativa do Ar
31/05/2002	94	29,45	76,169355
30/06/2002	85,7	29,59	74,983333
31/07/2002	92,8	26,47	73,75
31/08/2002	117,7	30,01	71,333333
30/09/2002	103,5	27,34	72,183333
31/10/2002	154,8	32,29	67,483871
30/11/2002	129	32,41	68,775
31/12/2002	129,3	32,83	72,435484
31/01/2003	113,5	32,82	75,580645
28/02/2003	174,2	37,24	63,633929
31/03/2003	119,6	32,689655	75,275862
30/04/2003	113,8	31,84	72
31/05/2003	105,9	28,448387	73,129032
30/06/2003		29,784211	73,223684
31/07/2003	81	28,093548	71,983871
31/08/2003	102,5	26,103226	72,846774
30/09/2003	102,3	27,77	73,658333
31/10/2003	109,7	29,164516	72,314516
30/11/2003	122,4	31,006667	71,766667
31/12/2003	116,6		72,637097
31/01/2004	128,6	31,464516	71,524194
29/02/2004	103,1	31,4	74,327586
31/03/2004	129,9	31,309677	
30/04/2004	91	31,17	75,241667
31/05/2004	94	27,596774	74,540323
30/06/2004	89,1	27,903333	71,116667
31/07/2004	82,6	26,490323	73,903226
31/08/2004	96,4	27,425806	74,112903
30/09/2004	120,6	30,266667	68,783333
31/10/2004	105,1	28,822581	71,653226
30/11/2004	131,4	30,686667	72,8
31/12/2004	115,5	31,074194	73,741935
31/01/2005	115,2	32,432258	73,064516
28/02/2005	102,7	32,064286	71,258929
31/03/2005	101,8	32,696774	74,693548
30/04/2005	102,2	32,326667	73,583333
31/05/2005	95,4	30,329032	73,483871
30/06/2005	83,6	29,083333	72,016667
31/07/2005	78,8	27,029032	75,056452
31/08/2005	114,1	30,9	68,693548

30/09/2005	80,8	26,83	76,791667
31/10/2005	109,6	31,222581	72,854839
30/11/2005	108,2	30,216667	70,783333
31/12/2005	129,3	31,729032	70,112903
31/01/2006	144,7	34,193548	68,056452
28/02/2006	107	33,971429	72,678571
31/03/2006	122	33,419355	72,241935
30/04/2006	98,3	30,833333	73,1
31/05/2006	92,8	27,487097	73,016129
30/06/2006	93	27,506667	72,283333
31/07/2006	97,6	28,825806	70
31/08/2006	110,5	29,251613	68,516129
30/09/2006	106,2	28,29	70,116667
31/10/2006	102,6	29,319355	72,645161
30/11/2006	99,8	30,126667	74,433333
31/12/2006			
31/01/2007			
28/02/2007	125	34,707692	66,305556
31/03/2007	150,8	35,316129	64,435484
30/04/2007	110,8	32,806667	71,483333
31/05/2007	94,6	28,632258	71,943548
30/06/2007	82,8	28,566667	72,783333
31/07/2007	94,9	27,341935	70,040323
31/08/2007	99,8	29,074194	69,258065
30/09/2007	118,8	29,926667	67,116667
31/10/2007	122,3		68,774194
30/11/2007	105,1	30,48	71,383333
31/12/2007	119,7	32,787097	70,822581
31/01/2008	122	32,009677	71,459677
29/02/2008	96	32,348276	73,862069
31/03/2008	104,9	32,270968	72,685484
30/04/2008	89,1	31,01	76,025
31/05/2008	91,1	28,470968	71,419355
30/06/2008	76,5	27,556667	74,35
31/07/2008	91,6	28,270968	69,959677
31/08/2008	110,9	29,519355	70,572581
30/09/2008	85	27,166667	72,283333
31/10/2008	106,3	30,180645	70,75
30/11/2008	86,1	28,856667	75,45
31/12/2008	97	29,993548	74,669355
31/01/2009	107,7	32,787097	
28/02/2009	124,1	34,807407	67,955357
31/03/2009	111,5	32,945161	71,758065
30/04/2009	86,6	30,356667	73,675

31/05/2009	94,7	29,935484	70,774194
30/06/2009	80,6	26,76	72,183333
31/07/2009	79,4	27,048387	74,217742
31/08/2009	90,9	28,4	73,241935
30/09/2009	99,6	30,31	71,775
31/10/2009	83,2	29,206452	
30/11/2009	142,5	35,33	67,808333
31/12/2009	100,3	32,283871	74,604839
31/01/2010	143	36,158065	67,169355
28/02/2010	160,3	36,589286	62,75
31/03/2010	109,1	33,216129	73,451613
30/04/2010	88	30,706667	74,566667
31/05/2010	86,4	29,183871	73,927419
30/06/2010	77,7	26,94	72,783333
31/07/2010	87,9	28,912903	71,919355
31/08/2010	104,4	28,009677	68,096774
30/09/2010	102,1	29,163333	70,308333
31/10/2010	105,8	28,416129	69,733871
30/11/2010	93,6	29,89	74,591667
31/12/2010	106,6	32,967742	75,774194
31/01/2011	122,7	34,735484	68,919355
28/02/2011	136,6	35,310714	
31/03/2011	76,3	30,567742	78,983871
30/04/2011	87,9	31,91	75,066667
31/05/2011	77,1	27,496774	77,266129
30/06/2011	78,1	26,966667	76,841667
31/07/2011	79,2	27,348387	74,266129
31/08/2011	102,3		70,879032
30/09/2011	97,4	27,3	69,833333
31/10/2011	104,9	29,370968	73,475806
30/11/2011	94,1	27,95	72,8
31/12/2011	106,6	31,058065	74,120968
31/01/2012	107,2	31,551613	75,879032
29/02/2012	132,6	34,837931	65,612069
31/03/2012	115,2	32,590323	74,362903
30/04/2012	86,6	30,893333	77,366667
31/05/2012	81,4	28,258065	
30/06/2012	70	28,3	79,033333
31/07/2012	93,4	28,296774	73,241935
31/08/2012	103	29,096774	71,137097
30/09/2012	114,6	29,603333	68,575
31/10/2012	138,4		69,056452
30/11/2012	90,9	28,873333	77,166667
31/12/2012	142,1		68,733871

APÊNDICES

APÊNDICE A - Configurações das RNAs utilizadas na estimativa dos dados de evaporação.

f1= Função da 1ª Camada , **f2**= Função da 2ª Camada , **lr**= Taxa de Aprendizado, **tm**= Taxa de momentum, **train**= Tipo de treinamento, **λ** = Fator de Auxilio

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de evaporação de Campos dos Goytacazes.

Modelos	f1	f2	train	lr	tm	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>satlin</i>	<i>logsig</i>	<i>traingd</i>	0,102	0,0153	Sim

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de evaporação da região de Cordeiro.

Modelos	f1	f2	train	lr	tm	(τ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>satlin</i>	<i>satlin</i>	<i>trainbfg</i>	0,6072	0,4857	Sim

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de evaporação da região de Itaperuna.

Modelos	f1	f2	train	lr	tm	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>purelin</i>	<i>trainrp</i>	0,5517	0,0552	Sim

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de evaporação da estação do Rio de Janeiro.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>logsig</i>	<i>satlin</i>	<i>trainbfg</i>	0,1831	0,0183	Não

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de evaporação da estação de Paty do Alferes.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>purelin</i>	<i>trainrpf</i>	0,2866	0,0344	Não

A tabela abaixo apresenta os parâmetros por cada metodologia aplicada no preenchimento dos dados de evaporação da estação de Resende.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>logsig</i>	<i>trainbfg</i>	0,0712	0,0107	Não

APÊNDICE B - Configurações das RNAs utilizadas na estimativa dos dados de temperatura máxima do ar.

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de temperatura máxima do ar da região de Campos dos Goytacazes.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>tansig</i>	<i>satlin</i>	<i>trainbfg</i>	0,0882	0,0106	Sim

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de temperatura máxima do ar da região de Cordeiro.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>satlin</i>	<i>satlin</i>	<i>trainrp</i>	0,688	0,5504	Sim

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de temperatura máxima do ar da região de Itaperuna.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>logsig</i>	<i>satlin</i>	<i>trainbfg</i>	0,2721	0,2177	Não

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de temperatura máxima do ar da região do Rio de Janeiro.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>logsig</i>	<i>trainlm</i>	0,0876	0,0131	Sim

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de temperatura máxima do ar da região de Paty do Alferes.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>satlin</i>	<i>traingd</i>	0,1616	0,0242	Não

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de temperatura máxima do ar da região de Resende.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>purelin</i>	<i>trainrp</i>	0,5726	0,0573	Não

APÊNDICE C - Configurações das RNAs utilizadas na estimativa dos dados de umidade relativa do ar.

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de umidade relativa do ar da região de Campos dos Goytacazes.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>satlin</i>	<i>trainbfg</i>	0,195	0,0292	Sim

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de umidade relativa do ar da região de Cordeiro.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>logsig</i>	<i>trainrp</i>	0,5605	0,056	Não

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de umidade relativa do ar da região de Itaperuna.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>purelin</i>	<i>trainbfg</i>	0,0094	0,0014	Sim

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de umidade relativa do ar da região do Rio de Janeiro.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>tansig</i>	<i>traingd</i>	0,4167	0,3334	Sim

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de umidade relativa do ar da região de Paty do Alferes.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>purelin</i>	<i>satlin</i>	<i>traingd</i>	0,0605	0,0091	Não

A tabela abaixo apresenta os parâmetros adotados por cada metodologia aplicada no preenchimento dos dados de umidade relativa do ar da região de Resende.

Modelos	<i>f1</i>	<i>f2</i>	<i>train</i>	<i>lr</i>	<i>tm</i>	(λ)
MED	X	X	X	X	X	X
RLM	X	X	X	X	X	X
MLP	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,45	X	X
RBF	<i>radbas</i>	X	X	X	X	X
MLP-AG	<i>tansig</i>	<i>logsig</i>	<i>trainbfg</i>	0,2178	0,0218	Sim

APÊNDICE D – Dados estimados de evaporação para estação de Campos Goytacazes com o modelo MLP-AG

Série de dados estimados da estação de Campos dos Goytacazes-(MLP-AG)						
DATA	Evaporação	Evaporação Est.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2010	104,80	95,43	9,37	87,86	0,09	0,88
31/07/2010	90,50	92,90	2,40	5,76	0,03	
31/08/2010	140,50	125,99	14,51	210,48	0,10	CONCORDÂNCIA (D)
30/09/2010	159,20	146,63	12,57	158,07	0,08	0,93
31/10/2010	147,40	138,68	8,72	76,08	0,06	
30/11/2010	100,20	97,62	2,58	6,65	0,03	CONFIANÇA (C)
31/12/2010	112,10	103,59	8,51	72,39	0,08	0,82
31/01/2011	120,20	110,85	9,35	87,37	0,08	
31/03/2011	67,50	80,30	12,80	163,88	0,19	DESEMPENHO
30/04/2011	75,30	98,96	23,66	559,63	0,31	Muito Bom
31/05/2011	91,00	91,04	0,04	0,00	0,00	
30/06/2011	85,60	84,99	0,61	0,37	0,01	
31/08/2011	102,10	124,74	22,64	512,73	0,22	
30/09/2011	125,30	146,17	20,87	435,48	0,17	
31/10/2011	99,20	119,65	20,45	418,17	0,21	
30/11/2011	90,00	94,74	4,74	22,45	0,05	
31/12/2011	65,30	93,60	28,30	801,12	0,43	
29/02/2012	126,80	124,16	2,64	7,00	0,02	
31/03/2012	123,60	122,96	0,64	0,40	0,01	
30/04/2012	93,20	91,15	2,05	4,20	0,02	
31/05/2012	71,30	74,53	3,23	10,45	0,05	
30/06/2012	75,10	76,83	1,73	2,99	0,02	
31/07/2012	112,10	103,86	8,24	67,83	0,07	
31/08/2012	82,40	90,38	7,98	63,60	0,10	
31/10/2012	144,80	145,98	1,18	1,40	0,01	
30/11/2012	87,10	90,86	3,76	14,16	0,04	
31/12/2012	136,50	116,65	19,85	394,02	0,15	
			EMA	REMQ	EMP	
			9,39	12,45	9,67	

APÊNDICE E – Dados estimados de evaporação para estação de Cordeiro com o modelo MLP-AG

Série de dados utilizados da estação de Cordeiro-(MLP-AG)						
DATA	Evaporação	Evaporação Est.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2010	37,30	44,65	7,35	54,04	0,20	0,96
31/07/2010	36,60	42,92	6,32	39,97	0,17	
31/08/2010	60,20	58,08	2,12	4,49	0,04	CONCORDÂNCIA (D)
30/09/2010	61,30	64,66	3,36	11,26	0,05	0,97
31/10/2010	55,90	58,29	2,39	5,73	0,04	
30/11/2010	40,80	37,64	3,16	9,97	0,08	CONFIANÇA (C)
31/12/2010	50,40	48,36	2,04	4,16	0,04	0,93
31/01/2011	50,20	52,87	2,67	7,12	0,05	
31/03/2011	33,40	32,99	0,41	0,17	0,01	DESEMPENHO
30/04/2011	41,20	35,95	5,25	27,54	0,13	Ótimo
31/05/2011	34,10	36,59	2,49	6,21	0,07	
30/06/2011	32,60	34,88	2,28	5,21	0,07	
31/08/2011	61,80	59,89	1,91	3,65	0,03	
30/09/2011	73,80	65,20	8,60	74,01	0,12	
31/10/2011	59,70	59,69	0,01	0,00	0,00	
30/11/2011	45,20	45,53	0,33	0,11	0,01	
31/12/2011	38,40	37,61	0,79	0,63	0,02	
29/02/2012	56,60	60,69	4,09	16,72	0,07	
31/03/2012	57,30	55,03	2,27	5,14	0,04	
30/04/2012	36,10	36,68	0,58	0,33	0,02	
31/05/2012	32,10	32,40	0,30	0,09	0,01	
30/06/2012	34,50	32,22	2,28	5,18	0,07	
31/07/2012	45,40	43,32	2,08	4,33	0,05	
31/08/2012	39,40	47,19	7,79	60,62	0,20	
31/10/2012	78,10	72,25	5,85	34,19	0,07	
30/11/2012	37,30	41,30	4,00	15,99	0,11	
31/12/2012	60,50	58,31	2,19	4,78	0,04	
			EMA	REMQ	EMP	
			3,07	3,86	6,65	

APÊNDICE F – Dados estimados de evaporação para estação de Itaperuna com o modelo MLP-AG

Série de dados utilizados da estação de Itaperuna-(MLP-AG)						
DATA	Evaporação	Evaporação Est.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2010	109,70	91,3356	18,36	337,25	0,17	0,97
31/07/2010	109,10	93,0276	16,07	258,32	0,15	
31/08/2010	146,60	147,3126	0,71	0,51	0,00	CONCORDÂNCIA (D)
30/09/2010	175,20	159,5514	15,65	244,88	0,09	0,98
31/10/2010	148,90	138,0771	10,82	117,14	0,07	
30/11/2010	82,90	96,8769	13,98	195,35	0,17	CONFIANÇA (C)
31/12/2010	115,30	113,0919	2,21	4,88	0,02	0,96
31/01/2011	129,30	114,1353	15,16	229,97	0,12	
31/03/2011	72,30	75,0078	2,71	7,33	0,04	DESEMPENHO
30/04/2011	83,90	88,149	4,25	18,05	0,05	Ótimo
31/05/2011	86,20	84,8496	1,35	1,82	0,02	
30/06/2011	79,00	82,65	3,65	13,32	0,05	
31/08/2011	159,50	147,1998	12,30	151,29	0,08	
30/09/2011	176,40	175,6677	0,73	0,54	0,00	
31/10/2011	148,60	145,7757	2,82	7,98	0,02	
30/11/2011	113,30	110,1168	3,18	10,13	0,03	
31/12/2011	92,70	90,2922	2,41	5,80	0,03	
29/02/2012	145,40	136,0044	9,40	88,28	0,06	
31/03/2012	137,10	126,4446	10,66	113,54	0,08	
30/04/2012	84,30	85,2444	0,94	0,89	0,01	
31/05/2012	60,60	65,7159	5,12	26,17	0,08	
30/06/2012	63,30	68,2962	5,00	24,96	0,08	
31/07/2012	99,00	101,0505	2,05	4,20	0,02	
31/08/2012	111,80	107,5929	4,21	17,70	0,04	
31/10/2012	188,40	194,4207	6,02	36,25	0,03	
30/11/2012	102,60	107,6916	5,09	25,92	0,05	
31/12/2012	141,10	151,7823	10,68	114,11	0,08	
			EMA	REMQ	EMP	
			6,87	8,73	6,01	

APÊNDICE G – Dados estimados de evaporação para estação do Rio de Janeiro com o modelo MLP

Série de dados utilizados da estação do Rio de Janeiro - (MLP)						
DATA	Evaporação	Evaporação Est.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2010	77,7	84,62	6,92	47,88	0,09	0,84
31/07/2010	87,9	90,13	2,23	4,98	0,03	
31/08/2010	104,4	109,47	5,07	25,71	0,05	CONCORDÂNCIA (D)
30/09/2010	102,1	110,05	7,95	63,27	0,08	0,91
31/10/2010	105,8	104,31	1,49	2,21	0,01	
30/11/2010	93,6	94,39	0,79	0,63	0,01	CONFIANÇA (C)
31/12/2010	106,6	104,75	1,85	3,42	0,02	0,76
31/01/2011	122,7	103,75	18,95	359,09	0,15	
31/03/2011	76,3	83,84	7,54	56,82	0,10	DESEMPENHO
30/04/2011	87,9	93,27	5,37	28,81	0,06	Muito Bom
31/05/2011	77,1	82,68	5,58	31,15	0,07	
30/06/2011	78,1	82,15	4,05	16,40	0,05	
31/08/2011	102,3	114,86	12,56	157,71	0,12	
30/09/2011	97,4	119,33	21,93	480,85	0,23	
31/10/2011	104,9	106,67	1,77	3,13	0,02	
30/11/2011	94,1	100,11	6,01	36,17	0,06	
31/12/2011	106,6	97,10	9,50	90,20	0,09	
29/02/2012	132,6	108,54	24,06	578,71	0,18	
31/03/2012	115,2	103,47	11,73	137,62	0,10	
30/04/2012	86,6	85,28	1,32	1,75	0,02	
31/05/2012	81,4	68,90	12,50	156,36	0,15	
30/06/2012	70	83,21	13,21	174,57	0,19	
31/07/2012	93,4	99,14	5,74	33,00	0,06	
31/08/2012	103	101,89	1,11	1,24	0,01	
31/10/2012	138,4	139,79	1,39	1,94	0,01	
30/11/2012	90,9	99,87	8,97	80,54	0,10	
31/12/2012	142,1	131,93	10,17	103,51	0,07	
			EMA	REMQ	EMP	
			7,77	9,96	7,89	

APÊNDICE H – Dados estimados de evaporação para estação de Paty do Alferes com o modelo MLP-AG

Série de dados utilizados da estação de Paty do Alferes-(MLP-AG)						
DATA	Evaporação	Evaporação Est.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2010	60	72,73	12,73	162,05	0,21	0,95
31/07/2010	69,5	77,15	7,65	58,59	0,11	
31/08/2010	95,8	95,40	0,40	0,16	0,00	CONCORDÂNCIA (D)
30/09/2010	101,3	108,47	7,17	51,44	0,07	0,97
31/10/2010	86,7	96,05	9,35	87,35	0,11	
30/11/2010	67,9	62,03	5,87	34,43	0,09	CONFIANÇA (C)
31/12/2010	77,1	82,23	5,13	26,36	0,07	0,93
31/01/2011	78,8	93,18	14,38	206,75	0,18	
31/03/2011	58,60	56,47	2,13	4,54	0,04	DESEMPENHO
30/04/2011	65	64,93	0,07	0,00	0,00	Ótimo
31/05/2011	60,5	61,12	0,62	0,38	0,01	
30/06/2011	61,4	58,05	3,35	11,24	0,05	
31/08/2011	108,1	108,13	0,03	0,00	0,00	
30/09/2011	119,9	113,10	6,80	46,22	0,06	
31/10/2011	106,4	102,95	3,45	11,89	0,03	
30/11/2011	81,1	81,24	0,14	0,02	0,00	
31/12/2011	77	76,31	0,69	0,48	0,01	
29/02/2012	96,5	103,99	7,49	56,16	0,08	
31/03/2012	83,30	95,25	11,95	142,84	0,14	
30/04/2012	60,7	62,12	1,42	2,01	0,02	
31/05/2012	47,6	50,44	2,84	8,09	0,06	
30/06/2012	46,9	48,52	1,62	2,63	0,03	
31/07/2012	66,2	69,46	3,26	10,66	0,05	
31/08/2012	85,5	83,80	1,70	2,88	0,02	
31/10/2012	138,9	127,54	11,36	129,14	0,08	
30/11/2012	86,5	74,70	11,80	139,36	0,14	
31/12/2012	113,3	102,47	10,83	117,32	0,10	
			EMA	REMQ	EMP	
			5,34	6,97	6,54	

APÊNDICE I – Dados estimados de evaporação para estação de Resende com o modelo MLP-AG

Série de dados utilizados da estação de Resende - (MLP-AG)						
DATA	Evaporação	Evaporação Est.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2010	101,60	104,79	3,19	10,16	0,03	0,88
31/07/2010	111,10	106,98	4,12	16,98	0,04	
31/08/2010	158,30	153,34	4,96	24,59	0,03	CONCORDÂNCIA (D)
30/09/2010	168,50	167,50	1,00	0,99	0,01	0,92
31/10/2010	150,50	149,56	0,94	0,89	0,01	
30/11/2010	114,00	107,31	6,69	44,82	0,06	CONFIANÇA (C)
31/12/2010	122,00	128,31	6,31	39,86	0,05	0,82
31/01/2011	126,10	135,18	9,08	82,39	0,07	
31/03/2011	86,40	90,43	4,03	16,22	0,05	DESEMPENHO
30/04/2011	113,00	107,60	5,40	29,21	0,05	Muito Bom
31/05/2011	95,50	96,10	0,60	0,35	0,01	
30/06/2011	95,30	91,77	3,53	12,48	0,04	
31/08/2011	160,20	162,27	2,07	4,28	0,01	
30/09/2011	188,50	183,06	5,44	29,60	0,03	
31/10/2011	195,50	153,96	41,54	1725,83	0,21	
30/11/2011	146,40	120,60	25,80	665,73	0,18	
31/12/2011	125,00	108,01	16,99	288,60	0,14	
29/02/2012	160,20	150,30	9,90	98,04	0,06	
31/03/2012	140,30	145,19	4,89	23,93	0,03	
30/04/2012	94,20	100,55	6,35	40,33	0,07	
31/05/2012	90,90	81,77	9,13	83,35	0,10	
30/06/2012	67,60	87,58	19,98	399,35	0,30	
31/07/2012	96,20	117,37	21,17	448,37	0,22	
31/08/2012	171,90	115,24	56,66	3210,61	0,33	
31/10/2012	212,30	192,95	19,35	374,51	0,09	
30/11/2012	135,70	107,87	27,83	774,68	0,21	
31/12/2012	136,90	149,77	12,87	165,73	0,09	
			EMA	REMQ	EMP	
			12,22	17,86	9,25	

APÊNDICE J – Dados estimados de temperatura máxima do ar para estação de Campos dos Goytacazes com o modelo MLP-AG

Série de dados utilizados da estação de Campos dos Goytacazes-(MLP-AG)						
DATA	Temp. Máx. do Ar	Temp. Máx. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
31/07/2010	27,380645	27,39598829	0,015343293	0,000235417	0,000560370	0,99
31/08/2010	27,864516	27,86406491	0,000451086	2,03E-07	0,000016189	
30/09/2010	29,043333	28,96773023	0,075602769	0,005715779	0,002603102	CONCORDÂNCIA (D)
31/10/2010	29,083871	28,5551719	0,528699098	0,279522736	0,018178429	0,99
30/11/2010	29,553333	29,35635826	0,196974739	0,038799048	0,006665060	
31/12/2010	32,735484	32,69032731	0,045156692	0,002039127	0,001379442	CONFIANÇA (C)
31/01/2011	33,025806	32,80327831	0,222527683	0,04951857	0,006737994	0,98
28/02/2011	34,321429	34,06201196	0,259417041	0,067297201	0,007558457	
31/03/2011	30,658065	30,26379514	0,394269857	0,155448721	0,012860233	DESEMPENHO
30/04/2011	30,873333	30,78739005	0,085942945	0,00738619	0,002783727	Ótimo
30/06/2011	26,966667	26,34496962	0,621697374	0,386507625	0,023054290	
30/09/2011	27,973333	27,99041688	0,017083878	0,000291859	0,000610720	
31/10/2011	28,709677	28,98400283	0,274325836	0,075254665	0,009555170	
30/11/2011	28,046667	27,72335476	0,323312237	0,104530803	0,011527653	
31/12/2011	29,948387	29,89622577	0,052161224	0,002720793	0,001741704	
31/01/2012	30,729032	30,69645492	0,03257708	0,001061266	0,001060140	
29/02/2012	32,806897	33,02247984	0,215582843	0,046475962	0,006571266	
31/03/2012	32,709677	32,01453571	0,695141295	0,483221421	0,021251854	
30/04/2012	31,346667	30,20349079	1,143176204	1,306851834	0,036468828	
31/05/2012	27,393548	26,94514149	0,448406518	0,201068405	0,016369056	
30/06/2012	28,24	27,56254316	0,677456834	0,458947762	0,023989265	
31/07/2012	28,6	27,75589997	0,844100036	0,712504871	0,029513987	
30/11/2012	28,933333	28,68248108	0,250851918	0,062926685	0,008669997	
			EMA	REMQ	EMP	
			0,32	0,44	1,09	

APÊNDICE L – Dados estimados de temperatura máxima do ar para estação de Cordeiro com o modelo MLP-AG

Série de dados utilizados da estação de Cordeiro-(MLP-AG)						
DATA	Temp. Máx. do Ar	Temp. Máx. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
31/07/2010	24,735484	25,18088506	0,445401062	0,198382106	0,01800656	0,99
31/08/2010	25,619355	25,77423409	0,154879093	0,023987534	0,00604539	
30/09/2010	26,953333	27,25415122	0,300818224	0,090491604	0,01116071	CONCORDÂNCIA (D)
31/10/2010	26,877419	26,75755461	0,119864382	0,01436747	0,00445967	0,99
30/11/2010	26,886667	26,97475392	0,088086928	0,007759307	0,00327623	
31/12/2010	29,954839	29,64433092	0,310508077	0,096415266	0,01036587	CONFIANÇA (C)
31/01/2011	29,851613	30,45882834	0,60721534	0,368710469	0,02034112	0,98
28/02/2011	31,592593	31,58826477	0,004328239	1,87E-05	0,000137	
31/03/2011	27,712903	27,55428118	0,158621823	0,025160883	0,00572375	DESEMPENHO
30/04/2011	27,66	27,50096862	0,159031381	0,02529098	0,00574951	Ótimo
30/06/2011	23,213333	23,82042755	0,607094557	0,368563801	0,02615284	
30/09/2011	26,563333	26,53936803	0,023964972	0,00057432	0,00090218	
31/10/2011	26,751613	26,78124908	0,029636086	0,000878298	0,00110782	
30/11/2011	25,373333	25,71203611	0,338703105	0,114719793	0,01334878	
31/12/2011	26,906452	27,59278469	0,686332697	0,471052571	0,02550811	
31/01/2012	27,651613	28,25326806	0,601655057	0,361988808	0,02175841	
29/02/2012	30,151724	30,8142454	0,662521403	0,438934609	0,02197292	
31/03/2012	29,464516	29,50611318	0,041597183	0,001730326	0,00141177	
30/04/2012	27,323333	27,8751438	0,551810797	0,304495155	0,02019559	
31/05/2012	24,109677	24,57272699	0,463049987	0,214415291	0,01920598	
30/06/2012	25,02	25,28257383	0,262573828	0,068945015	0,01049456	
31/07/2012	25,319355	25,31219192	0,007163085	5,13E-05	0,00028291	
30/11/2012	26,64	27,09520082	0,455200819	0,207207786	0,01708712	
			EMA	REMQ	EMP	
			0,31	0,38	1,15	

APÊNDICE M – Dados estimados de temperatura máxima do ar para estação de Itaperuna com o modelo MLP-AG

Série de dados utilizados da estação de Itaperuna-(MLP-AG)						
DATA	Temp. Máx. do Ar	Temp. Máx. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
31/07/2010	28,125806	27,47292758	0,652878421	0,426250232	0,02321279	0,98
31/08/2010	28,945161	28,72257471	0,222586288	0,049544656	0,00768993	
30/09/2010	31,093333	29,74080571	1,352527288	1,829330065	0,04349895	CONCORDÂNCIA (D)
31/10/2010	29,590323	29,74631562	0,155992625	0,024333699	0,00527174	0,99
30/11/2010	29,16333299	29,32095072	0,157617729	0,024843348	0,00540465	
31/12/2010	32,487097	33,19662001	0,709523008	0,503422899	0,02184015	CONFIANÇA (C)
31/01/2011	33,335484	33,45007578	0,114591783	0,013131277	0,00343753	0,97
28/02/2011	35,217857	35,26173353	0,043876532	0,00192515	0,00124586	
31/03/2011	30,003226	30,32485596	0,321629959	0,10344583	0,01071985	DESEMPENHO
30/04/2011	29,95	30,54084435	0,59084435	0,349097046	0,01972769	Ótimo
30/06/2011	26,45	26,29490925	0,155090753	0,024053142	0,00586354	
30/09/2011	29,033333	29,08182072	0,048487719	0,002351059	0,00167007	
31/10/2011	29,703226	29,56118271	0,14204329	0,020176296	0,00478208	
30/11/2011	28,453333	28,53854379	0,085210782	0,007260877	0,00299476	
31/12/2011	29,996774	29,89508314	0,101690862	0,010341031	0,00339006	
31/01/2012	31,05806501	30,44607394	0,611991069	0,374533068	0,01970474	
29/02/2012	33,513793	33,6285968	0,114803793	0,013179911	0,00342557	
31/03/2012	32,493548	32,76354124	0,269993239	0,072896349	0,00830913	
30/04/2012	30,45	30,30832624	0,141673763	0,020071455	0,00465267	
31/05/2012	26,825806	27,11368158	0,287875579	0,082872349	0,01073129	
30/06/2012	27,783333	28,14623834	0,362905341	0,131700286	0,01306198	
31/07/2012	28,922581	28,33577917	0,58680183	0,344336388	0,02028871	
30/11/2012	29,41	29,54906091	0,139060909	0,019337937	0,00472835	
			EMA	REMQ	EMP	
			0,32	0,44	1,07	

APÊNDICE N – Dados estimados de temperatura máxima do ar para estação do Rio de Janeiro com o modelo MLP-AG

Série de dados utilizados da estação de Rio de Janeiro - (MLP-AG)						
DATA	Temp. Máx. do Ar	Temp. Máx. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
31/07/2010	28,912903	27,62228197	1,290621	1,665703	0,04463824	0,97
31/08/2010	28,009677	27,92074752	0,0889295	0,007908	0,00317496	
30/09/2010	29,163333	29,83961368	0,6762807	0,457356	0,02318942	CONCORDÂNCIA (D)
31/10/2010	28,416129	29,13020117	0,7140722	0,509899	0,02512912	0,99
30/11/2010	29,890000	29,54560284	0,3443972	0,118609	0,01152215	
31/12/2010	32,967742	33,18621426	0,2184723	0,047730	0,00662685	CONFIANÇA (C)
31/01/2011	34,735484	34,12281696	0,612667	0,375361	0,01763807	0,96
28/02/2011	35,310714	35,66526016	0,3545462	0,125703	0,01004075	
31/03/2011	30,567742	30,73389663	0,1661546	0,027607	0,00543562	DESEMPENHO
30/04/2011	31,910000	30,81519508	1,0948049	1,198598	0,03430915	Ótimo
30/06/2011	26,966667	27,62673668	0,6600697	0,435692	0,02447724	
30/09/2011	27,300000	28,00093229	0,7009323	0,491306	0,02567518	
31/10/2011	29,370968	28,7693697	0,6015983	0,361921	0,02048275	
30/11/2011	27,950000	27,99759126	0,0475913	0,002265	0,00170273	
31/12/2011	31,058065	30,18485367	0,8732113	0,762498	0,02811545	
31/01/2012	31,551613	31,44776384	0,1038492	0,010785	0,00329141	
29/02/2012	34,837931	34,36559863	0,4723324	0,223098	0,01355799	
31/03/2012	32,590323	33,19957839	0,6092554	0,371192	0,01869437	
30/04/2012	30,893333	31,25732501	0,363992	0,132490	0,01178222	
31/05/2012	28,258065	27,74812752	0,5099375	0,260036	0,01804573	
30/06/2012	28,300000	28,1924848	0,1075152	0,011560	0,00379912	
31/07/2012	28,296774	28,50208712	0,2053131	0,042153	0,00725571	
30/11/2012	28,873333	28,9219435	0,0486105	0,002363	0,00168358	
			EMA	REMQ	EMP	
			0,47	0,58	1,57	

APÊNDICE O – Dados estimados de temperatura máxima do ar para estação de Paty do Alferes com o modelo MLP-AG

Série de dados utilizados da estação de Paty do Alferes-(MLP-AG)						
DATA	Temp. Máx. do Ar	Temp. Máx. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
31/07/2010	25,354839	25,46394401	0,109105	0,011903903	0,00430312	0,99
31/08/2010	26,067742	25,97857482	0,0891672	0,007950787	0,00342059	
30/09/2010	27,423333	27,60488858	0,1815556	0,032962432	0,00662048	CONCORDÂNCIA (D)
31/10/2010	27,183871	27,24203366	0,0581627	0,003382895	0,0021396	0,99
30/11/2010	27,766667	27,3928044	0,3738626	0,139773242	0,01346444	
31/12/2010	30,506452	30,51476396	0,008312	6,90887E-05	0,00027247	CONFIANÇA (C)
31/01/2011	30,593548	30,7389098	0,1453618	0,021130054	0,00475139	0,98
28/02/2011	32,964286	32,06267694	0,9016091	0,812898888	0,02735109	
31/03/2011	28,206452	28,11851426	0,0879377	0,007733046	0,00311765	DESEMPENHO
30/04/2011	28	28,25320279	0,2532028	0,064111653	0,00904296	Ótimo
30/06/2011	24	23,73510613	0,2648939	0,07016876	0,01103724	
30/09/2011	27,466667	26,78369059	0,6829764	0,466456774	0,02486565	
31/10/2011	27,341935	27,31440362	0,0275314	0,000757977	0,00100693	
30/11/2011	26	26,04089339	0,0408934	0,001672269	0,00157282	
31/12/2011	28,393548	27,76470557	0,6288424	0,395442793	0,02214737	
31/01/2012	28,487097	28,37884508	0,1082519	0,01171848	0,00380003	
29/02/2012	31,151724	30,99924062	0,1524834	0,023251181	0,00489486	
31/03/2012	30,035484	30,28157188	0,2460879	0,060559244	0,00819324	
30/04/2012	28,236667	28,30949053	0,0728235	0,005303267	0,00257904	
31/05/2012	24,819355	24,75029581	0,0690592	0,004769171	0,00278247	
30/06/2012	25,336667	25,14129461	0,1953724	0,038170371	0,00771105	
31/07/2012	25,148387	25,74437759	0,5959906	0,355204788	0,02369896	
30/11/2012	28	27,16162259	0,8383774	0,70287667	0,02994205	
			EMA	REMQ	EMP	
			0,27	0,38	0,95	

APÊNDICE P – Dados estimados de temperatura máxima do ar para estação de Resende com o modelo MLP-AG

Série de dados utilizados da estação de Resende - (MLP-AG)						
DATA	Temp. Máx. do Ar	Temp. Máx. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
31/07/2010	25,7806450	25,75533042	0,0253146	0,0006408	0,00098192	0,97
31/08/2010	26,5161290	26,47950182	0,0366272	0,0013416	0,00138132	
30/09/2010	28,3366670	27,85798885	0,4786781	0,2291328	0,01689254	CONCORDÂNCIA (D)
31/10/2010	27,5741940	27,61465795	0,040464	0,0016373	0,00146746	0,98
30/11/2010	28,0300000	28,2061034	0,1761034	0,0310124	0,00628268	
31/12/2010	30,9354840	30,98985552	0,0543715	0,0029563	0,00175758	CONFIANÇA (C)
31/01/2011	31,0967740	31,07833948	0,0184345	0,0003398	0,00059281	0,95
28/02/2011	32,7071430	33,48719901	0,780056	0,6084874	0,02384971	
31/03/2011	28,1677420	28,65318028	0,4854383	0,2356503	0,01723384	DESEMPENHO
30/04/2011	28,7666670	28,44361299	0,323054	0,1043639	0,01123015	Ótimo
30/06/2011	23,8066670	24,37917191	0,5725049	0,3277619	0,02404809	
30/09/2011	28,2000000	27,90106657	0,2989334	0,0893612	0,01060048	
31/10/2011	29,0354840	27,77416194	1,2613221	1,5909334	0,04344071	
30/11/2011	27,6333330	26,41081032	1,2225227	1,4945617	0,04424087	
31/12/2011	28,7225810	28,8429551	0,1203741	0,0144899	0,00419092	
31/01/2012	28,7096770	28,93842464	0,2287476	0,0523255	0,00796761	
29/02/2012	32,4758620	31,64533542	0,8305266	0,6897744	0,02557366	
31/03/2012	30,2193550	30,51134355	0,2919886	0,0852573	0,0096623	
30/04/2012	28,0600000	28,68345111	0,6234511	0,3886913	0,0222185	
31/05/2012	25,0967740	25,21161974	0,1148457	0,0131895	0,00457612	
30/06/2012	24,6600000	25,73670222	1,0767022	1,1592877	0,04366189	
31/07/2012	25,0838710	25,54576313	0,4618921	0,2133443	0,01841391	
30/11/2012	27,9833330	28,44361299	0,46028	0,2118577	0,01644836	
			EMA	REMQ	EMP	
			0,43	0,57	1,55	

APÊNDICE Q – Dados estimados de umidade relativa do ar para estação de Campos dos Goytacazes com o modelo MLP-AG

Série de dados utilizados da estação de Campos dos Goytacazes (MLP-AG)						
DATA	Umid. Relat. do Ar	Umid. Relat. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2008	77,616667	77,1773685	0,439299	0,192983177	0,0056598	0,89
31/10/2008	78,62096799	76,45427824	2,16669	4,694544503	0,0275587	
31/08/2009	75,637097	76,46984717	0,83275	0,693472837	0,0110098	CONCORDÂNCIA (D)
30/09/2009	77,42499999	76,77084646	0,654154	0,427916843	0,0084489	0,94
30/11/2009	74,04166701	73,15366529	0,888002	0,788547049	0,0119933	
28/02/2010	69,375	69,88764996	0,51265	0,262809979	0,0073895	CONFIANÇA (C)
30/04/2010	76,37500001	76,09619287	0,278807	0,077733419	0,0036505	0,84
31/05/2010	77,758065	76,54250217	1,215563	1,477592998	0,0156326	
30/06/2010	76,225	76,11003192	0,114968	0,01321766	0,0015083	DESEMPENHO
31/08/2010	72,556452	72,46171289	0,094739	0,008975499	0,0013057	Muito Bom
30/09/2010	73,441667	73,26437767	0,177289	0,031431506	0,002414	
30/11/2010	80,02499999	80,08356858	0,058569	0,003430279	0,0007319	
31/12/2010	75,217742	76,8400417	1,6223	2,631856319	0,021568	
31/03/2011	80,701613	82,90500449	2,203391	4,854934028	0,0273029	
30/04/2011	76,533333	78,42807246	1,894739	3,590037603	0,024757	
31/05/2011	74,790323	78,10458471	3,314262	10,98433068	0,044314	
			EMA	REMQ	EMP	
			1,03	1,39	1,35	

APÊNDICE R – Dados estimados de umidade relativa do ar para estação de Cordeiro com o modelo MLP-AG

Série de dados utilizados da estação de Cordeiro-(MLP-AG)						
DATA	Umid. Relat. do Ar	Umid. Relat. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2008	84,05	82,63310271	1,416897285	2,007597916	0,01685779	0,97
31/10/2008	78,50000001	77,97012856	0,529871449	0,280763753	0,006749955	
31/08/2009	78,790323	78,91792004	0,127597049	0,016281007	0,001619451	CONCORDÂNCIA (D)
30/09/2009	75,941667	77,05832915	1,116662156	1,246934371	0,014704209	0,98
30/11/2009	77,841667	75,9985623	1,8431047	3,397034935	0,023677611	
28/02/2010	71,232143	70,8716902	0,360452801	0,129926222	0,005060255	CONFIANÇA (C)
30/04/2010	80,32500001	81,83127911	1,506279101	2,26887673	0,018752308	0,95
31/05/2010	84,758065	82,25718541	2,500879591	6,254398726	0,029506096	
30/06/2010	80,52500001	77,57021654	2,954783472	8,730745369	0,036693989	DESEMPENHO
31/08/2010	71,120968	70,8616924	0,2592756	0,067223837	0,003645558	Ótimo
30/09/2010	71,9	71,34358639	0,55641361	0,309596105	0,007738715	
30/11/2010	82,758333	84,14077103	1,382438026	1,911134895	0,016704518	
31/12/2010	82,153226	82,83305872	0,679832718	0,462172524	0,00827518	
31/03/2011	86,73387101	86,15632761	0,577543401	0,33355638	0,006658799	
30/04/2011	83,875	83,47491751	0,400082482	0,160065993	0,004769985	
31/05/2011	84,056452	83,80484493	0,251607067	0,063306116	0,002993311	
			EMA	REMQ	EMP	
			1,03	1,31	1,28	

APÊNDICE S – Dados estimados de umidade relativa do ar para estação de Itaperuna com o modelo MLP-AG

Série de dados utilizados da estação de Itaperuna-(MLP-AG)						
DATA	Umid. Relat. do Ar	Umid. Relat. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2008	76,383333	77,21861948	0,8352865	0,697703512	0,01093545	0,96
31/10/2008	70,54032301	72,41337758	1,8730546	3,508333416	0,02655296	
31/08/2009	70,443548	71,41835274	0,9748047	0,950244286	0,0138381	CONCORDÂNCIA (D)
30/09/2009	69,07499999	69,57612272	0,5011227	0,251123987	0,00725476	0,98
30/11/2009	68,791667	69,86955576	1,0778888	1,161844188	0,01566889	
28/02/2010	61,88392901	61,85640673	0,0275223	0,000757476	0,00044474	CONFIANÇA (C)
30/04/2010	74,3	73,28926419	1,0107358	1,021586893	0,01360344	0,94
31/05/2010	75,80645199	77,90476743	2,0983154	4,402927684	0,02767991	
30/06/2010	68,99166699	73,40840242	4,4167354	19,50755182	0,06401839	DESEMPENHO
31/08/2010	63,733871	63,09853308	0,6353379	0,403654274	0,00996861	Ótimo
30/09/2010	62,33333299	64,20166483	1,8683318	3,490663867	0,02997324	
30/11/2010	79,85	77,04873719	2,8012628	7,847073302	0,03508156	
31/12/2010	75,29032301	74,42107736	0,8692456	0,755587993	0,01154525	
31/03/2011	82,40322601	80,97588622	1,4273398	2,037298872	0,01732141	
30/04/2011	77,99166701	76,57439054	1,4172765	2,00867259	0,01817215	
31/05/2011	75,90322601	75,97869939	0,0754734	0,005696232	0,00099434	
			EMA	REMQ	EMP	
			1,37	1,73	1,89	

APÊNDICE T – Dados estimados de umidade relativa do ar para estação do Rio de Janeiro com o modelo MLP-AG

Série de dados utilizados da estação de Rio de Janeiro - (MLP-AG)						
DATA	Umid. Relat. do Ar	Umid. Relat. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2008	74,3500000	75,41843307	1,0684331	1,1415492	0,01437032	0,90
31/10/2008	70,7500000	72,27412147	1,5241215	2,3229462	0,02154235	
31/08/2009	73,2419350	72,37507814	0,8668569	0,7514408	0,01183553	CONCORDÂNCIA (D)
30/09/2009	71,7750000	70,46015818	1,3148418	1,7288090	0,01831894	0,93
30/11/2009	67,8083330	70,08075652	2,2724235	5,1639086	0,03351245	
28/02/2010	62,7500000	66,48539659	3,7353966	13,9531877	0,05952823	CONFIANÇA (C)
30/04/2010	74,5666670	74,00829644	0,5583706	0,3117777	0,00748821	0,83
31/05/2010	73,9274190	76,08279306	2,1553741	4,6456374	0,02915527	
30/06/2010	72,7833330	72,98570312	0,2023701	0,0409537	0,00278045	DESEMPENHO
31/08/2010	68,0967740	68,43451155	0,3377375	0,1140666	0,00495967	Muito Bom
30/09/2010	70,3083330	68,81716988	1,4911631	2,2235675	0,02120891	
30/11/2010	74,5916670	73,61912478	0,9725422	0,9458384	0,01303822	
31/12/2010	75,7741940	76,10233306	0,3281391	0,1076752	0,00433049	
31/03/2011	78,9838710	74,83874642	4,1451246	17,1820577	0,05248065	
30/04/2011	75,0666670	75,6691964	0,6025294	0,3630417	0,00802659	
31/05/2011	77,2661290	77,14935137	0,1167776	0,0136370	0,00151137	
			EMA	REMQ	EMP	
			1,36	1,79	1,90	

APÊNDICE U – Dados estimados de umidade relativa do ar para estação de Paty do Alferes com o modelo MLP-AG

Série de dados utilizados da estação de Paty do Alferes-(MLP-AG)						
DATA	Umid. Relat. do Ar	Umid. Relat. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2008	83,0000000	83,0023218	0,0023218	0,0000054	2,7974E-05	0,94
31/10/2008	76,2822580	78,2753596	1,9931016	3,9724540	0,02612798	
31/08/2009	76,8064520	79,30803233	2,5015803	6,2579041	0,03256992	CONCORDÂNCIA (D)
30/09/2009	75,6833330	77,13466671	1,4513337	2,1063695	0,0191764	0,97
30/11/2009	76,4666670	75,85570801	0,610959	0,3732709	0,00798987	
28/02/2010	74,6696430	72,7447274	1,9249156	3,7053001	0,02577909	CONFIANÇA (C)
30/04/2010	81,8083330	80,53730175	1,2710312	1,6155204	0,0155367	0,92
31/05/2010	82,6370970	83,41711921	0,7800222	0,6084347	0,00943913	
30/06/2010	81,8750000	80,3709507	1,5040493	2,2621643	0,01837007	DESEMPENHO
31/08/2010	74,1290320	73,49006651	0,6389655	0,4082769	0,00861964	Ótimo
30/09/2010	74,7916670	74,74958157	0,0420854	0,0017712	0,0005627	
30/11/2010	82,0000000	82,36284245	0,3628425	0,1316547	0,00442491	
31/12/2010	81,2258060	81,45763351	0,2318275	0,0537440	0,00285411	
31/03/2011	83,6209680	83,24860776	0,3723602	0,1386522	0,00445295	
30/04/2011	82,3583330	82,77763987	0,4193069	0,1758182	0,00509125	
31/05/2011	81,9677420	82,20081221	0,2330702	0,0543217	0,00284344	
			EMA	REMQ	EMP	
			0,90	1,17	1,15	

APÊNDICE V – Dados estimados de umidade relativa do ar para estação de Resende com o modelo MLP-AG

Série de dados utilizados da estação de Resende - (MLP-AG)						
DATA	Umid. Relat. do Ar	Umid. Relat. do Ar Pred.	ERRO IND.	ERRO MD-QUADRA.	ERRO PERC.	CORRELAÇÃO (r)
30/06/2008	84,0750000	79,57932327	4,4956767	20,2111094	0,05347222	0,90
31/10/2008	78,7741940	80,32220661	1,5480126	2,3963430	0,01965127	
31/08/2009	75,5161290	75,71376822	0,1976392	0,0390613	0,00261718	CONCORDÂNCIA (D)
30/09/2009	77,9833330	77,47875657	0,5045764	0,2545974	0,00647031	0,93
30/11/2009	76,6750000	74,4918532	2,1831468	4,7661300	0,02847273	
28/02/2010	73,6607140	70,77487482	2,8858392	8,3280678	0,03917745	CONFIANÇA (C)
30/04/2010	76,9000000	77,47363324	0,5736332	0,3290551	0,00745947	0,84
31/05/2010	79,5967740	79,75863993	0,1618659	0,0262006	0,00203357	
30/06/2010	77,4583330	77,47107157	0,0127386	0,0001623	0,00016446	DESEMPENHO
31/08/2010	68,1693550	68,56671812	0,3973631	0,1578975	0,00582906	Muito Bom
30/09/2010	69,3833330	69,27117646	0,1121565	0,0125791	0,00161648	
30/11/2010	79,9333330	79,4999116	0,4334214	0,1878541	0,00542229	
31/12/2010	81,2016130	77,27894657	3,9226664	15,3873120	0,04830774	
31/03/2011	84,0806450	80,31964494	3,7610001	14,1451215	0,04473087	
30/04/2011	78,1083330	79,4230616	1,3147286	1,7285113	0,01683212	
31/05/2011	79,1290320	77,73492324	1,3941088	1,9435393	0,01761817	
			EMA	REMQU	EMP	
			1,49	2,09	1,87	