



Universidade Federal da Paraíba

Centro de Tecnologia

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL E

AMBIENTAL

– MESTRADO –

**PREVISÃO DOS NÍVEIS DE ÁGUA SUBTERRÂNEA EM POÇOS DO
AQUÍFERO OGALLALA UTILIZANDO REDES NEURAIAS
ARTIFICIAIS**

Por

Lairanne Costa de Oliveira

*Dissertação de Mestrado apresentada à Universidade Federal da Paraíba
para obtenção do grau de Mestre*

João Pessoa – Paraíba

Setembro de 2020



Universidade Federal da Paraíba

Centro de Tecnologia

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL E

AMBIENTAL

– MESTRADO –

**PREVISÃO DOS NÍVEIS DE ÁGUA SUBTERRÂNEA EM POÇOS DO
AQUÍFERO OGALLALA UTILIZANDO REDES NEURAIAS
ARTIFICIAIS**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Civil e Ambiental da Universidade Federal da Paraíba, como parte dos requisitos para a obtenção do título de Mestre.

Lairanne Costa de Oliveira

Orientador: Prof. Dr. Celso Augusto Guimarães Santos

João Pessoa – Paraíba

Setembro de 2020

Catálogo na publicação
Seção de Catalogação e Classificação

048p Oliveira, Lairanne Costa de.
Previsão dos níveis de água subterrânea em poços do
aquífero Ogallala utilizando redes neurais artificiais
/ Lairanne Costa de Oliveira. - João Pessoa, 2020.
142 f. : il.

Orientação: Celso Augusto Guimarães Santos.
Dissertação (Mestrado) - UFPB/CT.

1. Engenharia urbana e ambiental. 2. Redes neurais
artificiais. 3. Aquífero Ogallala. 4. Águas
subterrâneas. I. Santos, Celso Augusto Guimarães. II.
Título.

UFPB/BC

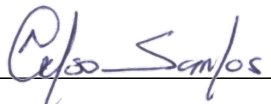
CDU 62:711 (043)

LAIRANNE COSTA DE OLIVEIRA

**PREVISÃO DOS NÍVEIS DE ÁGUA SUBTERRÂNEA EM POÇOS DO
AQUÍFERO OGALLALA UTILIZANDO REDES NEURAIAS
ARTIFICIAIS**

Dissertação aprovada em 25/09/20 como requisito para a obtenção do título de Mestre em Engenharia Civil e Ambiental do Centro de Tecnologia da Universidade Federal da Paraíba.

BANCA EXAMINADORA:



Prof. Dr. Celso Augusto Guimarães Santos – UFPB
(Orientador)



Prof. Dr. Richarde Marques da Silva – UFPB
(Examinador Interno)



Prof. Dr. Camilo Allyson Simões de Farias – UFCG
(Examinador Externo)

AGRADECIMENTOS

Agradeço a Deus por todo o seu cuidado e provisão em minha vida, pela saúde e disposição que me permitiram a realização deste trabalho.

Aos meus pais, Dione Meneses e Edmo José, por todo o amor, carinho e suporte recebido em minha formação pessoal e profissional.

Ao meu irmão, Lucas Costa, o primeiro amigo que tive e com quem compartilho as melhores recordações da nossa infância.

A João Vitor, por sempre querer o melhor para mim, pelo amor e apoio incondicional em todos os momentos.

Ao professor Celso, por toda sua dedicação, confiança, orientação e conhecimentos passados na elaboração desta dissertação.

Aos professores Richarde Marques e Camilo Allyson por terem aceitado fazer parte da banca de avaliação deste trabalho.

Ao grupo de orientandos do professor Celso, onde conheci pessoas maravilhosas que me inspiraram nessa jornada.

Aos meus queridos amigos Camila, Moana e Reginaldo, por todos os momentos alegres e de apoio mútuo que compartilhamos durante o mestrado.

Às minhas amigas Marília Michele e Cecília Flávia pela amizade sincera e por serem minha fonte diária de alegria, amor e boas risadas.

A todos os professores que fizeram parte da minha caminhada estudantil, contribuindo com o meu crescimento pessoal e intelectual.

Agradeço também a todos que de alguma forma contribuíram para a realização deste trabalho.

RESUMO

O Aquífero Ogallala, situado nos Estados Unidos da América, é o maior da América do Norte e serve como principal fonte de abastecimento de água agrícola e pública. Entretanto, devido a décadas de bombeamento para irrigação combinadas com recarga limitada, o Aquífero Ogallala vem enfrentando o declínio contínuo nos níveis de água subterrânea. A previsão dos níveis de água subterrânea surge como uma alternativa capaz de viabilizar um planejamento mais eficaz em relação ao uso dos recursos hídricos. Nesse sentido, as RNAs tem sido utilizadas por diversos autores para modelagem de águas subterrâneas. Diante do exposto, o presente estudo visa prever os níveis de água, em poços localizados na porção do Aquífero Ogallala contida no Texas, através da utilização de redes neurais artificiais. Para isso, foram obtidos dados de águas subterrâneas a partir da plataforma do *Texas Water Development Board* (TWDB) e, após um processo de seleção, foram escolhidos 403 poços para realizar a previsão dos níveis da água. A análise hierárquica de *clusters* foi realizada nas séries temporais dos 403 poços estudados, de modo a classificá-los em grupos de acordo com suas semelhanças nas flutuações dos níveis de água. A partir da análise de *clusters*, foram selecionados trinta grupos. Neste estudo, foram desenvolvidos três diferentes modelos de RNAs para realizar a previsão dos níveis de água subterrânea nos poços. O primeiro modelo consistiu em desenvolver 403 RNAs, uma para cada poço analisado, para realizar a previsão do nível d'água para um ou até dez anos à frente. O segundo modelo definiu uma RNA para cada um dos *clusters* selecionados, para prever os níveis d'água, nos poços pertencentes a cada *cluster*, para um ano à frente. O último modelo proposto estabeleceu uma única RNA, para todos os poços estudados no aquífero, para previsão dos níveis d'água para um ano à frente. O desempenho, de todos os modelos de RNA desenvolvidos nesse trabalho, foi avaliado a partir dos seguintes índices estatísticos: coeficiente de correlação linear de Pearson (R), a eficiência de Nash-Sutcliffe (NSE), erro de porcentagem absoluta média (MAPE), raiz do erro quadrático médio (RMSE) e o Bias. Ao analisar os resultados é possível concluir que os resultados das previsões, de todos os modelos propostos, mostram-se, no geral, satisfatórios, obtendo coeficiente de correlação e NSE próximos a 1, e MAPE e RMSE próximos a zero. O primeiro modelo foi o que apresentou, de modo geral, o melhor desempenho. Já o modelo que definiu uma única RNA, para todos os poços estudados, foi o que teve o menor desempenho, na maior parte dos casos, quando comparado aos demais.

PALAVRAS-CHAVE: Redes neurais artificiais; Aquífero Ogallala; Águas subterrâneas.

ABSTRACT

The Ogallala Aquifer, located in the United States of America, is the largest in North America and serves as the main source of agricultural and public water supply. However, due to decades of pumping for irrigation combined with limited recharge, the Ogallala Aquifer has been facing a decline in its groundwater levels (GWLs). The prediction of GWL appears to be an alternative to enable more efficient planning to better use water resources. In this regard, several authors have been using ANNs for modeling groundwater. Therefore, this study aims to use artificial neural networks to predict the GWLs of several wells in the Ogallala Aquifer located in Texas. For this purpose, groundwater data were obtained from the Texas Water Development Board (TWDB) platform, and 403 wells were chosen to perform the GWLs prediction after a selection process. Also, hierarchical cluster analysis was performed in the time series of the 403 wells studied to classify them in groups according to their similarities in the water level fluctuations. In this study, thirty groups were selected from cluster analysis, and it was developed three ANN models to forecast the GWLs of the wells. The first model consisted of developing 403 ANNs, one for each analyzed well, to predict the water level from one to ten years ahead. For the second model, it was determined an ANN for each of the selected clusters to forecast the GWLs of the wells belonging to each cluster for one year ahead. For the last model proposed, it is used a single ANN, to predict the GWLs for a year ahead, to all studied wells. The performance of the ANN models developed in this work was evaluated using the following statistical indices: Pearson's linear correlation coefficient (R), Nash-Sutcliffe efficiency (NSE), mean absolute percentage error (MAPE), the root of the mean square error (RMSE) and the Bias. By analyzing the results, it is possible to conclude that all the predictions for the proposed models are satisfactory, since it obtained a correlation coefficient and NSE close to 1, and MAPE and RMSE close to zero. Additionally, the first model presented the best performance, and the model that defined a single ANN had the lowest performance when compared to the others.

KEYWORDS: Artificial neural networks; Ogallala Aquifer; Groundwater.

SUMÁRIO

RESUMO

ABSTRACT

LISTA DE FIGURAS

LISTA DE TABELAS

1	INTRODUÇÃO	15
2	OBJETIVOS	16
2.1	OBJETIVO GERAL	16
2.2	OBJETIVOS ESPECÍFICOS.....	16
3	REFERENCIAL TEÓRICO	17
3.1	ÁGUAS SUBTERRÂNEAS: RELEVÂNCIA E UTILIZAÇÃO.....	17
3.2	A PROBLEMÁTICA DO USO IRRACIONAL DOS RECURSOS HÍDRICOS SUBTERRÂNEOS	17
3.3	A MODELAGEM NA HIDROLOGIA	18
3.4	A MODELAGEM HIDROLÓGICA NA GESTÃO DOS RECURSOS HÍDRICOS SUBTERRÂNEOS	19
3.5	REDES NEURAIS ARTIFICIAIS	21
3.5.1	Aspectos gerais	21
3.5.2	O neurônio biológico	22
3.5.3	O neurônio artificial.....	23
3.5.4	Caracterização das RNAs	24
3.5.4.1	Funções de ativação	24
3.5.4.2	Arquiteturas das RNAs.....	25
3.5.4.3	Aprendizagem da RNA	27
3.6	REDES NEURAIS ARTIFICIAIS PARA MODELAGEM DOS NÍVEIS DE ÁGUAS SUBTERRÂNEAS	28
3.7	ANÁLISE DE <i>CLUSTER</i>	29
3.7.1	Métodos por particionamento	29
3.7.2	Métodos hierárquicos	29
4	CARACTERÍSTICAS DA ÁREA DE ESTUDO	32
4.1	LOCALIZAÇÃO E CARACTERÍSTICAS GERAIS	32
4.2	CARACTERÍSTICAS HIDROGEOLÓGICAS	33
4.3	IMPORTÂNCIA E PROBLEMAS ENFRENTADOS NO AQUIFERO OGALLALA.....	34
5	MATERIAL E MÉTODOS	36
5.1	OBTENÇÃO E PRÉ-PROCESSAMENTO DOS DADOS	37

5.1.1	Preenchimento de falhas	40
5.2	ARQUITETURA DAS RNAs	43
5.2.1	Cr�terios de avalia��o das RNAs	47
5.3	DESENVOLVIMENTO DE RNAs PARA CADA PO�O EM ESPEC�FICO	48
5.4	APLICA��O DAS RNAs TREINADAS COM OS DADOS ESPEC�FICOS DE UM PO�O NOS OUTROS PO�OS	48
5.5	AN�LISE DE <i>CLUSTERS</i>	48
5.5.1	Escolha da m�trica e do m�todo	49
5.5.2	Escolha do n�mero de <i>clusters</i>	54
5.6	DESENVOLVIMENTO DE UMA RNA POR <i>CLUSTER</i> , TREINADA COM OS DADOS DE TODOS OS PO�OS DO <i>CLUSTER</i>	54
5.7	DESENVOLVIMENTO DE UMA RNA POR <i>CLUSTER</i> , TREINADA COM OS DADOS DE ALGUNS PO�OS DO <i>CLUSTER</i>	55
5.8	DESENVOLVIMENTO DE UMA �NICA RNA PARA TODOS OS PO�OS	55
6	RESULTADOS E DISCUSS�ES	56
6.1	RESULTADOS DAS PREVIS�ES DOS N�VEIS D'�GUA EM CADA PO�O UTILIZANDO RNAs TREINADAS COM OS DADOS ESPEC�FICOS DO PO�O ANALISADO	56
6.1.2	Resultados das previs�es dos n�veis d'�gua em cada po�o utilizando doze anos precedentes na camada de entrada.....	68
6.2	RESULTADOS DA APLICA��O DAS RNAs DESENVOLVIDAS DE MODO ESPEC�FICO PARA CADA PO�O PARA REALIZAR A PREVIS�O DOS N�VEIS D'�GUA EM OUTROS PO�OS ...	69
6.3	RESULTADOS E DISCUSS�ES DA AN�LISE DE <i>CLUSTERS</i>	72
6.3.1	Resultados das previs�es dos n�veis d'�gua em cada po�o utilizando uma RNA por <i>cluster</i> , treinada com os dados de todos os po�os do <i>cluster</i>	81
6.3.2	Compara��o entre os resultados obtidos pelas RNAs feitas de modo espec�fico para cada po�o com os das RNAs feitas por <i>cluster</i> , treinadas com os dados de todos os po�os do <i>cluster</i>	92
6.3.3	Resultados das previs�es dos n�veis d'�gua em cada po�o utilizando uma RNA por <i>cluster</i> , treinada com os dados de alguns po�os do <i>cluster</i>	95
6.4	RESULTADOS DAS PREVIS�ES DOS N�VEIS D'�GUA UTILIZANDO UMA �NICA RNA PARA TODOS OS PO�OS.....	116
7	CONCLUS�ES E RECOMENDA��ES	119
	REFER�NCIAS BIBLIOGR�FICAS	123
	AP�NDICE A	130
	AP�NDICE B.....	135

LISTA DE FIGURAS

Figura 1 – Neurônio biológico simplificado (Fonte: Adaptado de PATTERSON & GIBSON, 2017).....	22
Figura 2 – Modelo de um neurônio artificial (Fonte: HAYKIN, 2007).....	23
Figura 3 – Funções de ativação (Fonte: Adaptado de COPPIN, 2010).....	25
Figura 4 – Rede neural do tipo MLP com uma camada intermediária (Adaptado de SANTOS & SILVA, 2014).....	26
Figura 5 – Exemplo de dendrograma.	30
Figura 6 – Mapa da área de cobertura do Aquífero Ogallala (DE BRITO NETO, 2012)...	32
Figura 7 – Área de estudo: Aquífero Ogallala compreendido no estado do Texas (DE BRITO NETO et al., 2016).	33
Figura 8 – Fluxograma da metodologia aplicada.	36
Figura 9 – Distribuição espacial dos 403 poços selecionados para a pesquisa.	38
Figura 10 – Distribuição de falhas (células na cor preta) nas séries temporais do nível d'água de cada poço.	39
Figura 11 – Quantidade de falhas existentes nas séries históricas de cada um dos 403 poços.	40
Figura 12 – Representação gráfica do método da interpolação linear (Adaptado de JIANG et al., 2018).....	42
Figura 13 – Diferentes metodologias para o preenchimento das falhas do poço 329.	43
Figura 14 – Distribuição espacial dos 10 poços sem falhas.	44
Figura 15 – Esquema explicativo dos testes realizados para a determinação do número de neurônios na camada oculta das RNAs.	45
Figura 16 – Resumo dos testes realizados para a decisão da quantidade de anos precedentes a serem utilizados para a previsão com a RNA.	46
Figura 17 – Arquitetura final da RNA.....	46
Figura 18 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de um ano à frente.	57
Figura 19 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de dois anos à frente.	57
Figura 20 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de três anos à frente.	58
Figura 21 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de quatro anos à frente.....	58
Figura 22 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de cinco anos à frente.	59
Figura 23 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de seis anos à frente.	59
Figura 24 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de sete anos à frente.....	60
Figura 25 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de oito anos à frente.....	60
Figura 26 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de nove anos à frente.	61
Figura 27 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de dez anos à frente.	61
Figura 28 – Variação espacial e temporal do coeficiente de correlação (R), utilizando quatro anos precedentes na camada de entrada das RNAs.	64

Figura 29 – Amostra dos resultados de previsão para um ano à frente.	66
Figura 30 – Gráficos de dispersão para os poços de 1 a 6, para previsão de um ano à frente.	67
Figura 31 – Variação espacial e temporal do coeficiente de correlação (R), utilizando doze anos precedentes na camada de entrada das RNAs.	69
Figura 32 – Variação espacial do NSE para as RNAs do poço 1 ao 12 aplicadas nos demais poços.	71
Figura 33 – Dendrograma resultante da análise hierárquica de <i>clusters</i> das séries temporais de 403 poços do Aquífero Ogallala.	72
Figura 34 – Comportamento das séries temporais dos níveis d'água dos poços pertencentes a cada <i>cluster</i>	74
Figura 35 – Flutuações dos níveis das águas subterrâneas nos poços de diferentes <i>clusters</i>	75
Figura 36 – Distribuição espacial dos poços, no Aquífero Ogallala, divididos por <i>clusters</i> resultantes da análise hierárquica de suas séries temporais.	78
Figura 37 – Distribuição espacial dos poços, no Aquífero Ogallala, divididos por <i>clusters</i> resultantes da análise hierárquica de suas séries temporais.	79
Figura 38 – Arranjo que mostra a sobreposição de alguns dos <i>clusters</i> que foram considerados similares pela análise do dendrograma da Figura 33.	80
Figura 39 – Comparação dos resultados de correlação obtidos pelas RNAs individuais de cada poço e pelas RNAs dos <i>clusters</i>	93
Figura 40 – Comparação dos resultados de NSE obtidos pelas RNAs individuais de cada poço e pelas RNAs dos <i>clusters</i>	93
Figura 41 – Comparação dos resultados de MAPE obtidos pelas RNAs individuais de cada poço e pelas RNAs dos <i>clusters</i>	94
Figura 42 – Comparação dos resultados de RMSE obtidos pelas RNAs individuais de cada poço e pelas RNAs dos <i>clusters</i>	94
Figura 43 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 1.	98
Figura 44 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 2.	99
Figura 45 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 3.	100
Figura 46 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 4.	100
Figura 47 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 5.	101
Figura 48 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 6.	102
Figura 49 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 7.	103
Figura 50 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 8.	103
Figura 51 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 9.	104
Figura 52 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 10.	105
Figura 53 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 11.	105

Figura 54 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 12.....	106
Figura 55 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 13.....	107
Figura 56 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 14.....	107
Figura 57 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 15.....	108
Figura 58 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 16.....	109
Figura 59 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 17.....	109
Figura 60 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 18.....	110
Figura 61 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 19.....	111
Figura 62 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 20.....	111
Figura 63 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 21.....	112
Figura 64 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 22.....	113
Figura 65 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 23.....	113
Figura 66 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 24.....	114
Figura 67 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 26.....	115
Figura 68 – Comparação dos resultados obtidos pelas RNAs dos <i>clusters</i> , treinadas com todos os poços e treinadas com 80% dos poços, para o <i>cluster</i> 27.....	115
Figura 69 – Resultados dos índices estatísticos utilizando a RNA treinada com os dados de todos os poços.....	117
Figura 70 – Comparação dos resultados obtidos pela RNA desenvolvida para todos os poços e pelas RNAs específicas de cada poço.....	118
Figura 71 – Comparação dos resultados obtidos pela RNA desenvolvida para todos os poços e pelas RNAs feitas para cada <i>cluster</i> , treinadas com os dados de todos os poços do <i>cluster</i>	118

LISTA DE TABELAS

Tabela 1 – Resultado do cálculo do coeficiente de correlação cofenética para as combinações de método e métrica testadas	51
Tabela 2 – Resultados do cálculo da variação máxima de nível d'água entre os poços (em metros) de todos os <i>clusters</i> resultantes da combinação de cada método e métrica. ..	52
Tabela 3 – Índices estatísticos da RNA escolhida para cada poço, para previsão de um e dois anos.....	56
Tabela 4 – Quantidade de poços com valores de NSE abaixo de 0,7 por ano de previsão .	63
Tabela 5 – Quantidade e poços que se encontram em cada <i>cluster</i>	73
Tabela 6 – Resultados dos índices estatísticos das RNAs selecionadas para cada <i>cluster</i> ..	82
Tabela 7 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 1 e 2 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	83
Tabela 8 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 3 e 4 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	84
Tabela 9 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 5 e 6 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	85
Tabela 10 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 7 e 8 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	86
Tabela 11 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 9 e 10 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	87
Tabela 12 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 11 e 12 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	87
Tabela 13 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 13 e 14 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	88
Tabela 14 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 15 e 16 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	89
Tabela 15 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 17 e 18 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	89
Tabela 16 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 19 e 20 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	90
Tabela 17 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 21 e 22 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	90
Tabela 18 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 23 e 24 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	91
Tabela 19 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 25 e 26 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	91
Tabela 20 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 27 e 28 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	92
Tabela 21 – Resultados dos índices estatísticos das RNAs feitas para os <i>clusters</i> 29 e 30 aplicadas aos dados da série temporal de cada poço do <i>cluster</i>	92
Tabela 22 – Poços que foram retirados, por <i>cluster</i> , antes da realização da aprendizagem da RNA de cada <i>cluster</i>	96
Tabela 23 – Resultados dos índices estatísticos das RNAs selecionadas para cada <i>cluster</i>	97
Tabela 24 – Variação máxima do nível d'água de cada poço a partir do nível d'água mínimo e máximo	130
Tabela 25 – Poços que obtiveram resultados satisfatórios por meio do uso da RNA feita especificamente para outro poço	135

LISTA DE ABREVIATURAS E SIGLAS

BP	<i>Back propagation</i>
BR	<i>Bayesian regularization</i>
EUA	Estados Unidos da América
HCA	<i>Hierarchical Cluster Analysis</i>
IA	Inteligência Artificial
LM	<i>Levenberg-Marquardt</i>
MLP	<i>Multilayer Perceptron</i>
MAPE	<i>Mean absolute percentage error</i>
NA	Nível da água
NSE	<i>Nash–Sutcliffe efficiency</i>
PPGECAM	Programa de Pós-Graduação em Engenharia Civil e Ambiental
RMSE	<i>Root Mean Square Error</i>
RNA	Rede Neural Artificial
TANH	Tangente Hiperbólica
THP	<i>Texas High Plains</i>
TWDB	<i>Texas Water Development Board</i>
UFPB	Universidade Federal da Paraíba

1 INTRODUÇÃO

Em muitos países, a água subterrânea é tida como uma das principais fontes de abastecimento para os mais variados setores. Tal posição de destaque fez com que muitos aquíferos sofressem impactos ambientais negativos em função do seu uso indiscriminado, entre tais impactos pode-se citar a deterioração da qualidade da água, a depleção do nível das águas e a consequente diminuição do rendimento de poços (KONIKOW & KENDY, 2005).

O Aquífero Ogallala, situado nas planícies centrais dos Estados Unidos da América (EUA), serve como principal fonte de abastecimento de água agrícola e pública, ele tem sustentado o desenvolvimento econômico dos EUA por mais de 80 anos (CANO et al., 2018). Por estar localizado em uma região que possui clima semiárido e águas superficiais limitadas, há uma grande pressão sobre os seus recursos hídricos (STEWART; ALLEN, 2016) (MOGHADDAM et al., 2019).

Desse modo, o Aquífero Ogallala vem sofrendo com o declínio contínuo nos níveis de água subterrânea devido a décadas de bombeamento para irrigação combinadas com recarga limitada do aquífero, resultando, assim, na diminuição das capacidades dos poços (CHEN et al., 2018). Dessa forma, torna-se necessário um manejo sustentável dos recursos hídricos de maneira a garantir o abastecimento de água ao longo dos anos (MOHANTY et al., 2010).

A previsão dos níveis de água subterrânea surge como uma alternativa capaz de viabilizar um planejamento mais eficaz em relação ao uso dos recursos hídricos. Tal previsão permite a formulação de estratégias e ações que possibilitam, dentre outras práticas: A minimização dos impactos gerados pelo mau uso dos aquíferos, a exemplo da subsidência do solo e compactação do reservatório; trazer novas soluções para conflitos referentes ao uso da água; e também garantir um maior conhecimento sobre os fatores que influenciam o nível do lençol freático de determinado local (ADAMOWSKI & CHAN, 2011).

Nesse sentido, os modelos hidrológicos vêm sendo amplamente usados para a simulação do nível de água subterrânea em vários estudos (BARZEGAR et al., 2017). Os modelos de base física são mais comumente utilizados para esse tipo de análise, porém, tais modelos dependem de uma grande quantidade de informações e dados hidrológicos, que muitas vezes não estão disponíveis para a bacia hidrográfica em questão, ocasionando, assim, restrições ao seu uso. Dessa forma, modelos empíricos, a exemplo das redes neurais artificiais (RNAs), mostram-se como uma opção, uma vez que demandam uma menor

quantidade de dados e um menor tempo de computação, além de serem mais baratos (MOHANTY et al., 2010).

Segundo Maheswaran e Khosa (2013), entre a ampla variedade de métodos empíricos, as RNAs são as mais utilizadas para modelagem de águas subterrâneas, tendo sido aplicadas por vários autores como Daliakopoulos et al. (2005), Nayak et al. (2006), Ghadampour e Rakhshandehroo (2010), Mohanty et al. (2013), Moghaddam et al. (2019), entre outros.

Redes neurais artificiais são modelos empíricos do tipo “caixa preta” que procuram imitar o comportamento do cérebro humano e são capazes de aprender relações entre variáveis aplicando algoritmos de treinamento especiais nos dados analisados (WUNSCH; LIESCH; BRODA, 2018). RNAs são apropriadas para modelagem dinâmica de sistemas não-lineares, como é o caso, em geral, de sistemas hidrológicos, já que esses possuem características que variam no tempo e no espaço (WUNSCH; LIESCH; BRODA, 2018). Diante do exposto, o presente estudo visa prever os níveis de água em poços localizados no Aquífero Ogallala, através da utilização de redes neurais artificiais.

2 OBJETIVOS

2.1 Objetivo geral

Realizar a previsão dos níveis de água subterrânea em poços localizados na porção do Aquífero Ogallala contida no Texas, através da utilização de redes neurais artificiais.

2.2 Objetivos específicos

- (a) Definir uma RNA para cada poço estudado;
- (b) Estabelecer uma RNA para cada *cluster* definido;
- (c) Avaliar a capacidade de previsão das RNAs por *clusters*, treinadas sem parte dos poços do *cluster*; e
- (d) Definir uma RNA única para prever o nível d'água de todos os poços estudados no Aquífero Ogallala.

3 REFERENCIAL TEÓRICO

3.1 Águas subterrâneas: relevância e utilização

O crescimento e a manutenção da civilização humana dependem muito da disponibilidade de água (MUKHERJEE; RAMACHANDRAN, 2018). A água subterrânea é um recurso vital e constitui uma importante fonte de água doce em todo o mundo. Geralmente é mais limpa, confiável e mais facilmente disponível do que outros recursos de água potável (WUNSCH; LIESCH; BRODA, 2018).

Os aquíferos subterrâneos são o maior recurso de água doce do mundo, sendo que a maior parte da água doce líquida em nosso planeta (mais de 98%) ocorre como água subterrânea, enquanto menos de 2% ocorre na forma de água superficial, como em córregos e lagos que são frequentemente alimentados por água subterrânea (MOHANTY et al., 2013).

Durante as últimas décadas, as águas subterrâneas têm ganhado cada vez mais relevância para o suprimento das necessidades humanas, fornecendo água potável a mais de 1,5 bilhão de pessoas (MELAKU; WANG, 2019). Em regiões áridas e semiáridas, a água subterrânea torna-se ainda mais importante por ser um recurso flexível em períodos de seca e escassez de água (JEIHOUNI et al., 2019)

Os aquíferos são fonte de 1/3 de todas as captações e suprimentos de água doce. Estima-se que a água subterrânea forneça cerca de 36% do uso de água doméstico global atual, 27% do suprimento industrial e 42% da demanda de água na agricultura irrigada (MUKHERJEE; RAMACHANDRAN, 2018).

3.2 A problemática do uso irracional dos recursos hídricos subterrâneos

O crescimento rápido da população e o aumento da demanda de água tem levado à exploração excessiva dos recursos hídricos subterrâneos (MOGHADDAM et al., 2019). Vários efeitos do uso e gerenciamento inadequados das águas subterrâneas estão se tornando um problema sério em todo o mundo (ADAMOWSKI & CHAN, 2011).

Em muitas regiões, as águas subterrâneas foram retiradas a taxas muito acima da recarga, o que levou a impactos ambientais negativos, a exemplo de grandes quedas no nível da água de subterrânea o que implica na diminuição do volume de água armazenado pelos aquíferos (ADAMOWSKI & CHAN, 2011).

Os efeitos da depleção das águas subterrâneas são complexos e dependentes do aquífero. A redução do nível dos lençóis freáticos leva ao aumento dos custos com o

bombeamento dos poços, prejudicando os usuários; redução da descarga de águas subterrâneas em córregos, nascentes e zonas úmidas, afetando os ecossistemas; e subsidência do solo, que pode danificar estruturas (AESCHBACH-HERTIG; GLEESON, 2012).

A diminuição do nível dos lençóis freáticos também induz o fluxo de água subterrânea, o que pode levar à salinização por intrusão de água salgada nas regiões costeiras ou por vazamentos de camadas adjacentes que contêm água salina. Da mesma forma, esse fluxo induzido também pode promover a disseminação de outros tipos de poluição, contribuindo para a degradação da qualidade da água subterrânea (AESCHBACH-HERTIG; GLEESON, 2012).

Além disso, devido aos impactos provocados pelas mudanças climáticas no ciclo hidrológico, os recursos subterrâneos encontram-se sob maior estresse (JEIHOUNI et al., 2019). Portanto, espera-se uma rápida diminuição no nível das águas subterrâneas, principalmente em regiões áridas e semiáridas (MOGHADDAM et al., 2019).

Como consequência, muitas bacias hidrográficas estão enfrentando graves problemas ambientais, sociais e econômicos (LI et al., 2019). Assim, o gerenciamento adequado dos recursos hídricos subterrâneos em conjunto com as águas superficiais de uma bacia hidrográfica é muito importante para garantir sua sustentabilidade (MACKAY et al., 2015).

3.3 A modelagem na hidrologia

A modelagem pode ser considerada como um procedimento teórico que envolve técnicas elaboradas com o intuito de representar de modo simplificado o mundo (CHRISTOFOLETTI, 1999). Modelos são essencialmente representações mais simples de sistemas, uma vez que não é possível dar conta de toda a extensão da complexidade do mundo real (RAMESH et al., 2020).

Nesse sentido, a modelagem de sistemas ambientais tem como objetivo retratar os fenômenos da natureza e favorecer a compreensão desses (CHRISTOFOLETTI, 1999). O campo hidrológico foi amplamente favorecido com o surgimento da modelagem (TUCCI, 2005).

A hidrologia é a ciência que trata da água na Terra, sua ocorrência, circulação, distribuição espacial, suas propriedades físicas e químicas e sua relação com o ambiente, inclusive os seres vivos (COLLISCHONN & DORNELLES, 2015). Os hidrologistas precisam, constantemente, resolver problemas de previsão e estimativa de precipitação,

níveis de água, escoamento superficial, concentração de poluentes, entre outros (ASCE, 2000b).

Os processos hidrológicos possuem alta complexidade e dinamismo, e, portanto, as previsões hidrológicas sempre representaram um desafio para os hidrologistas (NOURANI et al., 2014). Desse modo, os modelos que visam simular tais processos hidrológicos também são complexos, porém, extremamente necessários na gestão ambiental e dos recursos hídricos (TUCCI, 2005).

Modelos hidrológicos são instrumentos que foram desenvolvidos com o propósito de compreender e retratar o comportamento de bacias hidrográficas. Os modelos podem ser classificados como (TUCCI, 2005):

i) Concentrado ou distribuído: Em um modelo concentrado toda a bacia e seus processos hidrológicos são modelados por variáveis concentradas no espaço. Já o modelo é considerado distribuído quando suas variáveis e parâmetros dependem do espaço e/ou tempo;

ii) Estocástico ou Determinístico: quando a chance de ocorrência das variáveis for levada em consideração, e o conceito de probabilidade é introduzido na formulação do modelo, o processo e o modelo são ditos estocásticos. O modelo é classificado como determinístico quando para uma mesma entrada, o modelo produz uma mesma saída (com condições iniciais iguais).

iii) Conceitual ou Empírico: os modelos são ditos conceituais quando levam em consideração os conceitos físicos relacionados aos processos hidrológicos. Já os modelos empíricos são aqueles que utilizam equações matemáticas que não tem relação com o comportamento dos processos físicos.

3.4 A modelagem hidrológica na gestão dos recursos hídricos subterrâneos

A previsão dos níveis das águas subterrâneas é essencial para fornecer orientação científica para o uso racional dos recursos hídricos subterrâneos (LI et al., 2019). A medição e análise do nível das águas subterrâneas pode ser usada para quantificar a disponibilidade de água. As variações no nível das águas nos poços fornecem uma medida direta do impacto do uso desse recurso e informações importantes sobre a dinâmica do aquífero estão regularmente incorporadas nas séries temporais de níveis de águas subterrâneas (RAJAEI; EBRAHIMI; NOURANI, 2019).

Assim, a modelagem e as avaliações precisas dos níveis de água subterrânea permitem que engenheiros e partes interessadas da água desenvolvam estratégias para evitar ou reduzir as consequências do uso insustentável dos aquíferos, obtenham uma melhor compreensão da dinâmica e dos fatores subjacentes que afetam os níveis das águas subterrâneas e mantenham um equilíbrio entre o suprimento e as demandas de água (ADAMOWSKI & CHAN, 2011).

Um sistema de águas subterrâneas geralmente envolve aspectos complexos, não-lineares, estocásticos e de múltiplas escalas, que são influenciados por fatores naturais e antropogênicos, o que dificulta a simulação do seu comportamento. Por conta disso, várias técnicas e modelos já foram propostos para simular esse processo (NOURANI; ALAMI; VOUSOUGHI, 2016).

Para modelagem de níveis de águas subterrâneas, modelos conceituais ou físicos são tradicionalmente empregados para descrever variáveis hidrológicas e entender os processos físicos que ocorrem em um sistema (DALIAKOPOULOS; COULIBALY; TSANIS, 2005). Os modelos numéricos de base física são ferramentas robustas e poderosas para simular variações espaciais e temporais nos níveis das águas subterrâneas (YOON et al., 2016).

No entanto, esses tipos de modelos possuem uma série de limitações práticas, incluindo a necessidade de uma grande quantidade de dados hidro geológicos e parâmetros de entrada. Assim, pode não ser viável ter acesso a todas as variáveis que influenciam o processo (RAJAEI; EBRAHIMI; NOURANI, 2019).

Nas bacias hidrográficas onde os dados são limitados e a obtenção de previsões precisas é mais importante do que a compreensão dos mecanismos subjacentes, os modelos empíricos são uma alternativa adequada (NOURANI; ALAMI; VOUSOUGHI, 2016).

Nos últimos anos, os métodos de inteligência artificial (IA), como as redes neurais artificiais (RNAs), têm sido utilizados por sua simplicidade e resultados aceitáveis (RAJAEI; EBRAHIMI; NOURANI, 2019). As RNAs são muito populares entre os modelos empíricos devido à sua capacidade de autoaprendizagem, auto adaptação e alta generalização e têm sido amplamente utilizadas em aplicações hidrológicas nas últimas décadas (ROSHNI, 2020; NATARAJAN; SUDHEER, 2019).

3.5 Redes Neurais Artificiais

3.5.1 Aspectos gerais

As redes neurais artificiais surgiram movidas pelo crescente interesse de compreender o funcionamento do cérebro humano e conseguir reproduzi-lo (ASCE, 2000a). A aptidão encontrada nos sistemas neurais biológicos, para processar tarefas complexas, tem sido remetida à capacidade de elaboração e análise das informações de modo paralelo e distribuído nos neurônios biológicos. Assim, as RNAs foram construídas de forma a imitar as estruturas e os processos de aprendizagem que ocorrem nos sistemas biológicos (SANTOS & SILVA, 2014).

As redes neurais são comumente implementadas por meio de componentes eletrônicos ou simuladas por programação em um computador. (AHMAD et al., 2019). Para atingirem bom desempenho, utilizam-se de uma interligação maciça de células computacionais simples, chamadas “neurônios” (HAYKIN, 2007). Desse modo, uma RNA é constituída por neurônios que estão conectados entre si de acordo com a organização da rede, tendo por objetivo transformar as entradas em saídas relevantes (ADAMOWSKI & CHAN, 2011).

O desenvolvimento de uma RNA é baseado nos seguintes fundamentos: (i) O processamento dos sinais ocorre nos neurônios; (ii) As informações são transmitidas entre os neurônios através de links de conexão; (iii) Cada link de conexão possui um peso que representa a sua influência na conexão; (iv) Cada neurônio geralmente aplica uma transformação não linear, denominada de função de ativação, nas entradas para obter as saídas (ASCE, 2000a).

O conhecimento adquirido pela RNA é obtido a partir do seu ambiente por meio de um processo de aprendizagem (HAYKIN, 2007). O sistema pode aprender com experiências passadas e aplicá-las a eventos futuros (MOGHADDAM et al., 2019). Os pesos na conexão entre os neurônios são utilizados para armazenar esse conhecimento a longo prazo (HAYKIN, 2007). A atualização dos pesos é a principal maneira pela qual a rede neural aprende novas informações (PATTERSON & GIBSON, 2017).

As RNAs podem ser classificadas de acordo com a quantidade de camadas como sendo de camada única, camada dupla ou camadas múltiplas. Também podem ser agrupadas com base na direção do fluxo e do processamento de informações (ASCE, 2000a).

3.5.2 O neurônio biológico

O neurônio biológico é uma célula nervosa que fornece a unidade funcional fundamental para o sistema nervoso (PATTERSON & GIBSON, 2017). O cérebro humano possui mais de 10 bilhões de neurônios, cada qual conectado, através de sinapses, a milhares de outros neurônios (COPPIN, 2010).

Cada neurônio consiste em um corpo celular ou soma, que possui um núcleo celular. Ramificando-se a partir do corpo celular, há uma série de fibras chamadas dendritos e uma única fibra longa chamada axônio. O axônio se estende por uma longa distância. Em geral, um axônio tem 1 cm de comprimento, mas pode alcançar até 1 metro (Figura 1) (RUSSEL & NORVING, 2013).

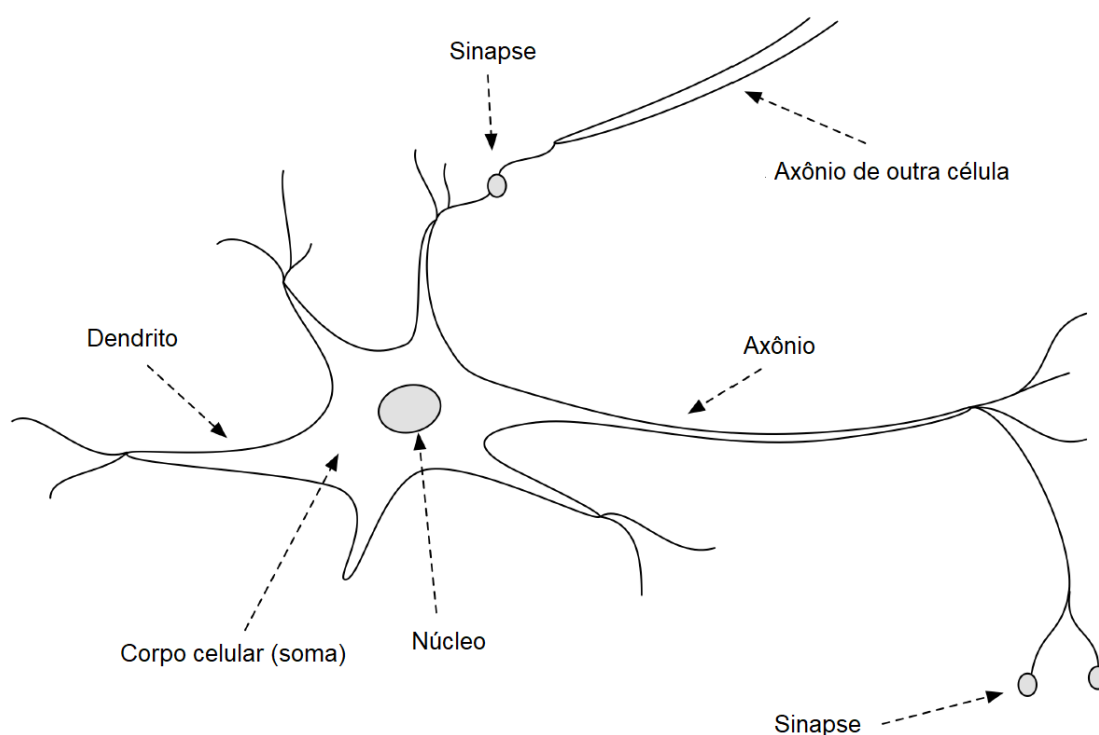


Figura 1 – Neurônio biológico simplificado (Fonte: Adaptado de PATTERSON & GIBSON, 2017).

Os dendritos são especializados na recepção de estímulos nervosos vindos do ambiente ou de outros neurônios. Tais estímulos são transmitidos ao corpo celular, que combina e processa as informações recebidas (FACELI et al., 2011). Quando o sinal de entrada ultrapassa um limiar, o corpo celular gera um novo impulso elétrico, chamado de

potencial de ação, que é enviado ao axônio (a saída do neurônio), rumo às sinapses, unidades que medeiam as interações entre os neurônios (COPPIN, 2010; FACELI et al., 2011)

As sinapses podem ser excitatórias, quando aumentam o potencial de ação, ou inibitórias quando o diminuem (FACELI et al., 2011). A plasticidade diz respeito às mudanças que ocorrem na força das conexões em resposta aos estímulos de entrada ao longo do tempo. Tais mudanças nas forças das conexões em conjunto com as novas conexões que podem se formar entre os neurônios proporciona o aprendizado no cérebro biológico (PATTERSON & GIBSON, 2017).

Ainda que cada neurônio, individualmente, seja simples, a grande e complexa rede de neurônios tem capacidade de processar informações de enorme complexidade (reconhecimento de padrões, percepção e controle motor) de modo muito mais ágil do que o mais rápido computador digital da atualidade (COPPIN, 2010; HAYKIN, 2007).

3.5.3 O neurônio artificial

O neurônio artificial (Figura 2) é uma unidade de processamento de informações, que computa funções matemáticas, fundamental para a operação de uma rede neural (FACELI et al., 2011; HAYKIN, 2007).

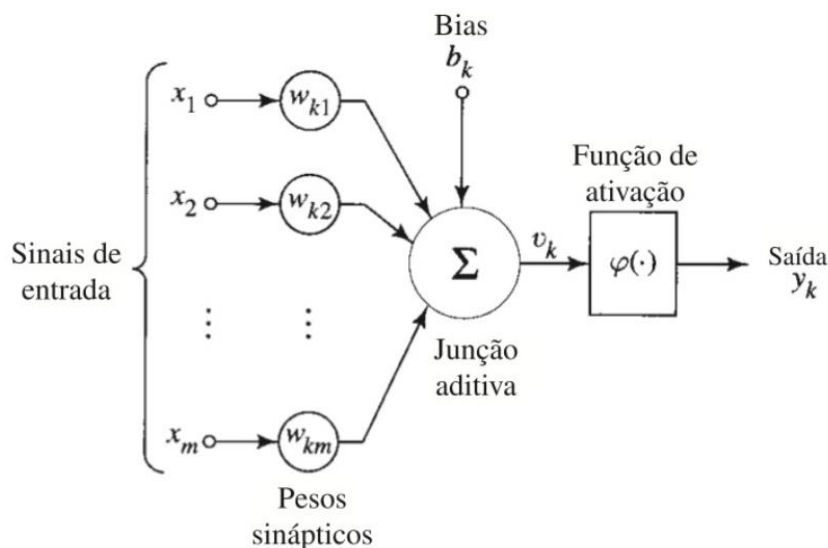


Figura 2 – Modelo de um neurônio artificial (Fonte: HAYKIN, 2007).

Cada neurônio (ou nó) em uma rede neural recebe uma série de sinais ($x_1, x_2, \dots, x_m$) em seus terminais de entrada. Um sinal x_j na entrada da sinapse j conectada ao neurônio k é multiplicada pelo peso sináptico w_{kj} (onde k refere-se ao neurônio analisado e j ao terminal

de entrada da sinapse). Após a multiplicação dos sinais de entrada com os respectivos pesos sinápticos, que podem ser positivos ou negativos, é aplicado um somador para somar os sinais de entrada ponderados (HAYKIN, 2007).

Uma função de ativação é então empregada ao resultado do somador com o intuito de restringir, a um valor finito, a amplitude da saída de um neurônio. O modelo neural também possui um *bias*, b_k , aplicado externamente que tem o efeito de aumentar, caso seja positivo, ou diminuir, se for negativo a entrada líquida da função de ativação (HAYKIN, 2007).

Matematicamente, um neurônio k pode ser descrito da seguinte maneira (Equações 1, 2 e 3):

$$u_k = \sum_{j=1}^m w_{kj} * x_j \quad (1)$$

$$y_k = \varphi(u_k + b_k) \quad (2)$$

$$v_k = u_k + b_k \quad (3)$$

Onde $x_1, x_2, \dots, x_m$ são os sinais de entrada; $w_1, w_2, \dots, w_m$ são os pesos sinápticos do neurônio k ; u_k é a saída do combinador linear; b_k é o *bias*; $\varphi()$ é a função de ativação e y_k é o sinal de saída do neurônio; v_k é o resultado da combinação do *bias* (b_k) com u_k .

3.5.4 Caracterização das RNAs

Uma rede neural é caracterizada pela função de ativação, pela sua arquitetura, que consiste no seu padrão de conexões entre os neurônios e pelo seu método de determinar os pesos nas conexões (denominado de treinamento ou aprendizado) (FAUSETT, 1994).

3.5.4.1 Funções de ativação

Há vários tipos de funções de ativação que podem ser utilizadas nos neurônios. As mais comuns são apresentadas na Figura 3, onde o eixo das abscissas de cada gráfico representa o valor de entrada e o eixo das ordenadas representa a saída ou o nível de ativação do neurônio (COPPIN, 2010).

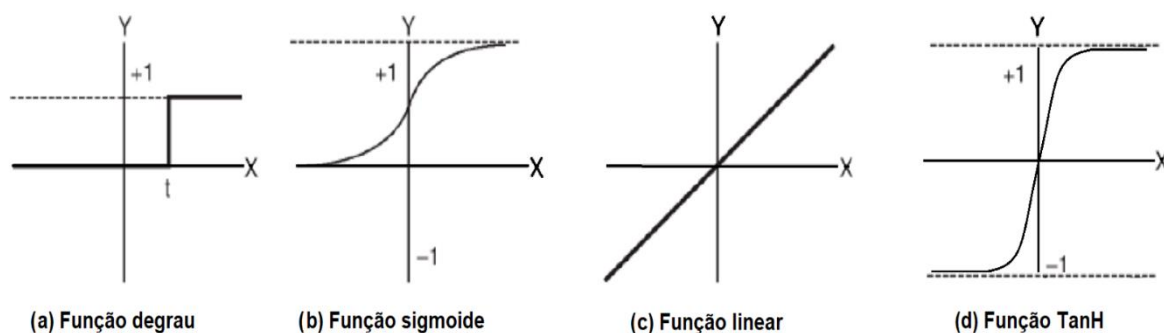


Figura 3 – Funções de ativação (Fonte: Adaptado de COPPIN, 2010).

Para a função degrau (Figura 3a), o valor limiar estabelecido define quando o resultado da função degrau será igual a 1 ou a 0 (ou -1). Caso a soma das entradas ponderadas e combinadas com o *bias* ultrapasse o limiar estabelecido, o neurônio é ativado (saída +1) (FACELI et al., 2011).

A função sigmoide (Figura 3b) representa uma aproximação contínua e diferenciável da função degrau, onde diferentes inclinações podem ser usadas. Já o uso da função linear (Figura 3 c) significa que a saída do neurônio irá possuir o mesmo valor da entrada (FACELI et al., 2011).

A função tangente hiperbólica (TanH) (Figura 3d) usa um tipo semelhante de não linearidade em forma de S como a função sigmoide, porém em vez de variar de 0 a 1, a saída dos neurônios da função TanH varia de -1 a 1 (BUDUMA, 2017).

3.5.4.2 Arquiteturas das RNAs

Redes alimentadas adiante

Em uma rede alimentada adiante (*feedforward networks*), os neurônios estão, geralmente, dispostos em camadas, sendo a primeira delas, a camada de entrada, e a última a de saída, podendo existir várias camadas escondidas. Nesse caso, os neurônios de uma camada não possuem conexões entre si, mas apenas com os da camada seguinte (MAIER et al., 2010).

Em sua forma mais simples, temos uma camada de entrada com neurônios de fonte que se projeta sobre uma camada de saída de neurônios (nós computacionais). Essa rede é

chamada de rede de camada única, referindo-se apenas à camada de saída, já que na camada de entrada não é realizada qualquer computação (HAYKIN, 2007).

Um exemplo de rede neural sem alimentação que é amplamente utilizada na modelagem hidrológica é a *multilayer perceptron* (MLP) (NAYAK; SATYAJI RAO; SUDHEER, 2006). Os sinais de entrada ($x_1, x_2, \dots, x_n$) são transmitidos para as demais camadas e geram as saídas ($y_1, y_2, \dots, y_k$) na última camada. Os neurônios na camada de entrada são conectados aos da camada oculta por meio de valores de ponderação ajustados w_{ji} , já os neurônios da camada oculta estão interligados aos da camada de saída por meio de pesos w_{kj} (Figura 4) (SANTOS & SILVA, 2014).

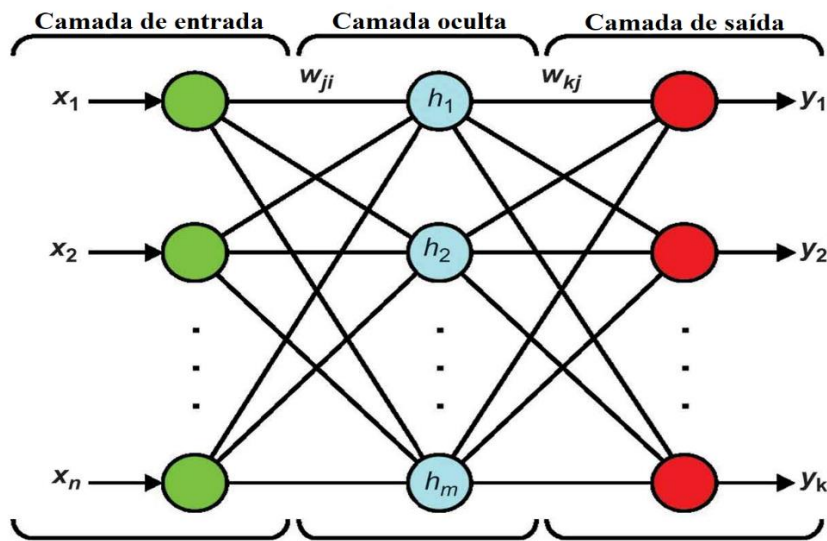


Figura 4 – Rede neural do tipo MLP com uma camada intermediária (Adaptado de SANTOS & SILVA, 2014).

A Equação 4 demonstra a correlação para os valores de saída:

$$y_k = f_o \left(\sum_{j=1}^s w_{kj} \cdot f_h \left(\sum_{i=1}^s w_{ji} x_i + b_j \right) + b_k \right) \quad (4)$$

Sendo: w_{ji} o peso que conecta o i -ésimo neurônio na camada de entrada ao j -ésimo neurônio na camada oculta; w_{ki} o peso que conecta o j -ésimo neurônio na camada oculta com

o k -ésimo neurônio na camada de saída; b_j é o viés para o j neurônio oculto; b_k é o viés para o k -ésimo neurônio de saída; f_h é a função de ativação dos neurônios na camada oculta; e f_o é a função de ativação dos neurônios na camada de saída (SANTOS & SILVA, 2014).

A rede neural da Figura 4 é dita totalmente conectada, no sentido de que cada um dos neurônios de uma camada da rede está conectado a todos os neurônios da camada seguinte. Porém, se algumas conexões sinápticas estiverem faltando, a rede é chamada de parcialmente conectada (HAYKIN, 2007).

Redes recorrentes

Redes neurais recorrentes distinguem-se de uma rede neural alimentada adiante por ter ao menos um laço de realimentação. As informações fluem pelos neurônios nas duas direções, isso acontece por meio de uma reciclagem que transforma as saídas de uma camada em entradas das camadas anteriores (MAIER et al., 2010).

3.5.4.3 Aprendizagem da RNA

A aprendizagem é um processo contínuo de estimulação pelo ambiente no qual a RNA está contida, através do qual os pesos de conexão são reajustados. Após a realização desse procedimento, espera-se que a RNA esteja pronta para obter resultados satisfatórios a partir de novos dados de entrada. Existem basicamente dois tipos de aprendizagem, supervisionada e não supervisionada (ASCE, 2000a).

Na aprendizagem supervisionada, a rede neural é apresentada a um vetor de treinamento que possui um vetor de saída associado, representando a resposta desejada a ser obtida pela rede neural. Os pesos são então ajustados de acordo com um algoritmo de aprendizado (FAUSETT, 1994).

Um dos algoritmos de treinamento mais utilizados para previsão e treinamento de redes de séries temporais é o de *Levenberg-Marquardt* (LM) (GUZMAN; PAZ; TAGERT, 2017). Este método foi apresentado em 1963 como um dos algoritmos de treinamento mais rápidos (KOUZIOKAS; CHATZIGEORGIOU; PERAKIS, 2018). O algoritmo de LM é uma variação do algoritmo de Gauss Newton que encontra a função mínima e otimiza a solução (GUZMAN; PAZ; TAGERT, 2017).

Para a aprendizagem não supervisionada, uma sequência de vetores de entrada é fornecida, mas nenhum vetor de saída associado é especificado. A rede modifica os pesos

para que os vetores de entrada mais semelhantes sejam atribuídos à mesma saída (ou unidade de *cluster*). A rede neural produzirá um vetor exemplar (representativo) para cada *cluster* formado (FAUSETT,1994).

3.6 Redes neurais artificiais para modelagem dos níveis de águas subterrâneas

Nos últimos anos, as RNAs têm sido aplicadas com sucesso na modelagem dos níveis de água subterrânea em diferentes partes do mundo (SUN et al., 2016). Entre esses estudos, há diferentes abordagens de modelagem, como por exemplo na escolha dos dados de entrada, na quantidade e qualidade dos dados usados, diferentes áreas de estudo e aquíferos, tempo de previsão, divisão de dados, entre outros (RAJAEI; EBRAHIMI; NOURANI, 2019). Alguns trabalhos que utilizaram RNAs para a previsão dos níveis das águas subterrâneas são citados abaixo.

Lallahem et al. (2005) avaliaram a viabilidade do uso da metodologia de RNA para estimar o nível mensal das águas subterrâneas em um aquífero não confinado no norte da França. Os dados de entrada foram os níveis das águas subterrâneas de 13 piezômetros, precipitação, temperatura média, precipitação efetiva e evapotranspiração potencial. O objetivo principal do trabalho foi simular o nível de água subterrânea em um piezômetro selecionado. As simulações revelaram o mérito do uso de modelos do tipo MLP.

Sun et al. (2016) aplicaram uma RNA do tipo MLP treinada com o algoritmo de LM para prever o nível d'água subterrânea diário em uma floresta pantanosa em Cingapura. As entradas para o modelo foram os níveis dos reservatórios circundantes e a precipitação. Os resultados revelam que a RNA é capaz de produzir uma previsão precisa com para 1 dia à frente, enquanto o desempenho diminui quando o tempo de previsão aumenta para 3 e 7 dias à frente.

Kouziokas, Chatzigeorgiou e Perakis (2018) usaram várias RNAs do tipo *feedforward networks* com múltiplas estruturas de rede e algoritmos de treinamento para prever o nível de água subterrânea diário de um poço localizado no Condado de Montgomery, Pensilvânia, EUA. Os modelos de previsão foram desenvolvidos aplicando diferentes estruturas em relação ao número de neurônios em cada camada oculta e ao número de camadas ocultas utilizadas. Como entrada do modelo foram usadas a umidade, a precipitação e a temperatura. A RNA do tipo *feedforward network* treinada com o algoritmo de LM foi o melhor modelo.

3.7 Análise de *cluster*

A análise de *cluster* cria grupos (*clusters*), a partir de dados, de modo que objetos pertencentes ao mesmo *cluster* são mais parecidos entre si do que com objetos de outros grupos. A maioria das técnicas de análise de *cluster* se baseia em uma medida de similaridade entre os elementos a serem agrupados (LIU; BARAHONA, 2020).

A análise de *cluster* pode ser realizada usando dois métodos diferentes, baseados em técnicas hierárquicas e de particionamento (não hierárquicas) (DE BRITO NETO et al., 2016).

3.7.1 Métodos por particionamento

Há várias técnicas de análise de *cluster* que usam métodos por particionamento, como por exemplo as técnicas *k-medoids*, *fuzzy c-means* e *fuzzy c-medoids* (DE BRITO NETO et al., 2016). A técnica clássica de agrupamento por particionamento é chamada de *k-means*. O agrupamento *k-means* começa a partir da atribuição inicial de k centroides, onde o número k é definido pelo usuário. Os pontos de dados são então atribuídos aos centroides mais próximos para formar os *clusters*, e as posições dos centroides são atualizadas com base nos pontos atribuídos aos diferentes *clusters*. Esse processo é repetido até que os centroides permaneçam os mesmos. Uma medida de proximidade deve ser definida para quantificar a distância de cada ponto de dados aos centroides (FABBROCINO et al., 2019).

É importante tomar cuidado com a seleção dos centroides iniciais já que arranjos completamente diferentes podem surgir de pequenas mudanças na escolha aleatória inicial. Para aumentar a chance de encontrar um mínimo global, é comum executar o algoritmo várias vezes com diferentes escolhas iniciais e escolher o que obteve o melhor resultado no final do processo (WITTEN; FRANK; HALL, 2011).

3.7.2 Métodos hierárquicos

A análise hierárquica de agrupamentos, *hierarchical cluster analysis* (HCA), é uma das técnicas de análise de agrupamentos mais populares e pode ser aplicada com sucesso para dividir as amostras em grupos semelhantes. A divisão em *clusters* é baseada na similaridade entre objetos dentro de um mesmo *cluster* e em diferenças entre diferentes *clusters* (ABU-ALNAEEM et al., 2018).

Existem dois tipos principais de métodos hierárquicos de agrupamento, o método aglomerativo (ascendente) e o método divisivo (descendente). O algoritmo aglomerativo considera cada item como um *cluster* e depois funde gradualmente os *clusters* (de baixo para cima). Por outro lado, o algoritmo divisivo começa com todos os objetos como um único *cluster* e depois divide o *cluster* para alcançar *clusters* com apenas um objeto (de cima para baixo) (AGHABOZORGI; SEYED SHIRKHORSHIDI; YING WAH, 2015).

No algoritmo de *cluster* hierárquico básico, o processo de combinação começa considerando cada objeto individual como um *cluster*. Os grupos são então vinculados, de acordo com sua similaridade, até restar apenas um *cluster*. O resultado da análise de *cluster* hierárquico pode ser observado como um dendrograma (diagrama em árvore) (Figura 5), cujas raízes são todos os objetos (ABU-ALNAEEM et al., 2018).

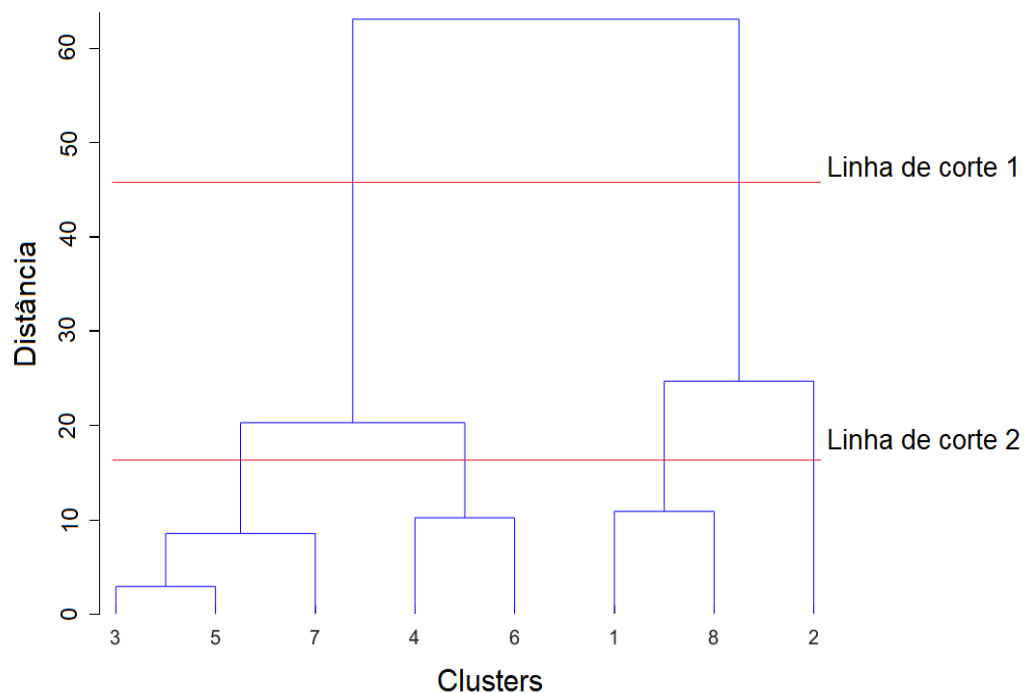


Figura 5 – Exemplo de dendrograma.

O dendrograma não é um conjunto único de *clusters*, mas uma hierarquia multinível, em que *clusters* em um nível são unidos como *clusters* no próximo nível. Isso permite que o usuário decida o nível ou a escala do armazenamento em *cluster* mais apropriado para seus objetivos (MATLAB, 2017).

O número de *clusters* é determinado com base na linha de corte, e a alteração na localização da linha de corte no dendrograma altera o número de *clusters* (YANG et al.,

2020). Na Figura 5, a linha de corte 1 equivale a dois *clusters*, o *cluster* 1 engloba os seguintes grupos: 3, 5, 7, 4 e 6. Já o *cluster* dois possui os grupos 1, 8 e 2. Na linha de corte 2 o número de *clusters* aumenta para quatro, sendo: *cluster* 1 (grupos 3,5 e 7); *cluster* 2 (grupos 4 e 6); *cluster* 3 (grupos 1 e 8) e *cluster* 4 (grupo 2).

A HCA dispõe de várias métricas de dissimilaridade e métodos de agrupamento disponíveis para categorizar diferentes variáveis em grupos homogêneos. As métricas de dissimilaridade referem-se às medidas estatísticas que relacionam a distância entre os pares de objetos que formarão cada grupo, criando cada *cluster* específico e incluem: distância euclidiana, correlação, geometria do táxi (*cityblocks*) e parâmetros de Spearman (SANTOS et al., 2019).

Já os métodos de agrupamento correspondem a algoritmos utilizados para calcular a distância entre *clusters* definidos. Os métodos hierárquicos mais usados incluem: *single linkage*, *complete linkage*, *average linkage*, centroide e *Ward's minimum variance* (SANTOS et al., 2019).

O *single linkage* é definido como medida da distância mínima entre grupos, isto é, a distância entre os dois membros mais próximos. Já o *complete linkage* é a medida da distância máxima entre os *clusters*. O centroide usa a distância do centroide entre seus membros, representando um comprometimento entre os extremos da distância mínima e máxima entre os membros do *cluster*. Já o *average linkage* calcula a distância média entre cada par de membros de dois *clusters* (WITTEN; FRANK; HALL, 2011).

Por fim, o *Ward's minimum variance* assume que um *cluster* é representado por seu centroide e mede a proximidade entre dois *clusters* com o aumento na soma dos erros ao quadrado resultantes da mesclagem dos dois *clusters* (FABBROCINO et al., 2019).

Cada método e métrica tem suas vantagens e desvantagens em comparação uns com os outros e, dependendo do objetivo do estudo, uma certa combinação pode ser mais adequada que outra (SANTOS et al., 2019). A escolha de qual combinação usar de método e métrica não tem uma resposta fácil e mudanças nessas combinações afetam bastante o resultado dos *clusters* (MACHIWAL et al., 2018).

4 CARACTERÍSTICAS DA ÁREA DE ESTUDO

4.1 Localização e Características gerais

O Aquífero Ogallala é o maior aquífero da América do Norte, estende-se por uma área de 450.660 km² e abrange oito estados norte-americanos (Texas, Oklahoma, Kansas, Novo México, Colorado, Nebraska, Wyoming e Dakota do Sul) (Figura 6) (CANO et al., 2018). É a parte não confinada de um sistema maior chamado *High Plains Aquifer*. Os estados que ele engloba são comumente referidos como *High Plains Region* (GOWDA et al., 2019).

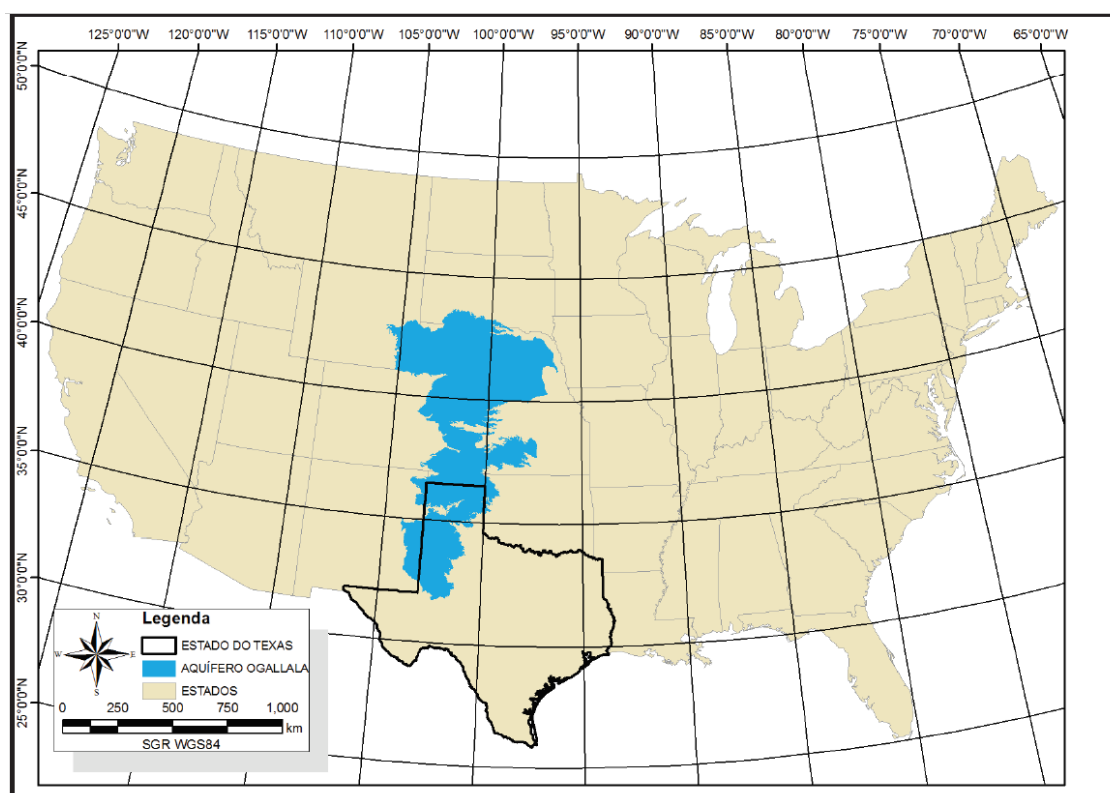


Figura 6 – Mapa da área de cobertura do Aquífero Ogallala (DE BRITO NETO, 2012).

A região em que o Aquífero Ogallala está situado é caracterizada como uma pradaria semiárida de clima temperado (STEWART; ALLEN, 2016). É um local conhecido por sua alta variabilidade climática e secas prolongadas, possui um gradiente de temperatura norte-sul, ficando mais quente para o sul (GOWDA et al., 2019). As temperaturas diárias podem variar de 15 ° a 30 °C (Celsius). Os extremos sazonais de temperatura chegam a variar cerca de 70 °C entre as baixas de inverno e as máximas de verão (GUTENTAG et al., 1984).

A região possui precipitação moderada, ventos frequentes, baixa umidade e uma alta taxa de evaporação (GUTENTAG et al., 1984). Além disso, possui recursos hídricos

superficiais limitados, como córregos efêmeros, lagos e poucos rios perenes (STEWART; ALLEN, 2016).

A área de estudo escolhida para esta pesquisa consiste na parte do Aquífero Ogallala compreendida no estado do Texas, que possui uma área de aproximadamente 90.000 km² (Figura 7) (DE BRITO NETO et al., 2016).

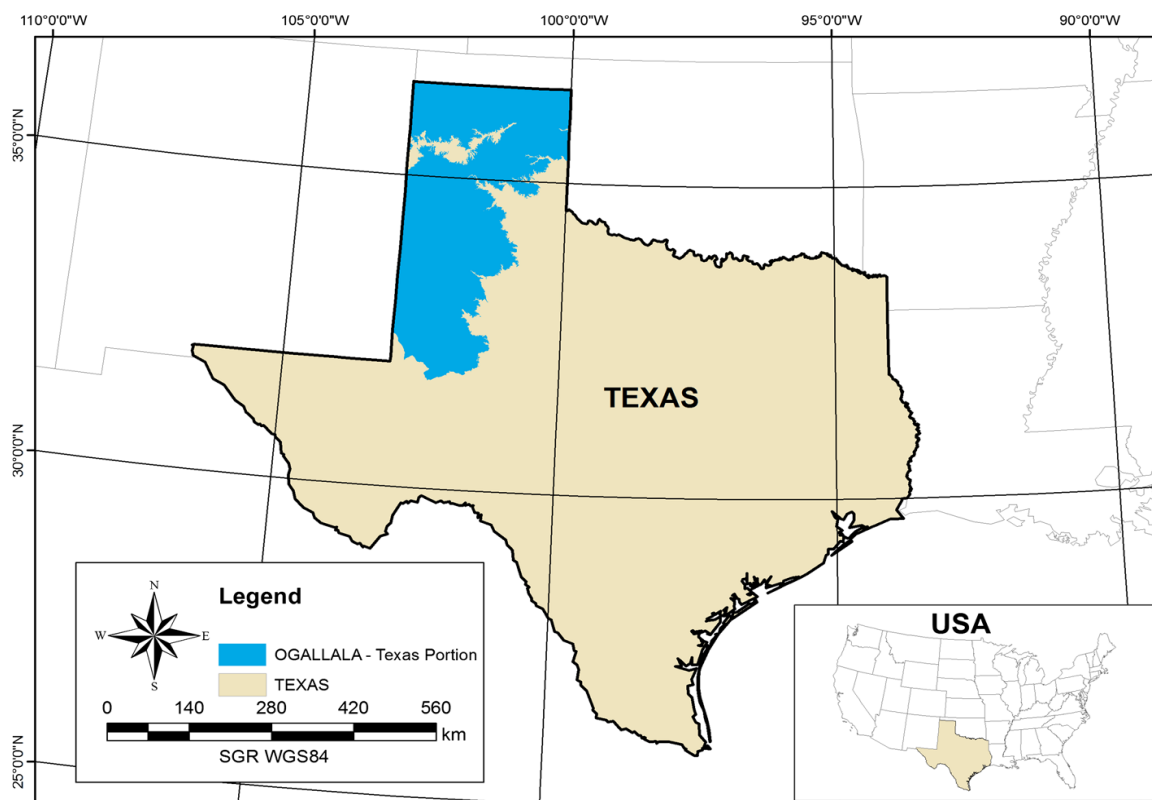


Figura 7 – Área de estudo: Aquífero Ogallala compreendido no estado do Texas (DE BRITO NETO et al., 2016).

A precipitação anual na região do Texas varia de cerca de 360 mm a oeste a 610 mm a leste. A maior parte da precipitação ocorre durante os meses de maio e setembro. Os meses de outubro a fevereiro são geralmente secos, já os meses de março, abril e maio são os que mais ventam (ADHIKARI et al., 2016).

Os solos da região do Texas são caracterizados como solos profundos e bem desenvolvidos, com elevado teor de argila e acúmulo de carbonato de cálcio nos horizontes do subsolo (ADHIKARI et al., 2016).

4.2 Características hidrogeológicas

O Ogallala é um aquífero não confinado que foi formado durante a idade quaternária ou terciária, há aproximadamente 10 milhões de anos atrás (MODALA et al., 2017). Ele foi

constituído a partir da erosão das *Rocky Mountains*, a oeste da região, pelo transporte e deposição subsequente dos sedimentos para elevações mais baixas a leste (STEWART; ALLEN, 2016).

O aquífero é composto por areia, cascalho, argila e silte não consolidados depositados em um ambiente fluvial continental (produto do intemperismo das *Rocky Mountains*) e eólico (TWDB, 2016; CHAUDHURI; ALE, 2014a). Sedimentos do Aquífero Ogallala são frequentemente cimentados com carbonato de cálcio, conhecido por sua resistência às intempéries (CHAUDHURI; ALE, 2014a).

O aquífero possui uma espessura máxima de 243,84 metros. A espessura saturada de água doce tem em média 28,96 metros (TWDB, 2016). A condutividade hidráulica e o rendimento específico do aquífero variam bastante, pois dependem dos tipos de sedimentos presentes, que mudam amplamente tanto na horizontal quanto na vertical. A condutividade hidráulica possui uma média de 18,3 metros por dia e o rendimento específico é em média cerca de 15% (GUTENTAG et al., 1984).

O fluxo de água no aquífero possui uma taxa média de cerca de 30,5 cm por dia. O fluxo é impulsionado pela recarga do aquífero de acordo com o gradiente hidráulico, que geralmente atua no sentido oeste-leste do Ogallala (STEWART; ALLEN, 2016).

A recarga do aquífero ocorre principalmente por infiltração de precipitação. O retorno da irrigação afeta a recarga apenas em escalas locais com base no uso da terra. Além disso, também ocorre infiltrações de formações hidrogeológicas adjacentes. A descarga do aquífero ocorre principalmente por vazamentos para outras formações hidrogeológicas, nascentes e por bombeamento de água para uso municipal e na irrigação (CHAUDHURI; ALE, 2014a; CHAUDHURI; ALE, 2014b).

4.3 Importância e problemas enfrentados no Aquífero Ogallala

O Aquífero Ogallala é considerado um dos maiores aquíferos de água doce do mundo e atua como um recurso valioso para a agricultura, produção animal e abastecimento público de água (CANO et al., 2018).

O Ogallala ocupa o primeiro lugar no país em total de captação de água subterrânea, atendendo a 23% das necessidades totais de água subterrânea dos EUA e 30% das necessidades de água para irrigação do país (CHAUDHURI; ALE, 2014a).

A agricultura dos EUA depende fortemente do Aquífero Ogallala. A região onde o aquífero está localizado representa 30% da produção total de culturas e animais dos EUA e mais de 90% da água bombeada do Aquífero Ogallala é usada na agricultura irrigada. A

produção agrícola irrigada tem um tremendo impacto na economia rural dessa região, aumentando os valores da produção da agrícola em mais de US \$ 12 bilhões anualmente (CANO et al., 2018).

O semiárido do *Texas High Plains* (THP) possui 1,8 milhão de hectares de terras irrigadas, sendo a principal fonte de irrigação o Aquífero Ogallala (KOTHARI et al., 2019). As principais culturas irrigadas nessa região incluem milho, algodão, sorgo e trigo com pequenas áreas de amendoim, silagem e soja (CHAUDHURI; ALE, 2014a).

A agricultura na região do Aquífero Ogallala pertencente ao Texas contribui com aproximadamente US\$ 7 bilhões anualmente em atividade econômica, produzindo 25% da produção de algodão dos EUA (MITCHELL-MCCALLISTER et al., 2020).

A água do Aquífero Ogallala vem sendo retirada a uma taxa muito acima da sua recarga natural (KOTHARI, 2019). Décadas de bombeamento combinadas com recarga mínima resultaram em declínios significativos na espessura saturada do aquífero, especialmente na sua porção sul (MAREK et al., 2020; KOTHARI, 2019).

O suprimento de água subterrânea do estado do Texas, proveniente do Ogallala, foi projetado para reduzir em cerca de 30% até 2060 (CHAUDHURI; ALE, 2014a). Por outro lado, espera-se que a população do estado duplique, e que haja uma demanda crescente de água em 80% até o ano 2060, havendo assim, a necessidade de uma avaliação sobre a disponibilidade futura de água subterrânea em todo o estado (CHAUDHURI; ALE, 2014b).

Além disso, projeções indicam mudanças consideráveis no clima do Texas no futuro que podem impactar potencialmente a agricultura nessa região (MODALA et al., 2017). Estudos de mudança climática para esta região preveem verões mais quentes e reduções na precipitação anual no futuro, e tais tendências exigem maiores captações de água subterrânea para maiores necessidades de evapotranspiração das culturas agrícolas (ADHIKARI et al., 2016).

A espessura saturada do aquífero e as capacidades dos poços influenciam diretamente o uso da terra e a composição das culturas (CHEN et al., 2018). A diminuição dos níveis de água impacta a produção agrícola, a exemplo da produção de algodão, devido ao aumento dos custos de bombeamento das águas subterrâneas e a menor disponibilidade de água (ADHIKARI et al., 2016).

Tais fatos ameaçam a sustentabilidade a longo prazo da produção de alimentos na região (GOWDA, 2019). Portanto, torna-se imperativo adotar práticas eficientes de gestão da água para sustentar o abastecimento de água subterrânea para uso agrícola, industrial e municipal nessa região (KOTHARI, 2019).

5 MATERIAL E MÉTODOS

Na Figura 8 é apresentado o fluxograma da metodologia aplicada neste trabalho.

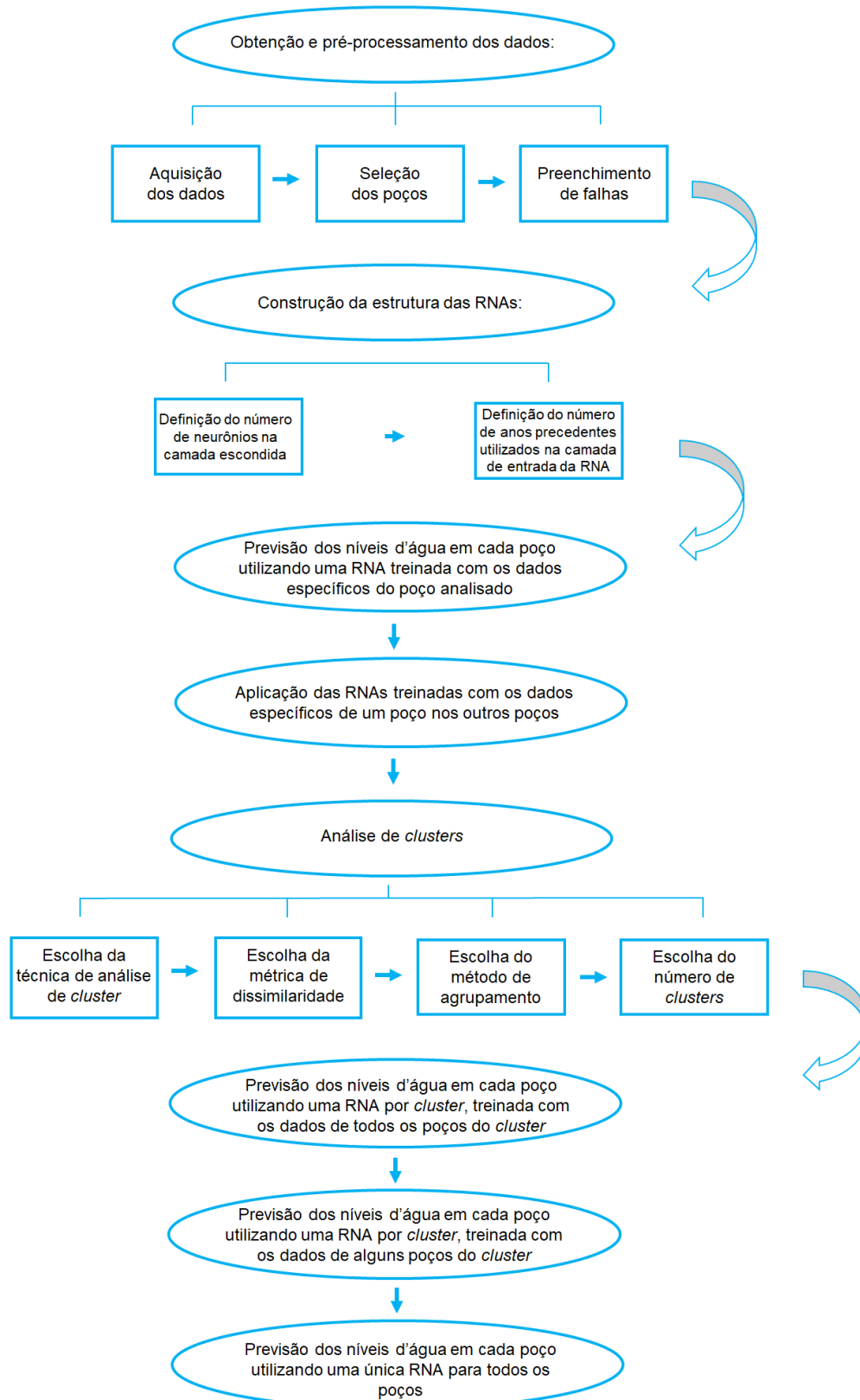


Figura 8 – Fluxograma da metodologia aplicada.

5.1 Obtenção e pré-processamento dos dados

Os dados utilizados neste trabalho foram obtidos a partir da plataforma do *Texas Water Development Board* (TWDB). Este website dispõe de um banco de dados de águas subterrâneas contendo informações sobre os poços, incluindo dados de localização, profundidade, tipo de poço, proprietário, perfurador, construção e conclusão, dados de aquíferos, nível de água e qualidade da água. Esse banco de dados de águas subterrâneas é um dos mais abrangentes em todo os Estados Unidos (TWDB, 2018).

Primeiramente, foi feito o download dos registros dos poços para o Aquífero Ogallala em formato (.xlsx) do Microsoft Excel. Tais registros possuem informações como: código do poço, município em que o poço está localizado, código do aquífero, profundidade do poço, latitude e longitude, data da coleta do dado, nível da água no poço (profundidade da água em metros abaixo da superfície da terra), elevação da água do poço (elevação do nível de água no poço em metros acima do nível do mar), entre outros.

Tal arquivo do Microsoft Excel possui 279.607 registros para 13.193 poços. Após uma análise preliminar, foi verificado que havia uma grande diferença na quantidade de registros pertencentes a cada poço, havendo poços que possuíam apenas um e outros que possuíam mais de 3.000 registros. Além disso, constatou-se que a grande maioria dos poços não possuíam séries históricas consistentes, possuindo poucos dados coletados em meses de anos específicos e/ou muitas falhas ao longo do período de coleta dos dados. Sendo necessária, dessa forma, uma filtragem dos poços que poderiam ser usados para esse estudo.

Para isso, foi necessário escolher quais os meses do ano em que ocorriam a coleta de dados dos poços que seriam utilizados para definir o nível de água anual. Normalmente os níveis de água nos poços são medidos, pelo distrito local, durante o inverno, quando o bombeamento de água nos poços, principalmente para irrigação, é parado e o lençol freático teve tempo de se recuperar da estação de crescimento das culturas agrícolas (STOUT, 2018). Assim, os níveis de água tendem a ser mais altos durante os meses de inverno e mais baixos durante a temporada agrícola de verão (STOUT, 2018).

Dessa forma, foram considerados os dados obtidos durante os meses de inverno, de novembro a fevereiro, que formavam a grande maioria dos registros encontrados, já que esse é o período preferencial para a obtenção dos dados (CHAUDHURI; ALE, 2014b). Para os níveis de água que foram coletados nos meses de novembro e dezembro foram atribuídos a data do ano seguinte. O motivo dessa alteração foi considerar o intervalo entre os meses de novembro a fevereiro como sendo pertencentes ao mesmo ano, visto que, caso não fosse

adotado esse procedimento, poderia haver até doze meses de defasagem entre um dado coletado e outro no mesmo ano. Para os poços em que havia mais de um dado, de elevação do nível da água, coletado entre os meses de novembro e fevereiro foi feita uma média dos dados registrados nesse período.

Preliminarmente, foram selecionados os poços que possuíam uma maior densidade de dados. Dentre os poços escolhidos inicialmente, foram excluídos aqueles que possuíam, ao longo da série histórica, mais de dois anos consecutivos com falhas, de maneira que, com o preenchimento das falhas, etapa seguinte a da seleção dos poços, não fosse necessário interpolar para mais de duas falhas consecutivas, tendo em vista que tais valores interpolados poderiam torna-se duvidosos e assim influenciar negativamente as previsões da rede neural. Finalmente, de um universo de 13.193 poços, foram selecionados 403 poços (Figura 9).

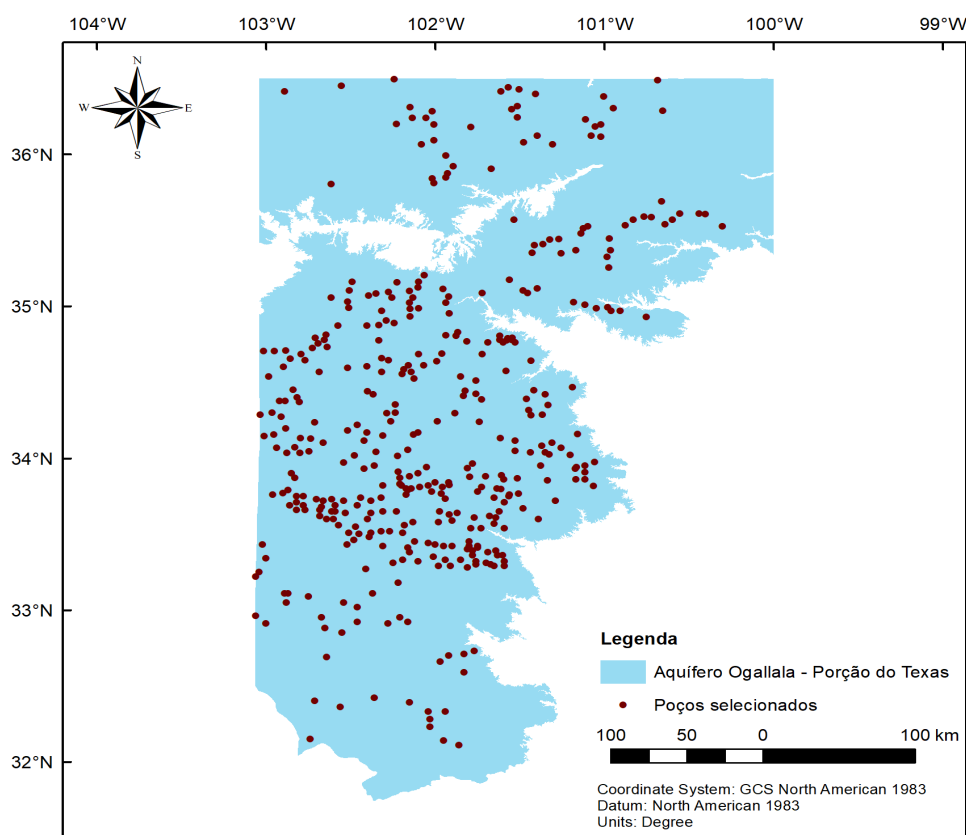


Figura 9 – Distribuição espacial dos 403 poços selecionados para a pesquisa.

A partir dos 403 poços selecionados foi formada uma matriz no MATLAB R2017a com 60 linhas, equivalente ao número de anos resultantes da série histórica (de 1960 a 2019) e 403 colunas, equivalente ao número de poços. Entretanto, alguns poços não possuem dados de nível d'água a partir de 1960, começando a apresentar dados depois dessa data, e nem

todos os poços selecionados possuem dados coletados até o ano de 2019. Assim, tais poços apresentam séries históricas menores do que 60 anos.

A Figura 10 apresenta onde se encontram as falhas nas séries históricas dos poços selecionados para este trabalho. Essa figura representa a matriz dos dados utilizados com 60 linhas e 403 colunas. Cada linha corresponde ao ano em que a medição do nível d'água foi efetuada (do ano de 1960 a 2019) e cada coluna representa a série histórica dos dados de nível d'água de um poço. As células que apresentam falhas estão marcadas na cor preta e as células que possuem dados estão na cor branca.

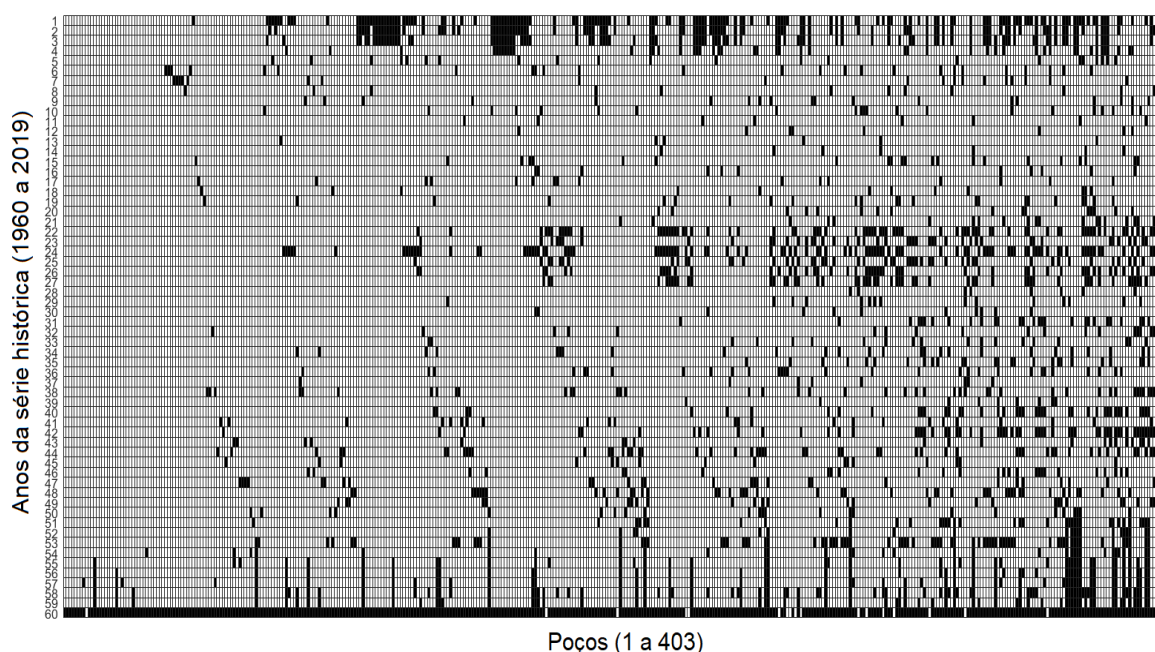


Figura 10 – Distribuição de falhas (células na cor preta) nas séries temporais do nível d'água de cada poço.

Na Figura 11 é apresentado um gráfico de dispersão que mostra a quantidade de falhas que cada poço possui ao longo dos 60 anos. O número máximo de falhas encontradas foi de 19. É importante notar que há poços cujas séries históricas são menores do que 60 anos e portanto o número total de falhas encontradas para cada poço pode não corresponder às falhas existentes apenas dentro dos limites inicial e final de sua série histórica (entende-se por limite inicial e final da série histórica o ano em que há o primeiro e o último dado coletado presente neste trabalho). Desse modo, o número total de falhas para cada poço apresentado na Figura 11 é uma soma das falhas existentes dentro da série histórica do poço com as falhas existentes no extremo inicial e/ou final da série.

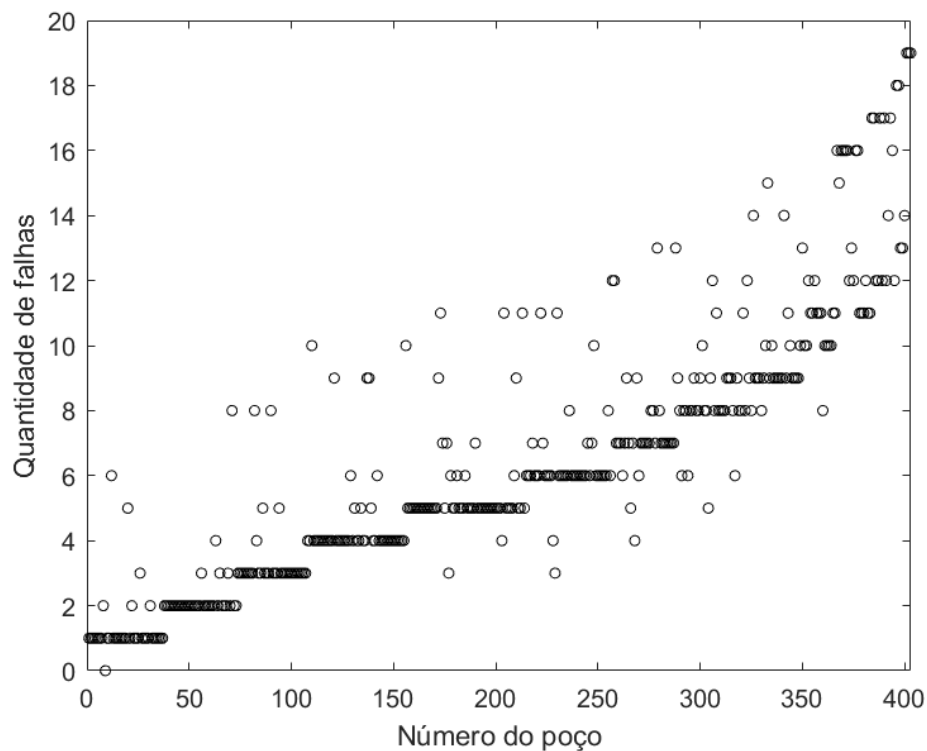


Figura 11 – Quantidade de falhas existentes nas séries históricas de cada um dos 403 poços.

De modo a otimizar e melhor aproveitar a distribuição temporal dos dados de nível da água, foi criada uma rotina, no MATLAB R2017a, que ignora as células que contenham falhas no extremo inicial e/ou final da série histórica de cada poço. Por exemplo se um poço possui falhas de 1960 a 1962, a rotina ignora esses anos iniciais com falhas e começa a trabalhar apenas a partir do ano de 1963. Se, em outro caso, um poço possui falha no ano de 2019, a rotina ignora esse último ano e considera como válido até o ano de 2018. Com isso evita-se de ter que extrapolar os valores para preencher essas falhas ou de excluir mais poços por não possuírem tais dados. Após isso, foi necessário fazer o preenchimento dos dados que apresentavam falhas, mas que não se encontravam nos extremos da série.

5.1.1 Preenchimento de falhas

Para realizar o preenchimento das falhas dos dados que não se encontravam nos extremos da série, foram testadas três metodologias distintas a fim de se escolher a que melhor poderia representar os dados de nível d'água nos poços.

A primeira das metodologias aplicadas foi o método da ponderação regional. Esse é um método normalmente utilizado para o preenchimento de falhas de séries mensais ou anuais de precipitação. Para isso, utilizam-se postos pluviométricos vizinhos pertencentes a

uma região climatológica semelhante ao posto a ser preenchido (TUCCI & SILVEIRA, 2004).

Seguindo a metodologia do método da ponderação regional, no caso do preenchimento de falhas nas séries históricas dos poços, utilizou-se três poços semelhantes, ou seja, aqueles que possuíam uma maior correlação com os dados do poço a ser preenchido. A Equação 5 mostra a fórmula do método da ponderação regional adaptado para encontrar o nível d'água no poço que possui a falha.

$$P_x = \frac{1}{3} \left(\frac{P_A}{M_A} + \frac{P_B}{M_B} + \frac{P_C}{M_C} \right) M_x \quad (5)$$

Onde:

P_x = Falha a ser corrigida

P_A, P_B, P_C = Nível d'água nos poços A, B, C, respectivamente

M_A, M_B, M_C = Média do nível d'água nos poços A, B, C, respectivamente

M_x = Média do nível d'água no poço a ser corrigido

O método utilizado para medir a correlação entre os poços e verificar a semelhança entre eles foi o coeficiente de correlação linear de Pearson (Equação 6), este coeficiente oscila de -1 a 1 , com valores absolutos mais altos indicando melhor dependência entre as variáveis (BERMUDEZ-EDO; BARNAGHI; MOESSNER, 2018). O coeficiente de Pearson entre duas variáveis aleatórias x e y é definido como:

$$P(x, y) = \frac{cov(x, y)}{\sqrt{var(x) * var(y)}} \quad (6)$$

Onde, $cov(x, y)$ é a covariância, $var(x)$ é a variância de x e $var(y)$ é a variância de y .

Os outros dois métodos testados para o preenchimento de falhas foram os métodos da interpolação linear e da interpolação *spline*. Ambos foram executados com a função nativa *interp1* do MATLAB R2017a.

O método de interpolação linear é mostrado na Figura 12. Para o caso específico desta pesquisa, os valores no eixo das abscissas representam os anos da série histórica em que as medições do nível d'água foi obtido. Já os valores do eixo das ordenadas, são os dados de nível d'água no poço.

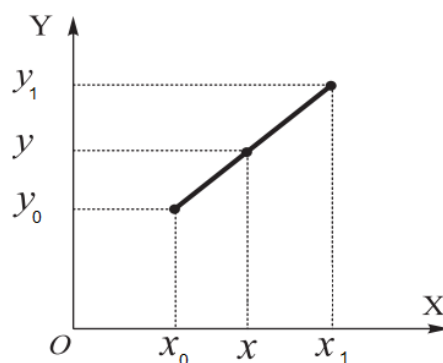


Figura 12 – Representação gráfica do método da interpolação linear (Adaptado de JIANG et al., 2018).

Para a interpolação linear (Equação 7) os valores de y_0 , y_1 , x_0 , x e x_1 são conhecidos, a variável não conhecida y , equivalente ao valor do nível d'água que se deseja encontrar, é dada por (JIANG et al., 2018):

$$y = y_0 + (y_1 - y_0) \frac{x - x_0}{x_1 - x_0} \quad (7)$$

Já para a interpolação *spline*, o algoritmo *spline* realiza uma interpolação cúbica para produzir polinômios por partes com derivadas contínuas de segunda ordem. O resultado é comparável a uma interpolação polinomial regular, mas é menos suscetível a oscilações pesadas entre pontos de dados para altos graus. O valor interpolado em um ponto de consulta é baseado em uma interpolação cúbica dos valores nos pontos de grade vizinhos em cada dimensão respectiva (MATLAB, 2017).

A Figura 13 mostra o resultado da aplicação das três metodologias descritas para o preenchimento das falhas do poço 329, escolhido aleatoriamente. Ao observar os resultados das interpolações percebeu-se que o método da ponderação regional não se adequava bem aos dados de nível d'água, entre a interpolação linear e a interpolação *spline*, optou-se por escolher a do tipo *spline*, já que esse tipo de interpolação deixa a curva mais suave.

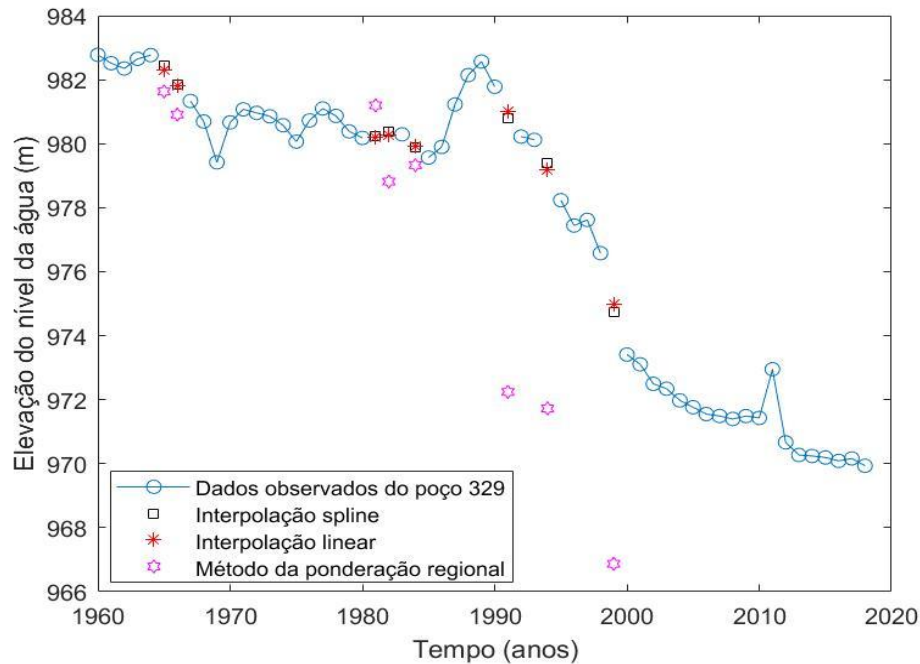


Figura 13 – Diferentes metodologias para o preenchimento das falhas do poço 329.

5.2 Arquitetura das RNAs

Os modelos de RNAs, para previsão do nível de água subterrânea do Aquífero Ogallala, foram desenvolvidos utilizando o software MATLAB. Os modelos construídos para as RNAs possuem uma camada de entrada, uma camada oculta única e uma camada de saída consistindo em um nó que indica o nível da água subterrânea anual. Em relação a direção de fluxo e de processamento das informações, as RNAs são do tipo *feedforward networks*, que são bastante aplicadas em estudos sobre recursos hídricos como ferramenta de previsão (NOURANI; ALAMI; VOUSOUGH, 2015). Para o processo de aprendizagem as RNAs geradas utilizam o algoritmo de treinamento de *Levenberg-Marquardt*. As séries de dados foram divididas de maneira aleatória, sendo 70% para treinamento, 15% para validação e 15% para testes.

A fim de se determinar o número de neurônios na camada oculta foram feitos testes, em uma base de tentativa e erro, de modo a encontrar um número que melhor atendesse a todos os poços. Assim, de um universo de 403 poços, foram escolhidos 10 poços que não possuíam falhas, visando evitar que as falhas que foram preenchidas, em poços que não possuíam uma série histórica completa, pudessem vir a influenciar de modo negativo a escolha do número ideal de neurônios. A Figura 14 mostra o número do poço e a distribuição

espacial dos 10 poços sem falhas que foram selecionados para determinar o número de neurônios da camada oculta das RNAs.

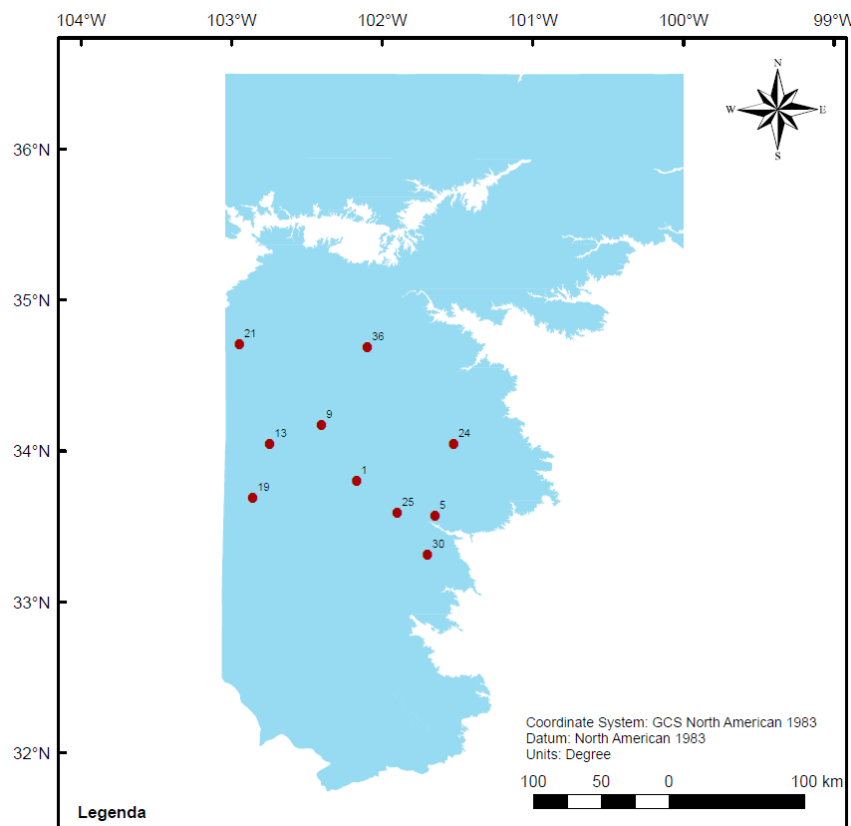


Figura 14 – Distribuição espacial dos 10 poços sem falhas.

Dessa maneira, para cada um dos 10 poços selecionados, foram feitas redes neurais testando de 1 a 36 neurônios na camada oculta, tendo sido cada número de neurônios testado 600 vezes, resultando num total de 216.000 testes. A Figura 15 mostra um esquema explicativo dos testes que foram feitos para determinar a quantidade de neurônios na camada oculta. Tomando como exemplo um poço qualquer, entre os 10 escolhidos, a Figura 15 mostra a estrutura de uma RNA com um nó na camada de entrada que representa o nível da água (NA) no ano t , um neurônio na camada oculta e um nó na camada de saída que equivale ao nível da água no ano $t + 1$, sendo o resultado previsto pela RNA. O número de neurônios na camada oculta é então testado progressivamente até 36 neurônios.

O número de 27 neurônios na camada oculta, foi o que exibiu, no geral, melhores índices estatísticos de qualidade de ajuste (Correlação, NSE, MAPE e RMSE) para a previsão do nível d'água e atendeu bem aos 10 poços escolhidos para o teste, portanto, 27 foi o número selecionado.

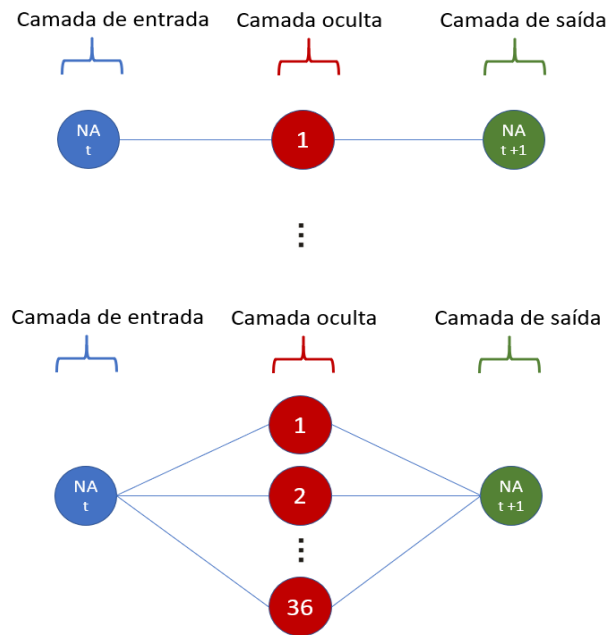


Figura 15 – Esquema explicativo dos testes realizados para a determinação do número de neurônios na camada oculta das RNAs.

Após ter fixado o número de neurônios na camada oculta, foram feitas várias combinações a fim de se definir a quantidade de anos precedentes usados pela rede neural para a previsão. Assim, para os 10 poços selecionados anteriormente, foram testados de um a doze anos precedentes (camada de entrada), para prever de um a dez anos à frente (camada de saída). O teste para cada ano precedente, em cada poço e para cada ano à frente, foi repetido 600 vezes, resultando num total de 720.000 testes.

A Figura 16 mostra um esquema simplificado dos testes que foram feitos para definir a quantidade de anos precedentes utilizados na camada de entrada da RNA. A Figura 16a mostra um nó na camada de entrada que representa o nível da água no ano t , correspondente a um ano precedente à previsão, 27 neurônios na camada oculta e 1 nó na camada de saída que é a previsão da RNA para o ano $t+1$. Então, a quantidade de anos previstos na camada de saída é incrementada progressivamente até dez anos à frente ($t+10$), Figura 16b. O mesmo procedimento é feito para dois anos precedentes na camada de entrada e assim por diante até doze anos precedentes, como mostrado na Figura 16c que mostra 12 nós na camada de entrada, correspondentes ao nível da água nos doze anos que antecedem a previsão para um ano à frente ($t+1$) e progressivamente até dez anos à frente ($t+10$) (Figura 16d).

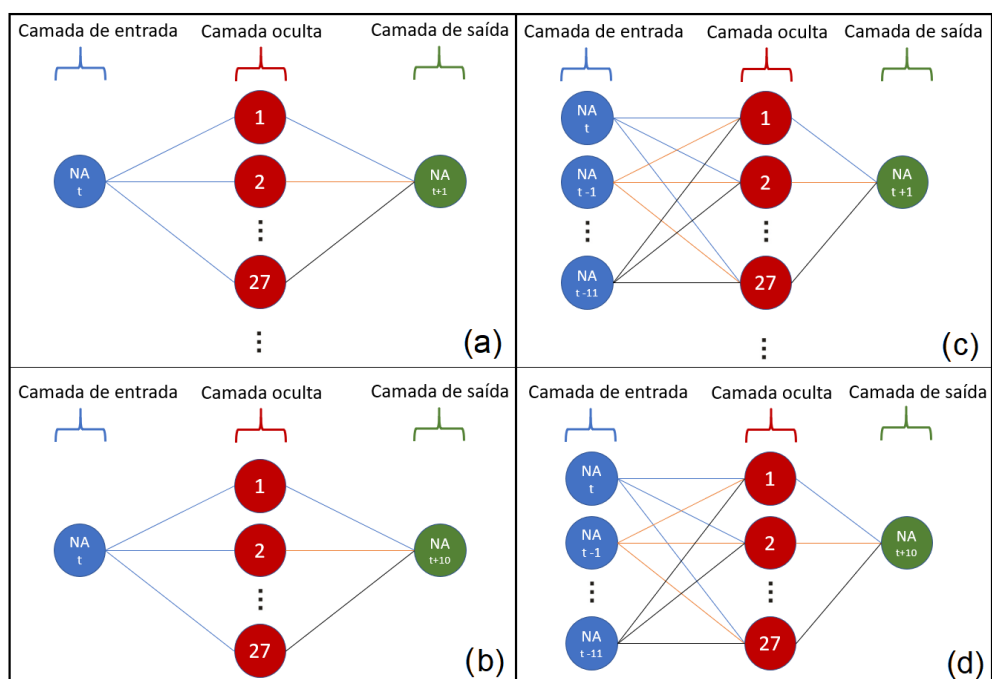


Figura 16 – Resumo dos testes realizados para a decisão da quantidade de anos precedentes a serem utilizados para a previsão com a RNA.

Feito isso, foram feitas análises dos resultados dos testes e os índices estatísticos foram avaliados. Desse modo, percebeu-se que a quantidade de anos precedentes que teve melhor desempenho, no geral, para prever de um a dez anos à frente, foram doze anos. Porém, por ser um número alto, a rede neural só poderia começar as previsões a partir do décimo terceiro ano da série histórica, já que necessitaria de doze anos da série histórica para fazer as previsões. Por conta disso, optou-se por uma quantidade menor de anos precedentes, mas que continuasse tendo, no geral, um bom desempenho, assim, a quantidade de quatro anos precedentes foi escolhida.

A Figura 17 mostra a arquitetura final escolhida para a previsão dos níveis de água nos poços. Com 4 nós na camada de entrada (*Input*), 27 neurônios na camada oculta (*Hidden Layer*) e 1 neurônio na camada de saída (*Output Layer*). Pela Figura 17 também é possível observar que a função de ativação empregada na camada oculta é a tangente hiperbólica (TanH), já na camada de saída é a função linear.

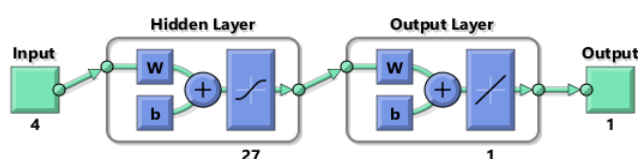


Figura 17 – Arquitetura final da RNA.

5.2.1 Critérios de avaliação das RNAs

De modo a avaliar a eficácia de cada rede neural gerada e a sua capacidade de fazer previsões precisas, foram utilizados quatro índices estatísticos. O primeiro deles foi o coeficiente de correlação linear de Pearson (R), já descrito anteriormente. Outro índice utilizado foi a eficiência de *Nash-Sutcliffe* (NSE) (Equação 8) (JEUNG et al., 2019):

$$NSE = 1 - \frac{\sum_{i=1}^n (E_i - O_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (8)$$

Onde, E_i e O_i são o valor estimado e o valor observado do poço i , respectivamente, n é o número de amostras e $\bar{O}$ é a média dos valores observados. O valor do NSE pode variar entre $-\infty$ e 1, com o $NSE = 1$, descrevendo um ajuste perfeito entre os valores estimados e os dados observados. Valores entre 0 e 1 são geralmente vistos como níveis aceitáveis de desempenho, enquanto valores menores que 0 indicam que o valor médio observado é um melhor preditor do que o modelo (MELAKU; WANG, 2019).

O desempenho das redes neurais também foi avaliado com relação ao erro de porcentagem absoluta média (MAPE) (Equação 9) e pelo cálculo da raiz do erro quadrático médio (RMSE) (Equação 10).

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|O_i - E_i|}{O_i} * 100 \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - E_i)^2} \quad (10)$$

Tanto o RMSE quanto o MAPE determinam a precisão do modelo comparando o desvio entre a previsão e os dados de medição. Quanto menor o RMSE e o MAPE, menor o erro entre os dados estimados e observados (ZHANG et al., 2019).

Além disso, também foi realizado o cálculo do *Bias* com a Equação 11 (DELNAZ; RAKHSHANDEHROO; NIKOO, 2020). O *Bias* indica um desvio positivo ou negativo do valor estimado em relação ao valor observado (MOJID; HOSSAIN; ASHRAF, 2019).

$$Bias = \frac{1}{n} \sum_{i=1}^n Oi - Ei \quad (11)$$

5.3 Desenvolvimento de RNAs para cada poço em específico

Após a arquitetura da RNA ter sido escolhida, foram geradas RNAs específicas para todos os 403 poços. Durante a aprendizagem da RNA de cada poço, foram utilizados os dados da série temporal de nível d'água específica do poço em questão. Para cada poço e para cada quantidade de anos previstos à frente (de um a dez anos) foram feitas 200 RNAs de modo a escolher a RNA que apresentou os melhores índices estatísticos (Correlação, NSE, MAPE, RMSE).

Por exemplo, para o poço 1, prevendo-se um ano à frente, foram feitas 200 RNAs e escolhida aquela que obteve melhor performance e assim por diante, para todos os poços. Após isso, a quantidade de anos previstos à frente era incrementada em um, até atingir o máximo escolhido de dez anos, e repetia-se o processo. Desse modo, para cada poço foram feitas 2000 RNAs, para previsão de um a dez anos à frente, resultando num total de 806000 RNAs criadas para todos os poços, dessas, 403 foram selecionadas para cada ano previsto à frente.

5.4 Aplicação das RNAs treinadas com os dados específicos de um poço nos outros poços

Após a melhor RNA para cada poço ter sido escolhida, cada RNA desenvolvida especificamente para um poço foi aplicada nos dados dos outros 402 poços, para avaliar se a previsão do nível d'água de um poço qualquer poderia ser feita com as RNAs de outros poços pertencentes ao Aquífero, a fim de se identificar padrões e poços compatíveis uns com os outros. Por exemplo, a RNA que foi desenvolvida especificamente para o poço 1, foi usada para prever os níveis d'água subterrânea dos outros 402 poços restantes do Aquífero Ogallala e assim por diante. Poços que obtiveram boas previsões, através da análise dos índices estatísticos, ao utilizar a RNA de outro poço, foram considerados compatíveis.

5.5 Análise de *clusters*

A análise de *cluster* é um método estatístico, que agrupa indivíduos com base na semelhança ou dissimilaridade associada às variáveis (PATHAK; DODAMANI, 2019). Um tipo especial de análise de *cluster* é o agrupamento de séries temporais (AGHABOZORGI;

SEYED SHIRKHORSHIDI; YING WAH, 2015). Séries temporais são definidas como uma coleção ordenada de observações ou sequência de pontos de dados feitas ao longo do tempo em intervalos de tempo frequentemente uniformes (ALI et al., 2019).

O agrupamento de dados de séries temporais é particularmente vantajoso e é utilizado principalmente para a descoberta de padrões interessantes que permitem aos analistas de dados extrair informações valiosas dos conjuntos de dados dessas séries (AGHABOZORGI; SEYED SHIRKHORSHIDI; YING WAH, 2015).

Em alguns estudos feitos com o intuito de avaliar os parâmetros das águas subterrâneas, as análises das séries temporais podem ser realizadas usando valores individuais dos poços de observação ou valores dos poços de observação das águas subterrâneas agrupados de acordo com algum critério (MIRZAVAND; GHAZAVI, 2014).

Os níveis de água subterrânea em uma região podem variar fortemente de um poço para outro (PATHAK; DODAMANI, 2019). Desse modo, a análise de *clusters* foi empregada no presente estudo para classificar grupos de poços de acordo com suas semelhanças na flutuação no nível da água.

Entre os vários métodos de análise de *cluster*, o agrupamento hierárquico é um método básico popularmente aplicado aos estudos de águas subterrâneas (PATHAK; DODAMANI, 2019). Bloomfield et al. (2015) utilizaram o agrupamento hierárquico usando como métrica de dissimilaridade, entre os hidrogramas do nível da água subterrânea, a matriz de correlação de suas séries temporais e como método de agrupamento o *complete linkage*.

Subba Rao e Chaudhary (2019) também fizeram uso da análise hierárquica de *clusters*, dispondo da distância euclidiana como métrica de dissimilaridade e do *complete linkage* como método de agrupamento para agrupar amostras de águas subterrâneas de modo a obter informações sobre processos hidrogeoquímicos que regulam a distribuição espacial da contaminação das águas subterrâneas.

5.5.1 Escolha da métrica e do método

Devido à grande variedade de métodos de agrupamentos para a análise de *cluster*, é difícil decidir sobre a melhor abordagem para fazer o agrupamento de um conjunto específico de dados. Alguns autores justificam a escolha de um determinado método de agrupamento alegando que o método escolhido produz agrupamentos mais interpretáveis fisicamente (BLOOMFIELD et al., 2015).

Hannah et al. (2000) justificaram sua escolha da análise hierárquica de *clusters*, utilizando o método de agrupamento *average linkage*, por essa combinação produzir

agrupamentos fisicamente mais interpretáveis do que outras alternativas a exemplo do método de agrupamento centróide e do *Ward's minimum variance*.

A análise de agrupamento hierárquico foi realizada para as séries temporais dos 403 poços do Aquífero Ogallala. Como explicado na seção 5.1 da metodologia, as séries temporais dos poços possuem tamanhos diferentes no que diz respeito à quantidade de anos que possuem dados observados. Para igualar o tamanho das séries temporais, de modo a possibilitar a análise de *clusters*, elas foram reduzidas para um tamanho em que todos os poços possuíssem dados de nível d'água. Desse modo, as séries temporais utilizadas para a análise de *cluster* ficaram com 44 anos, referente ao período de 1965 a 2008.

Após igualar o tamanho das séries temporais, foram testadas várias métricas de dissimilaridade combinadas com diferentes métodos de agrupamento, a fim de decidir qual seria a métrica e o método mais adequado. Para auxiliar nesta escolha, o coeficiente de correlação cofenética foi calculado para todas as combinações feitas.

O coeficiente de correlação cofenética compara a matriz de distância gerada pela métrica de dissimilaridade com as distâncias do método de agrupamento determinadas para a construção do dendrograma. O coeficiente de correlação pode variar de 0 a 1, valores próximos a um indicam alta correlação entre essas distâncias, o que sugere que o dendrograma resultante retrata com precisão as distâncias entre os parâmetros de entrada (KRAMER; SIEGEL, 2019).

Sendo d_{ij} a distância entre os objetos da série i e j com base na métrica de dissimilaridade escolhida e t_{ij} a distância dendrogramática, entre os dois objetos da série temporal i e j , que representa a altura do nó no qual esses dois pontos são unidos pela primeira vez. Assumindo que d'_{ij} seja a média de d_{ij} e t'_{ij} seja a média de t_{ij} , o coeficiente de correlação cofenética pode ser indicado pela Equação 12 (KUMAR; TOSHNIWAL, 2016):

$$c = \frac{\sum_{i < j} (d_{ij} - d'_{ij}) * (t_{ij} - t'_{ij})}{\sqrt{(\sum_{i < j} (d_{ij} - d'_{ij})^2) * (\sum_{i < j} (t_{ij} - t'_{ij})^2)}} \quad (12)$$

A Tabela 1 mostra o resultado do cálculo do coeficiente de correlação cofenética entre as distâncias dos métodos de agrupamento e das métricas de dissimilaridade testadas. Cores mais puxadas para o verde indicam uma correlação mais alta, enquanto cores mais puxadas para o vermelho indicam uma correlação menor.

Tabela 1 – Resultado do cálculo do coeficiente de correlação cofenética para as combinações de método e métrica testadas

Métrica Método	Average	Centroid	Complete	Median	Single	Ward	Weighted
<i>Euclidean</i>	0,71	0,71	0,70	0,66	0,38	0,69	0,67
<i>Squared euclidean</i>	0,57	0,56	0,58	0,55	0,31	0,55	0,54
<i>Seuclidean</i>	0,71	0,71	0,67	0,67	0,38	0,68	0,66
<i>Cityblock</i>	0,71	0,71	0,67	0,65	0,39	0,68	0,68
<i>Minkowski</i>	0,71	0,71	0,70	0,66	0,38	0,69	0,67
<i>Chebychev</i>	0,71	0,72	0,66	0,63	0,43	0,69	0,69
<i>Mahalanobis</i>	0,95	0,95	0,87	0,94	0,94	0,63	0,92
<i>Cosine</i>	0,76	0,75	0,46	0,69	0,31	0,59	0,59
<i>Correlation</i>	0,88	0,88	0,78	0,70	0,33	0,81	0,80
<i>Spearman</i>	0,89	0,84	0,79	0,81	0,42	0,79	0,73
<i>Hamming</i>	0,95	-0,16	0,91	0,05	0,84	0,06	0,95
<i>Jaccard</i>	0,95	-0,16	0,91	0,05	0,84	0,06	0,95

Ao analisar de maneira detalhada os resultados da Tabela 1 e da divisão dos *clusters* resultantes de cada combinação feita, percebeu-se que, embora a combinação de alguns métodos e métricas possuísem um coeficiente de correlação cofenética elevado, a divisão dos *clusters* não estava adequada em relação às flutuações no nível d'água e à sua distribuição espacial.

Para chegar a esta conclusão, o valor da variação máxima do nível d'água de cada uma das séries dos 403 poços foi calculado, Tabela 24 (encontra-se no Apêndice A). Por exemplo, para a série histórica do poço 1, a variação máxima do nível d'água é encontrada diminuindo o maior valor de nível d'água do menor valor encontrado na série, e assim por diante. Foi constatado que a maior variação na diferença de nível d'água de cada uma das séries temporais dos 403 poços é de 53,4 m, correspondente ao poço 344, que possui um valor mínimo de nível d'água de 1002,63 m, e um valor máximo observado de 1056 m.

Já a maior diferença na variação da elevação do nível d'água entre todos os poços do Aquífero Ogallala é de 589,6 m. Sendo 1339,57 m o maior valor registrado, correspondente ao poço 403, e 750 m o menor valor, referente ao poço 373. Com essas informações, é possível avaliar a coerência dos *clusters* resultantes das combinações de método e métrica. Pois, já que os *clusters* estão sendo avaliados de acordo com suas semelhanças na flutuação do nível da água, supõe-se que, para serem considerados adequados, cada *cluster* deve possuir poços que possuem níveis d'água semelhantes.

Dessa forma, fica claro que, na divisão do *clusters*, resultantes de uma combinação específica de método e métrica, ao analisar a variação máxima de nível d'água de cada *cluster* e selecionar a maior dentre todos, o menor valor que poderia ocorrer é de 53,4 m, já

que esse valor corresponde à maior variação de nível d'água existente na série temporal de um poço individual (poço 344).

Por outro lado, a maior variação de nível d'água que pode ocorrer em um *cluster* é de 589,6 m, que significa que os poços que possuem a maior diferença de nível d'água entre si, dentre todos os analisados do Aquífero Ogallala (poços 373 e 403), foram colocados juntos no mesmo *cluster*.

Assim, para os todos os *clusters* resultantes de cada combinação de métrica e método, a variação máxima no nível d'água foi calculada e a maior variação dentre todas é apresentada na Tabela 2.

Tabela 2 – Resultados do cálculo da variação máxima de nível d'água entre os poços (em metros) de todos os *clusters* resultantes da combinação de cada método e métrica.

Métrica Método	Average	Centroid	Complete	Median	Single	Ward	Weighted
<i>Euclidean</i>	64,4	68,5	68,5	68,5	157,3	57,5	68,5
<i>Squared euclidean</i>	78,9	68,5	68,5	68,5	157,3	68,5	68,5
<i>Seuclidean</i>	64,4	68,5	68,5	68,5	157,3	57,5	68,5
<i>Cityblock</i>	68,5	68,5	68,5	68,5	197,1	68,5	68,5
<i>Minkowski</i>	64,4	68,5	68,5	68,5	157,3	57,5	68,5
<i>Chebychev</i>	57,5	57,5	57,5	63,2	293,8	57,5	57,5
<i>Mahalanobis</i>	497,3	497,3	489,9	497,3	497,3	437,8	497,3
<i>Cosine</i>	505,7	505,7	505,7	505,7	589,6	505,7	505,7
<i>Correlation</i>	559,6	559,6	558,3	559,6	559,7	559,6	559,6
<i>Spearman</i>	559,6	559,6	558,3	559,6	559,7	558,3	558,3
<i>Hamming</i>	558,3	558,3	558,3	558,3	558,3	589,6	558,3
<i>Jaccard</i>	558,3	558,3	558,3	558,3	558,3	589,6	558,3

Por exemplo, utilizando o método *euclidean* e a métrica *average*, os *clusters* formados foram avaliados e a variação máxima do nível d'água de cada *cluster* foi calculada. A maior variação no nível d'água entre todos os *clusters* foi 64,4 m, o que demonstra que os poços que se encontram nos *clusters* formados por meio dessa combinação possuem níveis de água semelhantes, sendo 64,4 m o maior intervalo em que se encontram os poços de um *cluster* específico dessa combinação.

Em contrapartida, a variação máxima do nível d'água dentre todos os *clusters* resultantes da combinação do método *cosine* com a métrica *single* é de 589,6 m, isso indica que os poços que possuem a maior diferença de flutuação de nível d'água foram postos no mesmo *cluster*. O que demonstra que a divisão dos *clusters* com base na semelhança da flutuação do nível d'água não foi adequada.

Assim, na Tabela 2, as cores mais puxadas para o verde indicam as combinações de método e métrica que criaram *clusters* que possuem poços mais homogêneos em relação ao nível d'água, justificado por um menor valor de variação máxima do nível d'água entre os poços. Já as cores mais puxadas para o vermelho, mostram a combinação de métodos e métricas que formaram *clusters* com poços que possuem maiores diferenças na flutuação do nível d'água, sendo, portanto, mais heterogêneos.

Além disso, foram feitas outras análises a respeito da distribuição espacial dos *clusters* formados pelas combinações de método e métrica. Unindo tais informações com as informações das Tabelas 1 e 2, concluiu-se que a combinação do método *euclidean* com a métrica *average* seria adequada para este trabalho. Tendo em vista que, essa combinação possui um coeficiente de correlação cofenética de 0,71 (considerado adequado); a maior variação máxima de nível d'água dentre os *clusters* é de 64,4 m, o que demonstra que os *clusters* formados possuem uma boa similaridade em relação às flutuações no nível d'água, quando comparado às outras opções; e, além disso, os *clusters* formados por essa combinação são fisicamente bem interpretáveis.

Portanto, neste estudo, para a análise de *cluster*, o agrupamento hierárquico foi escolhido, usando a distância euclidiana (*euclidean distance*) como métrica de dissimilaridade e a ligação média (*average linkage*) como método de agrupamento, para identificar os poços homogêneos com base nas séries temporais do nível das águas subterrâneas.

A distância euclidiana e o método *average linkage* foram calculados no MATLAB R2017a utilizando as funções “pdist” e “linkage”, respectivamente. A distância euclidiana é uma métrica comumente usada para séries temporais e é uma das mais utilizadas na análise hierárquica de agrupamentos. Entretanto, para ser empregada, as duas séries temporais que estão sendo comparadas devem ter o mesmo comprimento (ALI et al., 2019; MACHIWAL et al., 2018). Ela é computada como a distância em linha reta entre dois pontos no espaço euclidiano. Seja x_i e v_j dois vetores de tamanho P , a distância euclidiana é calculada pela Equação 13 (HE; AGARD; TRÉPANIER, 2020).

$$d_E = \sqrt{\sum_{k=1}^P (x_i - v_j)^2} \quad (13)$$

O método *average linkage* (Equação 14) usa a distância média entre todos os pares de objetos de dois *clusters*. Na Equação 14, $d(r, s)$ representa a distância entre os *clusters* r e s , e n é o número de elementos em cada *cluster* (SANTOS et al., 2019).

$$d(r, s) = \frac{1}{n_r n_s} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} d(x_{ri}, x_{sj}) \quad (14)$$

No agrupamento hierárquico, a matriz de distância, baseada nos níveis das águas subterrâneas, é calculada primeiro, utilizando a distância euclidiana, e cada poço é considerado como um grupo separado. Posteriormente, cada grupo é mesclado aos grupos mais próximos, com base no método *average linkage* (PATHAK; DODAMANI, 2019).

5.5.2 Escolha do número de *clusters*

A seleção do número ideal de *clusters* é complexa e sem critério único disponível para tomar uma decisão objetiva (MCGREE; SCHREIDER; KULESHOV, 2016). O número de *clusters* resultantes é uma escolha do usuário. Para obter *clusters* diferentes, o dendrograma deve ser cortado em um nível específico da distância de ligação (*linkage distance*). A quantidade de *clusters* resultantes da linha de corte pode ser especificada por inspeção visual do dendrograma. Assim, a seleção da quantidade de *clusters* é subjetiva e varia de usuário para usuário (FABBROCINO et al., 2019; PATHAK; DODAMANI, 2019).

Torok (1996) sugere que uma escolha subjetiva de cerca de 10% do número total de elementos no conjunto de dados é considerada um número razoável de *clusters*. Para esse caso, foram escolhidos trinta *clusters*, pois essa quantidade foi considerada suficiente para agrupar os 403 poços do Aquífero Ogallala, por não ser um número pequeno, nem alto demais de *clusters*. Desse modo, trinta *clusters* foi considerado satisfatório para a análise de *clusters*.

5.6 Desenvolvimento de uma RNA por *cluster*, treinada com os dados de todos os poços do *cluster*

A partir da análise hierárquica de agrupamentos, os poços que estão sendo estudados do Aquífero Ogallala foram divididos em trinta *clusters*. Desse modo, foram criadas 200 RNAs para cada um dos trinta *clusters* com o intuito de selecionar a RNA que possuísse melhores índices estatísticos. Dentre as 200 RNAs foi escolhida 1 RNA por *cluster*, resultando em um total de trinta RNAs que seriam utilizadas para prever o nível d'água dos 403 poços. As RNAs, criadas para cada *cluster*, utilizaram os dados de todos os poços presentes no *cluster* durante a aprendizagem.

A arquitetura das RNAs feitas para este caso é semelhante ao do caso de uma RNA específica por poço. Sendo utilizados quatro anos precedentes (camada de entrada), 27 neurônios na camada oculta para prever um ano à frente (camada de saída). Após a aprendizagem, as RNAs selecionadas foram testadas com a série temporal de cada poço individualmente a fim de avaliar sua performance.

5.7 Desenvolvimento de uma RNA por *cluster*, treinada com os dados de alguns poços do *cluster*

Após criar uma RNA por *cluster*, treinada com os dados de todos os poços presentes no *cluster*, um novo teste foi feito. Dessa vez, para a etapa de aprendizagem das RNAs de cada *cluster*, foram utilizados os dados de 80% dos poços presentes no *cluster*. Por exemplo, se o *cluster* 1 possui 25 poços, foram utilizados os dados de apenas 20 poços na etapa de aprendizagem da RNA. Isso foi feito com o intuito de saber se uma RNA treinada dessa forma seria capaz de prever o nível da água subterrânea para todos os poços presentes no *cluster*, ou apenas para os poços cujos dados foram utilizados durante a aprendizagem.

Assim, foram feitas 200 RNAs para cada *cluster* e aquela que obteve melhores índices estatísticos foi selecionada, resultando em um total de 30 RNAs. Após isso, cada uma das 30 RNAs foi testada utilizando os dados da série temporal de um poço, pertencente ao seu *cluster*, por vez, de modo a avaliar sua performance para a previsão do nível d'água de cada poço especificamente.

5.8 Desenvolvimento de uma única RNA para todos os poços

Para este caso, o intuito foi desenvolver uma RNA capaz de prever o nível d'água de qualquer um dos 403 poços analisados no Aquífero Ogallala. Para isso, foram criadas 200 RNAs, utilizando os dados de nível d'água dos 403 poços estudados para a aprendizagem das redes, com o objetivo de escolher a que obtivesse melhores índices estatísticos.

A arquitetura dessa RNA é semelhante às das RNAs criadas para cada poço em específico. Esta RNA possui quatro neurônios da camada de entrada, representando quatro anos precedentes de dados de nível d'água, 27 neurônios na camada escondida e um neurônio na camada de saída, sendo a previsão do nível d'água, no poço desejado, para um ano à frente.

6 RESULTADOS E DISCUSSÕES

6.1 Resultados das previsões dos níveis d'água em cada poço utilizando RNAs treinadas com os dados específicos do poço analisado

Para cada poço estudado no Aquífero Ogallala foi desenvolvida uma RNA específica. Para fazer as RNAs utilizadas neste caso, os dados de níveis da água subterrânea de quatro anos antecedentes foram usados como variáveis de entrada. A saída do modelo da RNA é o nível da água subterrânea no ano desejado (de 1 a 10 anos à frente).

Na Tabela 3 é apresentada uma parte dos resultados obtidos pelos índices estatísticos (correlação, NSE, MAPE e RMSE) que foram utilizados para avaliar o desempenho das RNAs na previsão do nível d'água de cada poço especificamente. Assim, são exibidos os valores dos índices estatísticos da RNA que foi selecionada para cada poço, por ter apresentado o melhor desempenho, para um horizonte de previsão de um e dois anos à frente.

Tabela 3 – Índices estatísticos da RNA escolhida para cada poço, para previsão de um e dois anos

Poço	Previsão para um ano à frente				Previsão para dois anos à frente			
	R	NSE	MAPE	RMSE	R	NSE	MAPE	RMSE
1	0,9564	0,9065	0,00020	0,0415	0,9001	0,7752	0,00028	0,0259
2	0,9981	0,9961	0,00011	0,0184	0,9978	0,9949	0,00013	0,0544
3	0,9987	0,9973	0,00024	0,0525	0,9978	0,9953	0,00022	0,0739
4	0,9861	0,9721	0,00031	0,0390	0,9848	0,9699	0,00024	0,0142
5	0,9883	0,9760	0,00019	0,0526	0,9793	0,9573	0,00031	0,0850
6	0,9936	0,9870	0,00012	0,0171	0,9833	0,9665	0,00013	0,0097
7	0,9818	0,9542	0,00025	0,1386	0,9737	0,9472	0,00026	0,0467
8	0,9836	0,9666	0,00025	0,0402	0,9735	0,9476	0,00035	0,0011
9	1,0000	0,9999	0,00007	0,0204	1,0000	0,9999	0,00005	0,0127
10	0,9981	0,9962	0,00018	0,0097	0,9972	0,9943	0,00019	0,0062
11	0,9592	0,9133	0,00034	0,0665	0,9374	0,8778	0,00037	0,0332
12	0,9810	0,9606	0,00026	0,0072	0,9698	0,9401	0,00031	0,0048
13	0,9487	0,8993	0,00017	0,0235	0,9322	0,8686	0,00015	0,0111
14	0,9992	0,9984	0,00017	0,0345	0,9989	0,9977	0,00023	0,0844
15	0,8613	0,7219	0,00015	0,0322	0,8556	0,7069	0,00014	0,0431
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
398	0,9980	0,9954	0,00038	0,1330	0,9942	0,9858	0,00078	0,2262
399	0,9857	0,9649	0,00091	0,3398	0,9640	0,9283	0,00135	0,0757
400	0,9979	0,9954	0,00038	0,0683	0,9960	0,9918	0,00077	0,0531
401	0,9968	0,9934	0,00033	0,0663	0,9920	0,9840	0,00053	0,0446
402	0,9808	0,9522	0,00050	0,2649	0,9684	0,9344	0,00049	0,0632
403	0,9927	0,9844	0,00026	0,0831	0,9949	0,9875	0,00036	0,2699

As Figuras 18 a 27 dão uma visão geral e mostram os valores de todos os índices estatísticos obtidos pelas 403 RNAs selecionadas para cada poço, sendo cada figura correspondente a um ano diferente no horizonte de previsão (de 1 a 10 anos à frente). No eixo das abscissas de cada figura é informado o número do poço e no eixo das ordenadas, o índice estatístico analisado (correlação, NSE, MAPE, RMSE e Bias). Parte dos resultados das Figuras 18 e 19 já foram apresentados na Tabela 3.

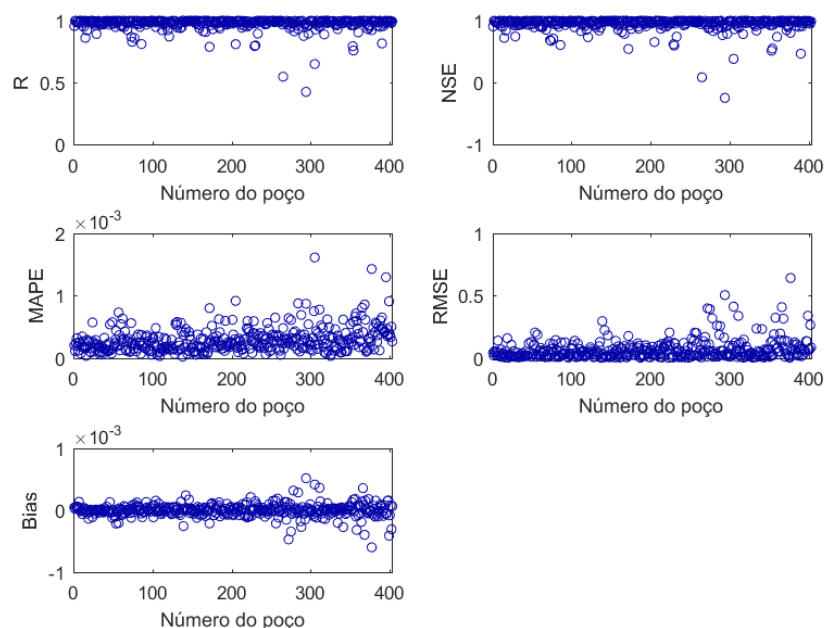


Figura 18 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de um ano à frente.

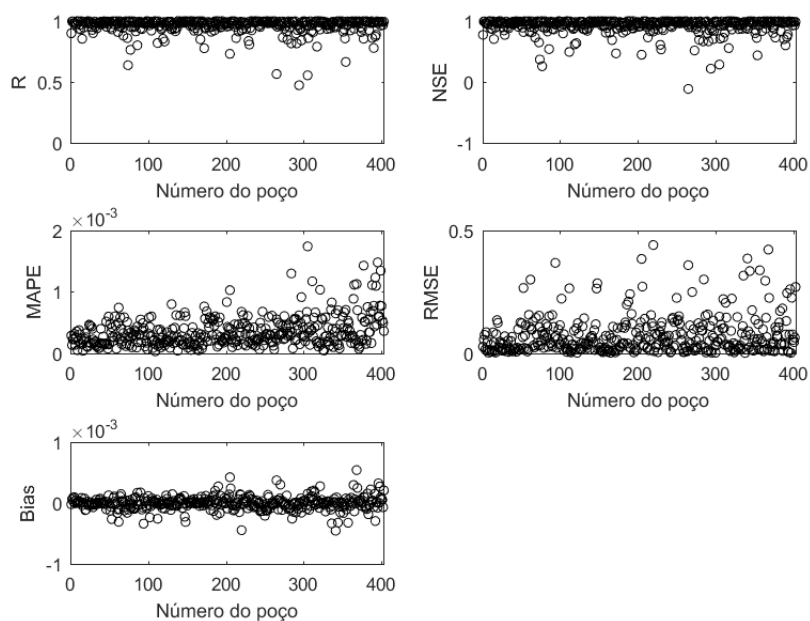


Figura 19 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de dois anos à frente.

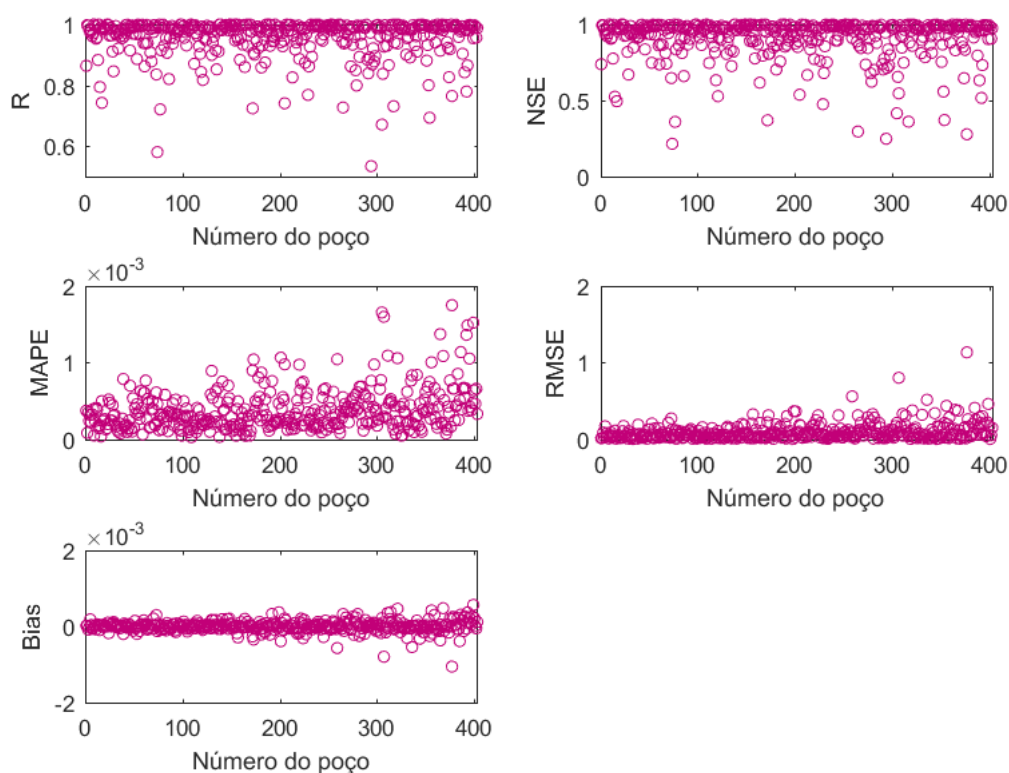


Figura 20 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de três anos à frente.

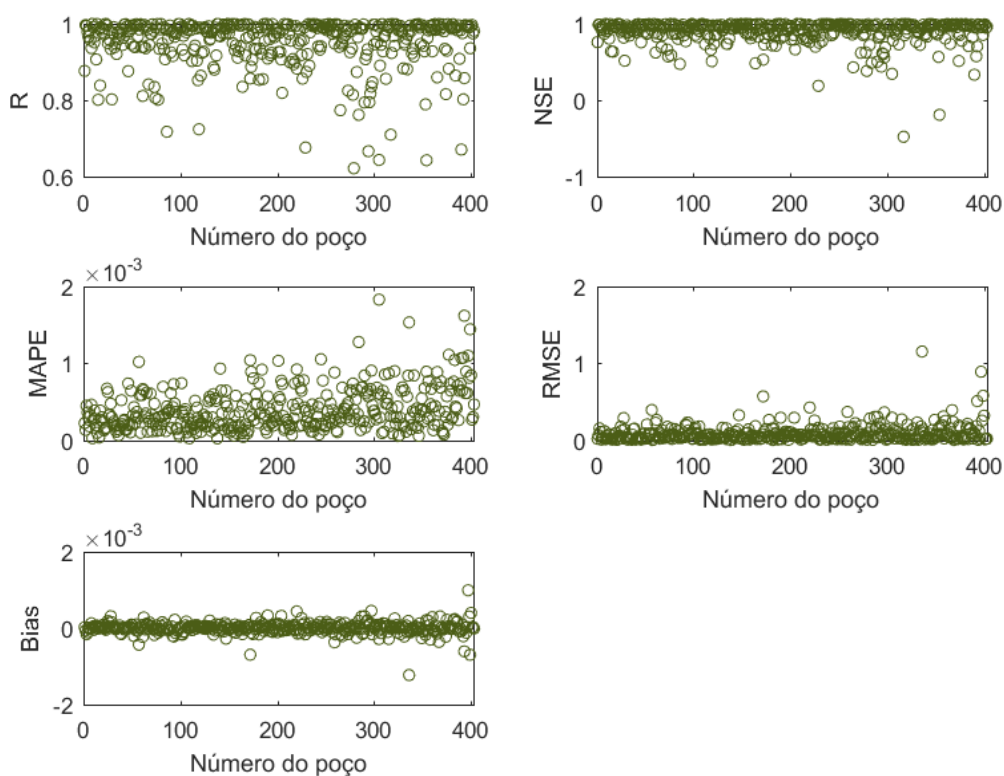


Figura 21 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de quatro anos à frente.

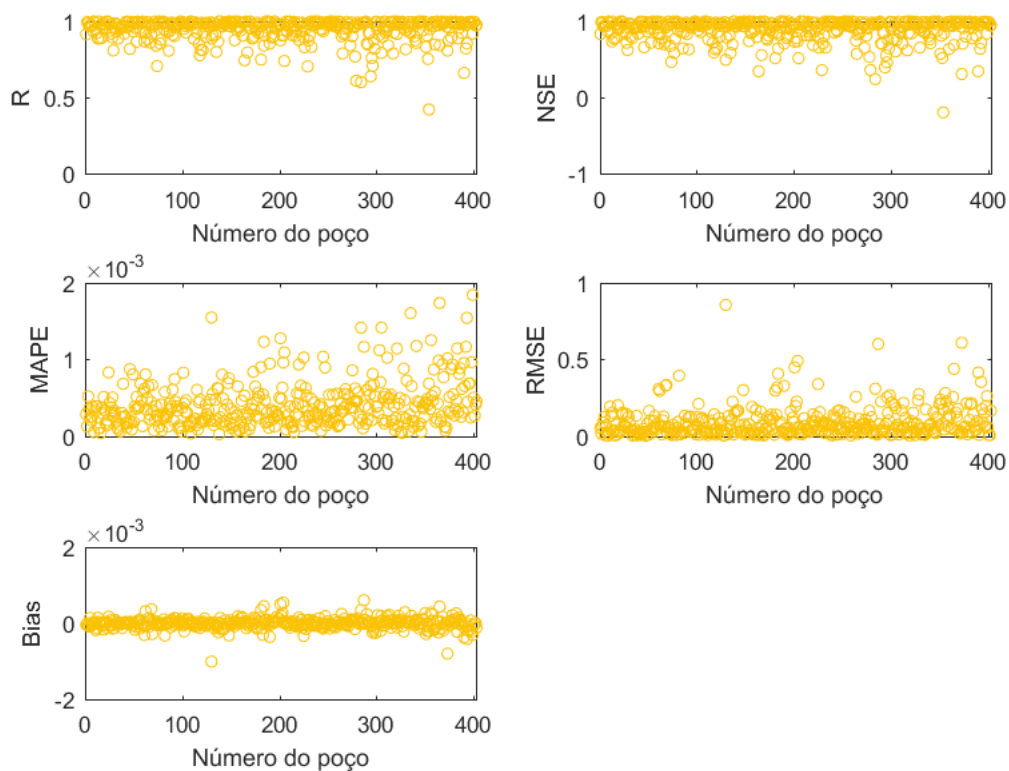


Figura 22 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de cinco anos à frente.

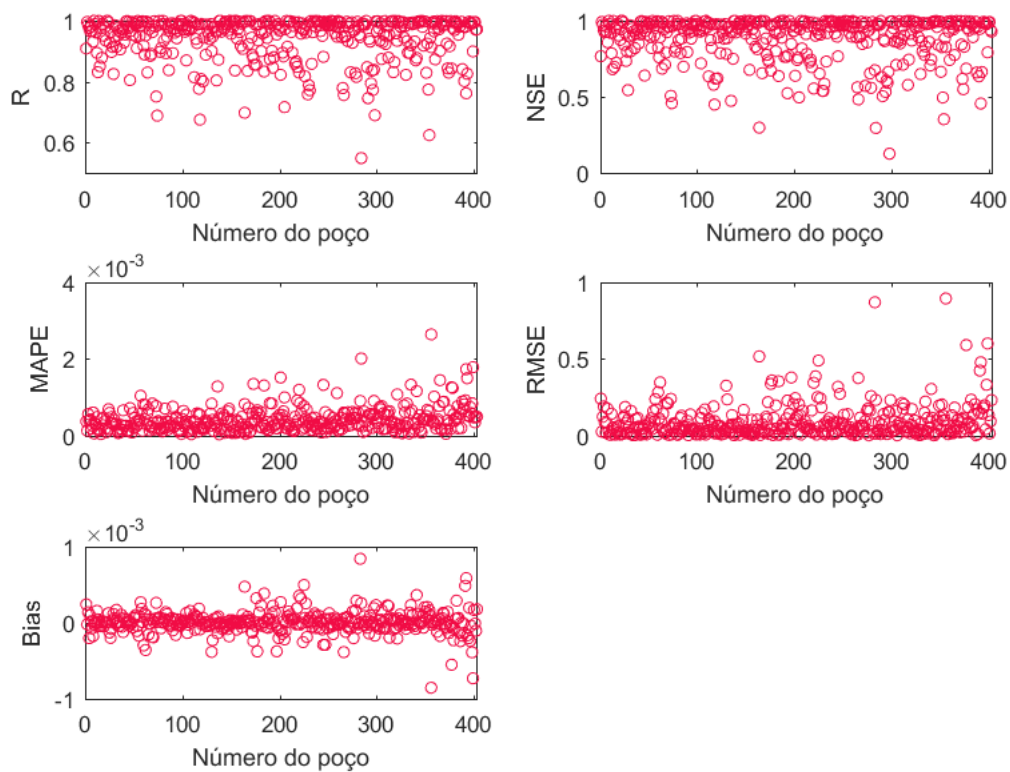


Figura 23 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de seis anos à frente.

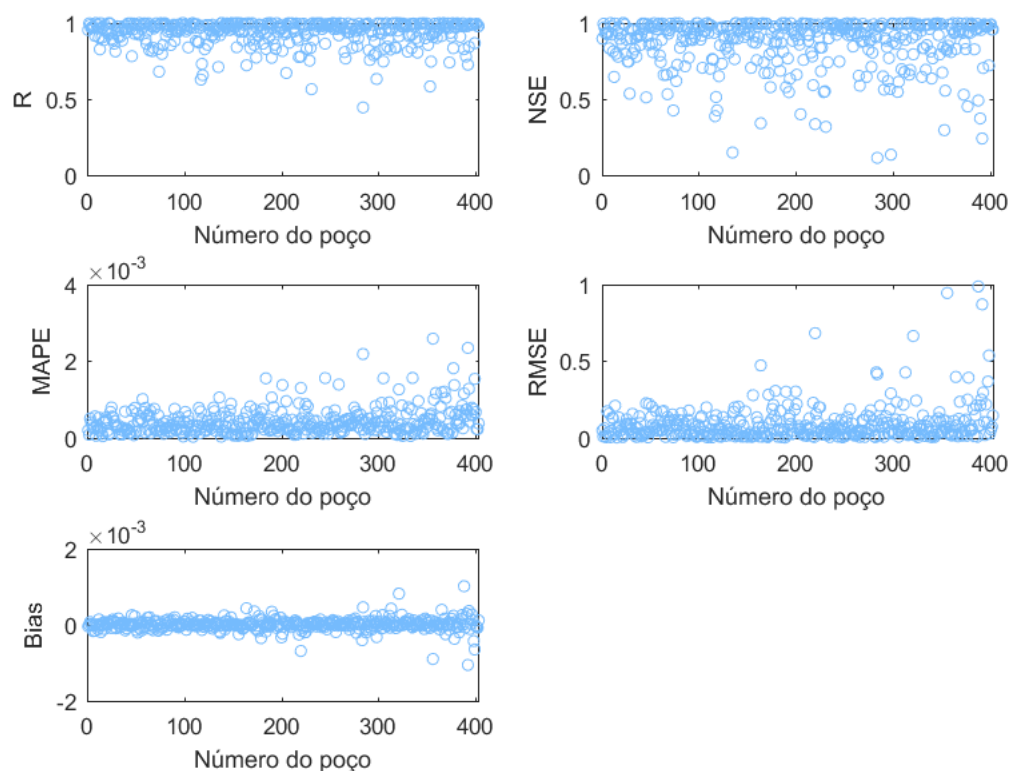


Figura 24 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de sete anos à frente.

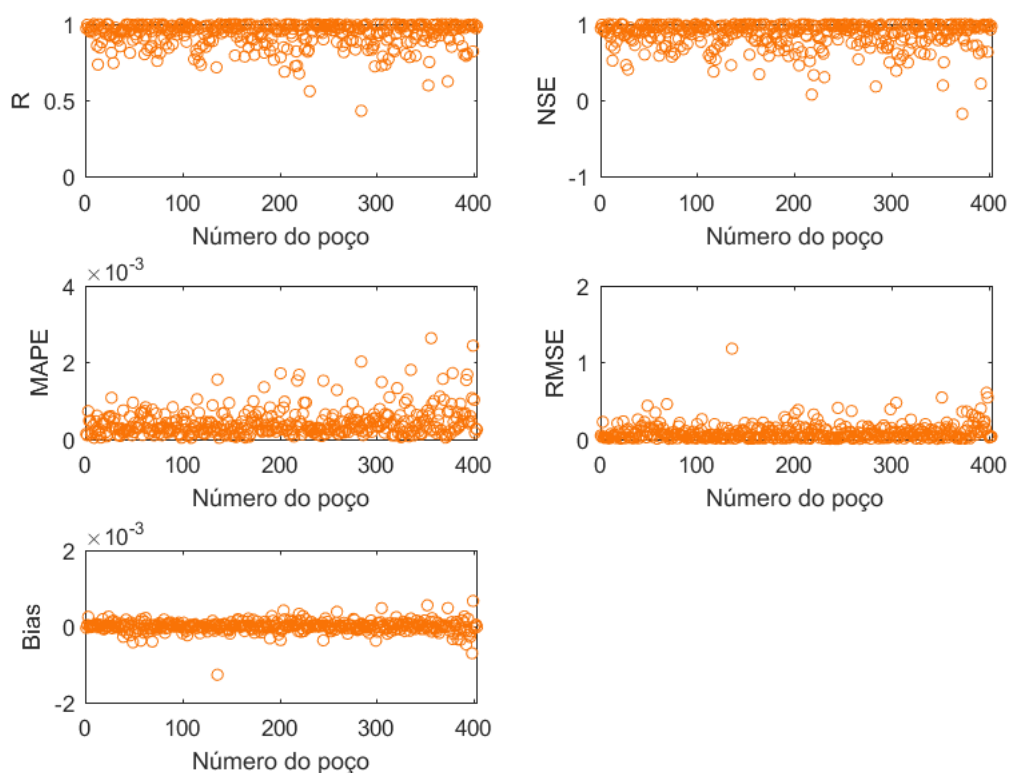


Figura 25 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de oito anos à frente.

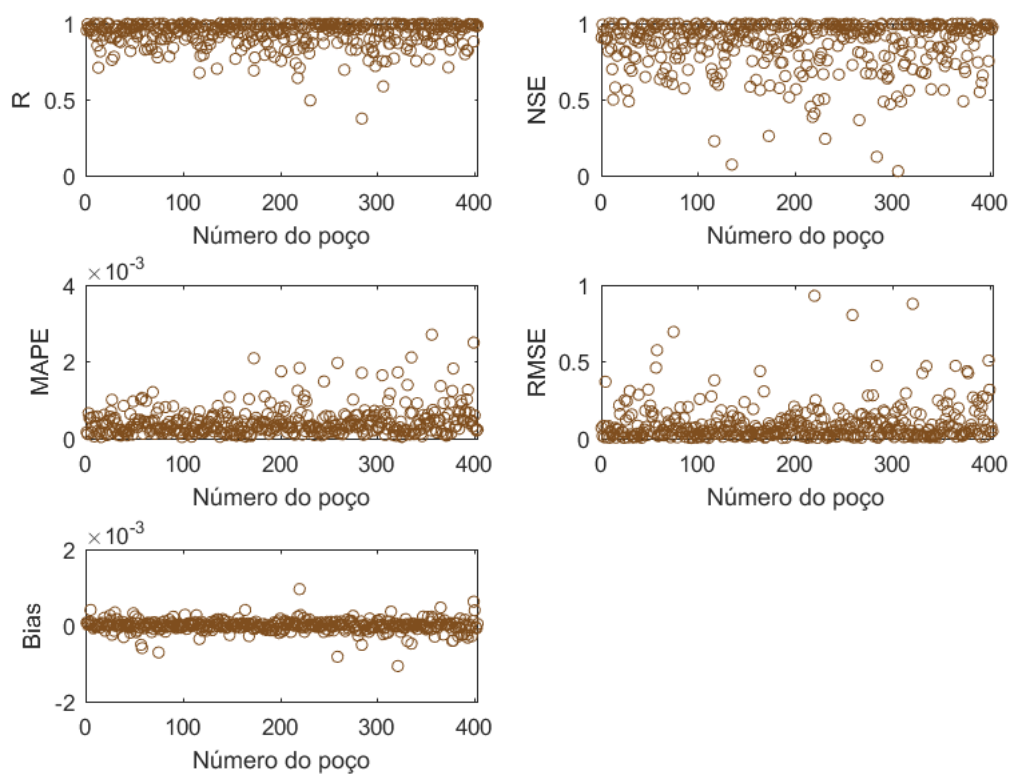


Figura 26 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de nove anos à frente.

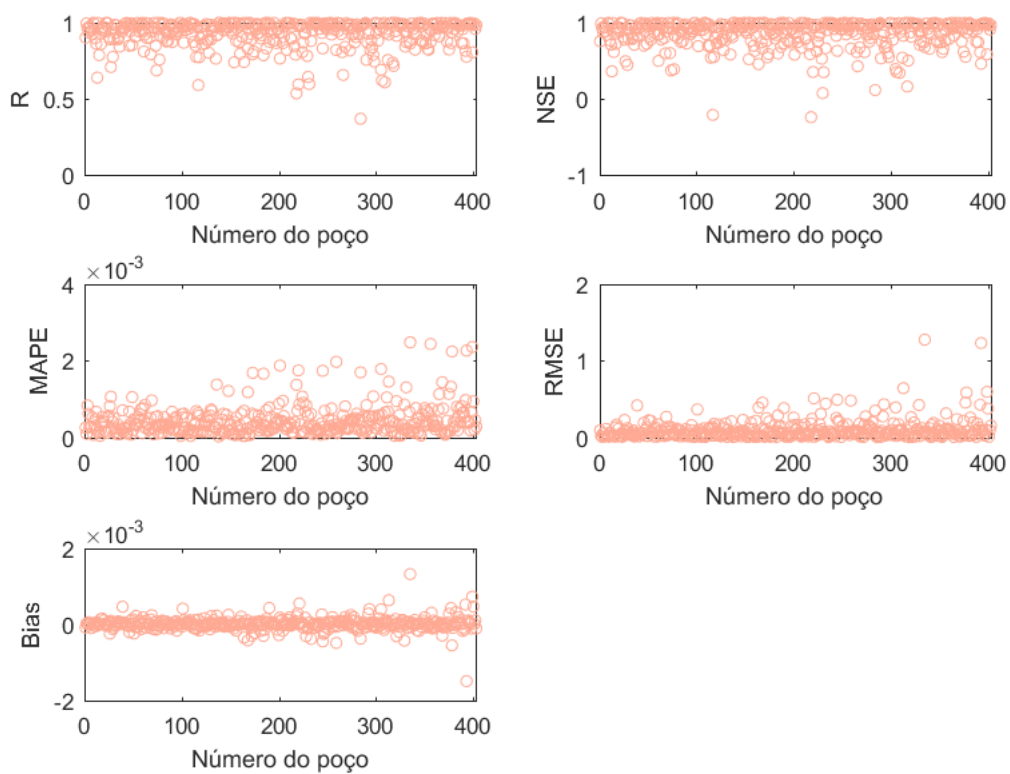


Figura 27 – Índices estatísticos da RNA escolhida para cada um dos 403 poços, para previsão de dez anos à frente.

Para a previsão de um ano à frente (Figura 18), pode-se dizer que a grande maioria das RNAs de cada poço fornecem uma previsão relativamente alta e precisa dos níveis das águas subterrâneas do Aquífero Ogallala. Com exceção dos resultados dos poços 265, 294 e 305, todos os demais possuem coeficiente de correlação superior a 0,7. Em relação ao NSE, apenas 13 poços apresentam RNAs com valores menores do que 0,7. Os maiores valores do MAPE e do RMSE obtidos pelas RNAs, e, portanto, piores resultados encontrados, foram 0,0016 para a RNA do poço 305 e 0,64 para a RNA do poço 377, nesta ordem.

Na Figura 19, para a previsão de dois anos à frente, o número de poços cujas RNAs apresentaram coeficiente de correlação inferior a 0,7 sobe de 3 para 5, sendo eles os poços 74, 265, 294, 305 e 354. Já para o NSE, 22 modelos de RNAs detêm de um valor menor do que 0,7. O MAPE e o RMSE possuem valores máximos de 0,0017 e 0,44 nas RNAs dos poços 305 e 220, respectivamente.

Na previsão de três anos à frente (Figura 20), os poços 74, 294, 305 e 354 possuem RNAs com coeficiente de correlação abaixo de 0,7. Além disso, há 29 poços nos quais as RNAs obtiveram valores do NSE menores do que 0,7. No tocante ao MAPE e ao RMSE, os maiores valores encontrados foram 0,0018 e 1,14 na RNA do poço 377.

Para a previsão de quatro anos à frente (Figura 21), o número de poços com modelos de RNA que apresentaram coeficiente de correlação e NSE inferiores a 0,7 foram 6 (poços 229, 279, 294, 305, 354 e 390) e 33 poços, respectivamente. No que se refere ao MAPE e ao RMSE, os valores mais altos obtidos foram 0,0018 na RNA do poço 305 e 1,16 na RNA do poço 336.

A Figura 22 mostra os índices estatísticos resultantes das RNAs para a previsão de cinco anos à frente e exibe 5 poços cujas RNAs tiveram coeficiente de correlação menor do que 0,7 (poços 279, 284, 294, 354, e 390) e 43 poços com NSE inferior a 0,7. Além do mais, a RNA que possui o MAPE mais elevado foi a do poço 399, com o valor de 0,0018, já o maior RMSE encontrado foi 0,86, na RNA do poço 130.

Na previsão de seis anos à frente (Figura 23), 6 poços (74, 118, 164, 284, 298 e 354) apresentaram RNAs com coeficiente de correlação abaixo de 0,7 e 54 poços tiveram NSE abaixo desse mesmo valor. Os piores resultados obtidos para o MAPE e RMSE foram 0,0026 e 0,90, nesta ordem, ambos na RNA do poço 356.

Já para a previsão de sete anos à frente (Figura 24), 8 (poços 74, 117, 119, 205, 231, 284, 298 e 353) e 56 poços obtiveram RNAs cujo coeficiente de correlação e o NSE encontram-se abaixo de 0,7, respectivamente. As RNAs que tiveram menor desempenho

quanto ao MAPE e ao RMSE, na devida ordem, foram as dos poços 356, com MAPE de 0,0026, e a do poço 388, com RMSE de 0,99.

A Figura 25 exhibe os resultados dos índices estatísticos provenientes das RNAs para a previsão de oito anos à frente. Com exceção dos resultados dos poços 205, 220, 231, 284, 353 e 373 todos os demais possuem coeficiente de correlação da RNA superior a 0,7. No que se refere ao NSE, 61 poços detêm de valores menores do que 0,7. A RNA cujo MAPE exibiu menor performance foi a do poço 356, com valor de 0,0026. Quanto ao RMSE, a RNA com maior valor foi a do poço 136, com um resultado de 1,18.

Já para a previsão de nove anos à frente (Figura 26), 7 poços (117, 173, 218, 231, 266, 284 e 306) apresentaram RNAs com coeficiente de correlação inferior a 0,7 e 64 poços possuem NSE abaixo de 0,7. As RNAs que obtiveram menores desempenho em relação ao MAPE e ao RMSE, de modo respectivo, foram as dos poços 356, com MAPE de 0,0027, e 220, com RMSE de 0,93.

Na previsão de dez anos à frente (Figura 27), 12 poços (13, 74, 117, 218, 220, 230, 231, 266, 284, 305, 306 e 309) têm coeficiente de correlação inferior a 0,7 e 62 poços dispõem de NSE abaixo de 0,7. A RNA do poço 335 foi a que apresentou os maiores valores tanto para o MAPE, com resultado de 0,0025, quanto para o RMSE, que teve um valor de 1,28.

No geral, percebe-se que os resultados das previsões foram satisfatórios, obtendo, em sua grande maioria, coeficiente de correlação e NSE próximos a 1, e MAPE e RMSE próximos a zero. É possível verificar que com o passar dos anos de previsão as RNAs passam a obter uma performance menor. Na previsão de um ano à frente, há 13 poços, o que equivale a 3,23% do total, que possuem NSE abaixo de 0,7 e a cada ano essa porcentagem aumenta gradativamente. Na previsão de nove anos à frente, a quantidade de poços com NSE inferior a 0,7 atinge seu valor máximo, com 64 poços, o equivalente a 15,9% do total. A Tabela 4 mostra a quantidade de poços cuja RNA obteve valores de NSE abaixo de 0,7, para cada ano de previsão à frente, e quanto isso representa em porcentagens totais.

Tabela 4 – Quantidade de poços com valores de NSE abaixo de 0,7 por ano de previsão

Ano previsto à frente	Quantidade de poços	(%) do total	Ano previsto à frente	Quantidade de poços	(%) do total
1	13	3,2	6	54	13,4
2	22	5,5	7	56	13,9
3	29	7,2	8	61	15,1
4	33	8,2	9	64	15,9
5	43	10,7	10	62	15,4

Na Figura 28, é possível observar a variação espacial e temporal dos resultados obtidos pelo cálculo da correlação (R) entre os dados observados e previstos pelas RNAs dos 403 poços. Os resultados da correlação foram separados em quatro faixas de valores, cada uma representada por uma cor, indicando melhores ou piores resultados. Os resultados são apresentados para a previsão de um a dez anos.



Figura 28 – Variação espacial e temporal do coeficiente de correlação (R), utilizando quatro anos precedentes na camada de entrada das RNAs.

Pela Figura 28, é possível facilmente visualizar como a eficiência das RNAs diminui à medida em que o período de previsão aumenta. Com o aumento do período de previsão, vão surgindo mais poços com coeficiente de correlação entre 0,8 e 0,9, 0,7 e 0,8 e abaixo de 0,7. Também se observa que esse fenômeno ocorre de modo generalizado por todas as áreas do Aquífero Ogallala.

A diminuição no desempenho das RNAs ao aumentar o horizonte de previsão também foi observado por Daliakopoulos, Coulibaly e Tsanis (2005). Eles testaram sete tipos de RNAs com arquiteturas e algoritmos de treinamento diferentes para avaliar a eficiência e a precisão da previsão dos modelos para o nível da água subterrânea no Vale de Messara, em Creta (Grécia). Os resultados mostraram que a eficiência de todos os métodos diminuiu à medida em que o período de previsão aumentou. Além disso, os resultados mostraram que a rede neural do tipo *feedforward* treinada com o algoritmo de *Levenberg-Marquardt* foi a que obteve as previsões mais precisas dos níveis das águas subterrâneas.

Ao analisar a Figura 28, observa-se que poços que apresentaram resultados de coeficiente de correlação abaixo de 0,7 para determinado ano, podem não apresentar um resultado semelhante no ano posterior ou em anos seguintes. Isso acontece pelo fato de os dados de entrada utilizados pelas RNAs serem divididos de maneira aleatória durante a etapa de aprendizagem o que influencia no desempenho das RNAs.

A Figura 29 mostra alguns exemplos dos resultados da previsão do nível da água para um ano à frente, nos poços de 1 a 6, de modo a comparar os valores observados com os calculados pelas RNAs. A linha vermelha mostra os valores previstos pela RNA, nível da água calculado (NAc), e a linha preta representa os dados observados, nível da água observado (NAo).

Já na Figura 30 são apresentados os gráficos de dispersão dos valores observados e calculados, para os poços de 1 a 6, para a previsão de um ano à frente. À esquerda dos gráficos são encontradas as equações de regressão linear, que é representada pela reta azul (linha *fit*), para cada poço. No eixo das ordenadas estão os dados previstos (*output*) e no eixo das abscissas os dados observados (*target*). No título de cada gráfico, além do número do poço ao qual o gráfico corresponde, também é encontrado o coeficiente de correlação (R). A linha diagonal pontilhada corresponde a reta (1:1), representando a linha de ajuste perfeito.

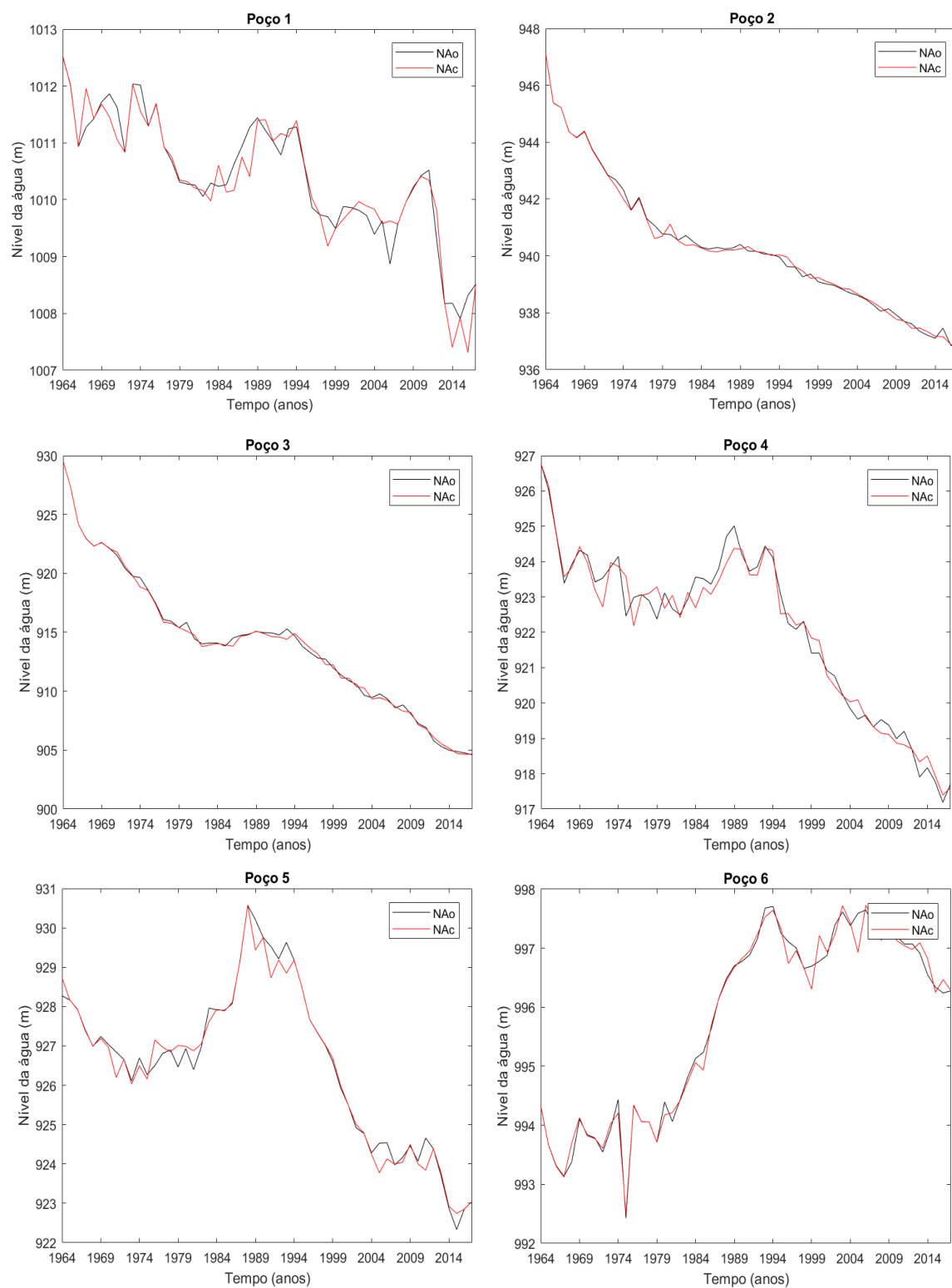


Figura 29 – Amostra dos resultados de previsão para um ano à frente.

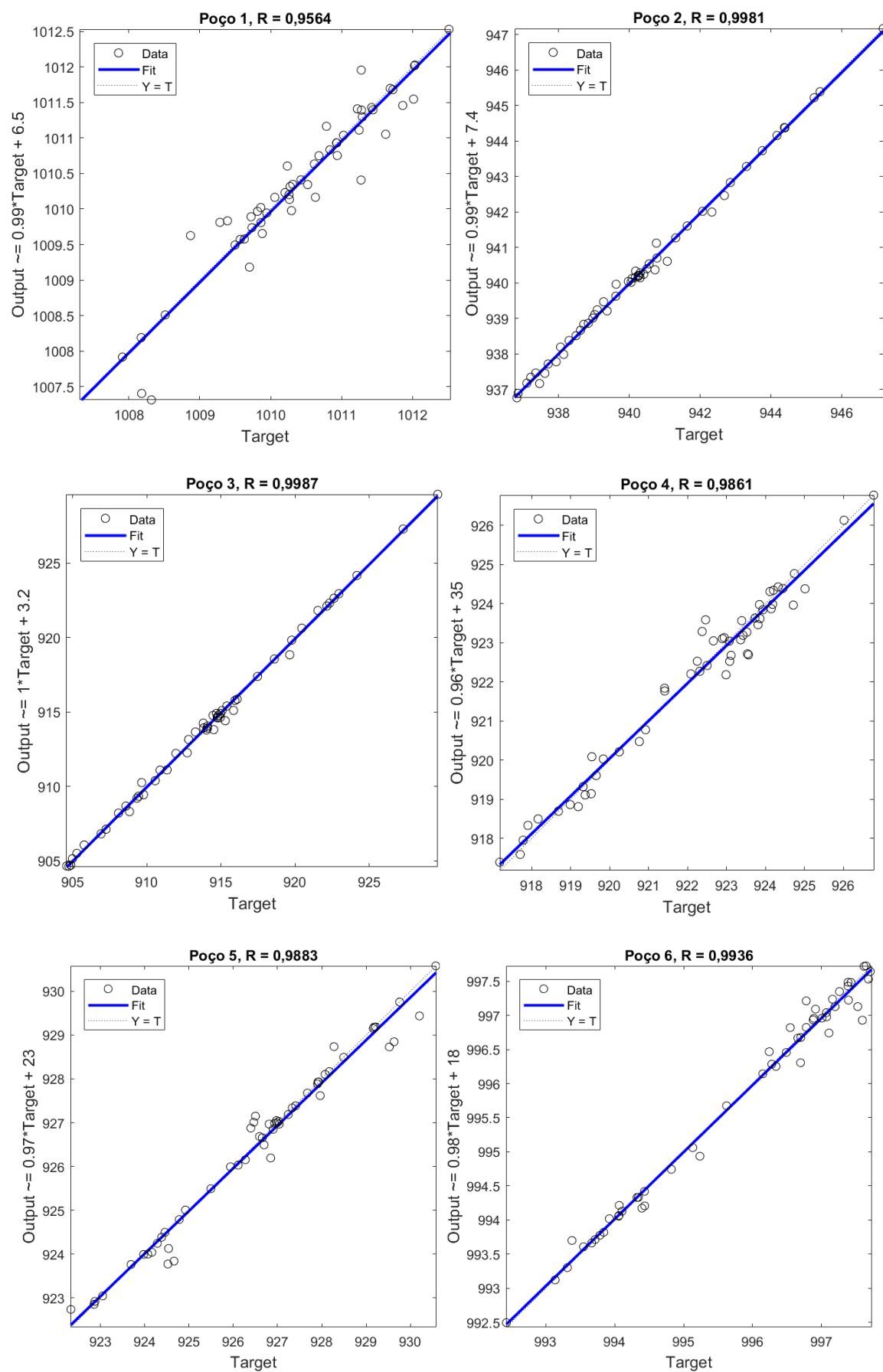


Figura 30 – Gráficos de dispersão para os poços de 1 a 6, para previsão de um ano à frente.

Pelas Figuras 29 e 30, é possível verificar que os modelos de RNA, para os poços de 1 a 6, em alguns momentos subestimam e em outros superestimam os níveis de água subterrânea. Por exemplo, a RNA do poço 1, para o ano de 1967, estima o valor observado, de 1011,28 m, em torno de 1011,96 m, havendo assim uma superestimativa desse valor pela RNA. Já no ano de 2016, a RNA do poço 1 previu um valor 1007,31 m, enquanto o valor observado era de 1008,32 m, havendo, dessa forma, uma subestimativa. Importante ressaltar, entretanto, que para estes exemplos apresentados, mesmo havendo superestimativas e subestimativas, os valores encontrados pelas RNAs não diferem muito do real, sendo os erros, em sua maioria, na casa de alguns centímetros.

Pode-se observar na Figura 30 que os modelos de RNA apresentados possuem estimativas com valores mais densos na vizinhança da linha reta (1:1), o que indica que tais modelos fornecem resultados de previsão satisfatórios.

6.1.2 Resultados das previsões dos níveis d'água em cada poço utilizando doze anos precedentes na camada de entrada

Conforme descrito na metodologia, a quantidade de anos precedentes que obteve os melhores resultados nas previsões foi doze. A Figura 31 mostra a variabilidade espacial e temporal dos valores do coeficiente de correlação para a previsão de um a dez anos à frente, utilizando doze anos precedentes na camada de entrada da RNA. Assim, é possível comparar tais resultados com os obtidos utilizando quatro anos precedentes.

Como discutido anteriormente, para a previsão de um ano à frente (Fig. 28a), nas RNAs com quatro anos precedentes na camada de entrada, três poços obtiveram coeficiente de correlação abaixo de 0,7. Já para a previsão de dez anos à frente (Fig. 28j), esse número aumentou para doze poços. Ao comparar esses resultados com os obtidos pelas RNAs que utilizaram doze anos precedentes na camada de entrada, visualiza-se pela Figura 31 que para a previsão de um ano à frente (Fig. 31a), apenas um poço, o poço 294, teve coeficiente de correlação abaixo de 0,7. Já para a previsão de dez anos à frente (Fig. 31j), dois poços obtiveram coeficiente de correlação abaixo de 0,7.

Assim, observa-se que, de modo geral, os resultados do coeficiente de correlação, utilizando-se doze anos precedentes, foram mais altos para todos os anos previstos, na maioria dos poços, do que utilizando-se quatro anos precedentes. Dessa forma, ao aumentar o período de previsão de um a dez anos à frente, a eficiência dos modelos de RNA, utilizando-se doze anos precedentes, mostra-se melhor do que a das RNAs que usam apenas quatro anos precedentes na camada de entrada.

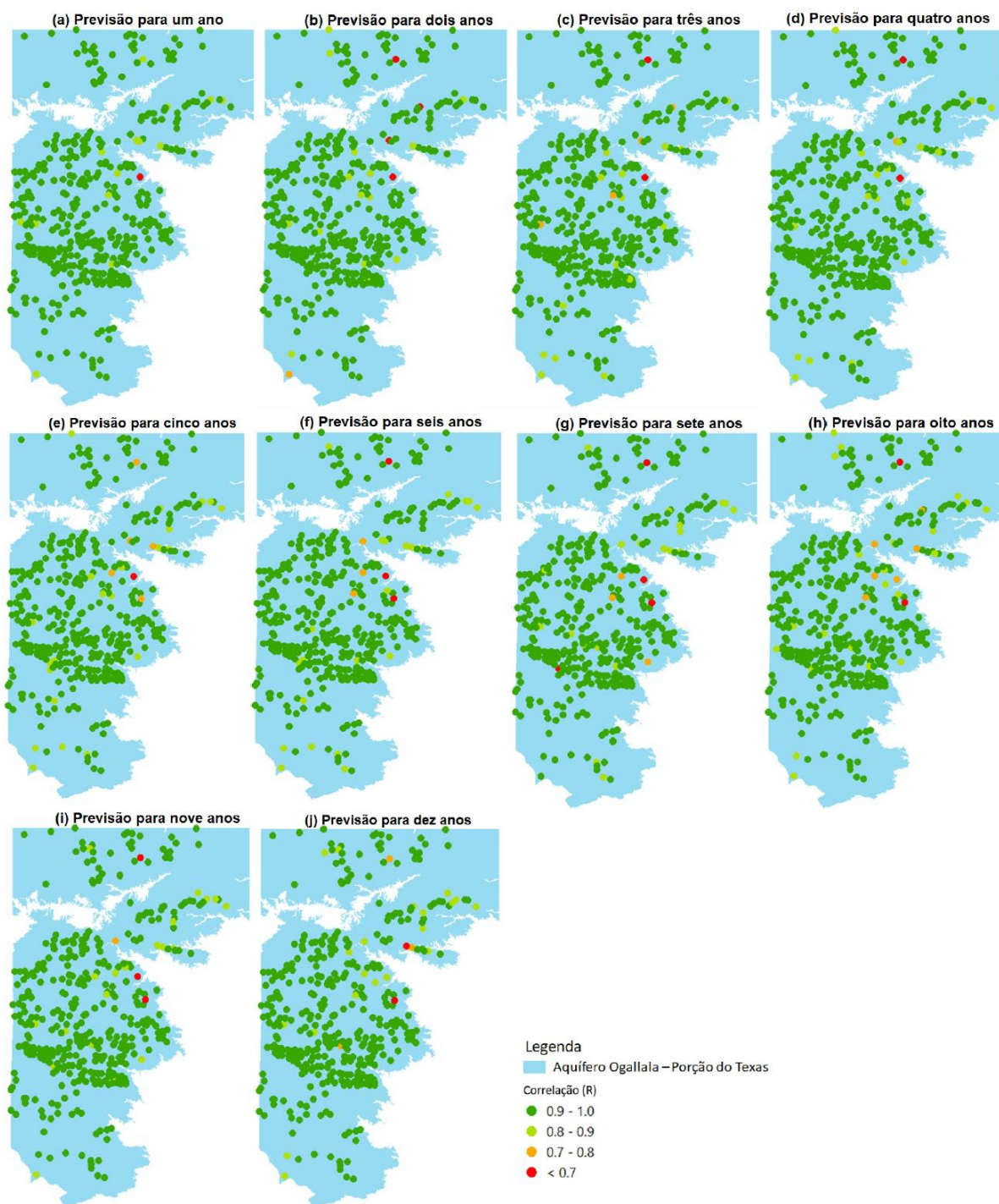


Figura 31 – Variação espacial e temporal do coeficiente de correlação (R), utilizando doze anos precedentes na camada de entrada das RNAs.

6.2 Resultados da aplicação das RNAs desenvolvidas de modo específico para cada poço para realizar a previsão dos níveis d'água em outros poços

As RNAs desenvolvidas de modo específico para cada um dos 403 poços foram testadas uma a uma, para fazer a previsão do nível d'água, em todos os outros poços

analisados no Aquífero Ogallala. Na Tabela 25, que se encontra no Apêndice B, são exibidos os poços que apresentaram resultados satisfatórios para previsão do nível d'água subterrânea ao se utilizar a RNA de cada um dos outros poços.

O critério utilizado para a seleção dos poços, que obtiveram bom desempenho ao se utilizar a RNA específica de outro poço, foi selecionar aqueles que alcançaram NSE acima de 0,5, pois foi observado que poços cuja RNA testada apresentava NSE acima de 0,5 possuía, no geral, os outros índices estatísticos com valores satisfatórios, como, por exemplo, coeficiente de correlação acima de 0,7 e o MAPE e o RMSE com valores próximos a zero. Portanto, para este caso, os poços em que as RNAs testadas obtiveram valores de NSE abaixo de 0,5 foram considerados insatisfatórios.

Ao analisar a Tabela 25, é possível visualizar que a RNA do poço 54 foi a que teve mais poços com índices estatísticos satisfatórios, com um total de 51 poços que apresentaram NSE acima de 0,5. Em seguida, também com uma grande quantidade de poços com índices estatísticos considerados adequados, estão as RNAs dos poços 207 e 315, ambas com 46 poços. Também houve RNAs que, ao serem testadas nos dados de outros poços, não obtiveram nenhum poço com bons resultados, são elas as RNAs dos poços 265, 294, 305 e 390, isso significa que nem com os dados de seu próprio poço essas RNAs obtiveram resultados razoáveis.

Salvo exceções, grande parte das RNAs não tiveram boas previsões ao serem utilizadas em outros poços. Em torno de 56% das RNAs específicas de cada poço, o equivalente a 226 RNAs, só apresentaram até 5 poços com índices estatísticos aceitáveis, dessas, 110 RNAs, o equivalente a 27,3% do total, só obtiveram bons resultados de previsão com os dados de seu próprio poço. Assim, mesmo que algumas RNAs tenham tido bons resultados para uma quantidade razoável de outros poços, a exemplo da RNA do poço 54, a porcentagem de poços em relação ao total ainda é considerada baixa.

Na Figura 32 é apresentada uma amostra da variação espacial dos valores do NSE, das RNAs desenvolvidas especificamente para os poços de 1 a 12, testadas nos outros 402 poços restantes. Para tanto, foram utilizados os valores de quatro anos precedentes na camada de entrada da RNA para a previsão de um ano à frente. Os valores do NSE foram agrupados por faixas, representadas por cores diferentes, indicando melhores ou piores adequações dos poços à RNA testada. Na figura, são apresentados apenas os poços que obtiveram valores de NSE maiores do que 0,5. A localização espacial dos demais poços, que apresentaram valores de correlação menores do que 0,5, é mostrada na cor branca. O número

do poço, cuja RNA foi utilizada nos demais, é indicado em cima da localização espacial do mesmo.



Figura 32 – Variação espacial do NSE para as RNAs do poço 1 ao 12 aplicadas nos demais poços.

6.3 Resultados e discussões da análise de *clusters*

A análise de *clusters* (grupos) foi realizada para os níveis anuais de água subterrânea dos 403 poços de modo a classificá-los de acordo com suas semelhanças na flutuação do nível da água. Empregou-se a análise hierárquica de agrupamentos, utilizando-se a distância euclidiana como métrica de dissimilaridade e o *average linkage* como método de agrupamento. O dendrograma resultante para 30 *clusters* é apresentado na Figura 33. A Tabela 5 mostra quantos e quais são os poços que estão presentes em cada um dos 30 *clusters*.

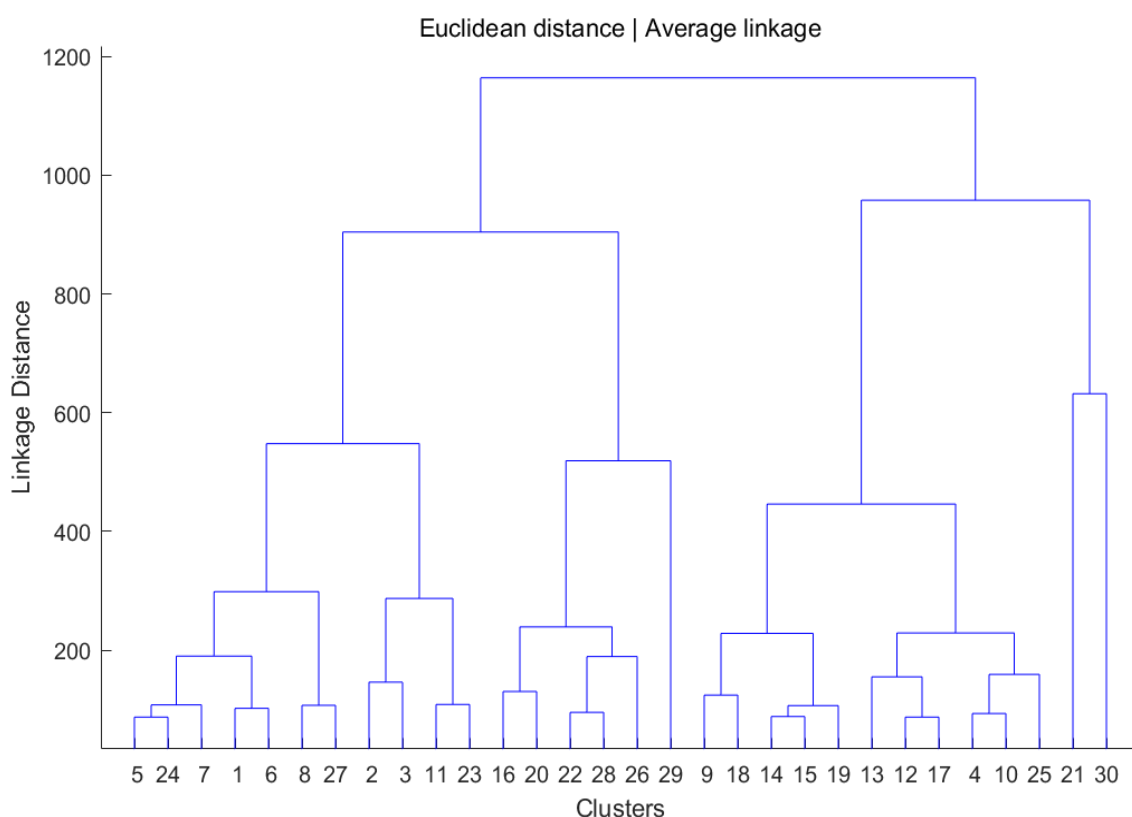


Figura 33 – Dendrograma resultante da análise hierárquica de *clusters* das séries temporais de 403 poços do Aquífero Ogallala.

Ao analisar o dendrograma, é possível observar que os *clusters* 21, 29 e 30 (respectivamente com 3, 1 e 1 poços) possuem uma maior dissimilaridade em relação aos demais, sendo os *clusters* 21 e 30 os mais diferentes quando comparados aos outros grupos, pois possuem a maior distância de ligação (*linkage distance*). Já os pares de *clusters* mais parecidos entre si são 5 e 24, 14 e 15 e 12 e 17.

Tabela 5 – Quantidade e poços que se encontram em cada *cluster*

Cluster Quantidade de poços	Numeração dos poços que estão no cluster especificado.
<i>Cluster 1</i> (25 poços)	1, 11, 49, 112, 116, 120, 131, 171, 183, 186, 213, 243, 261, 267, 278, 287, 299, 309, 331, 342, 358, 364, 370, 371 e 387
<i>Cluster 2</i> (38 poços)	2, 10, 17, 28, 40, 43, 52, 54, 60, 63, 73, 74, 78, 89, 102, 136, 142, 154, 177, 184, 185, 191, 192, 211, 216, 221, 236, 247, 256, 268, 288, 304, 311, 315, 324, 327, 336 e 395
<i>Cluster 3</i> (37 poços)	3, 4, 5, 12, 24, 26, 29, 30, 39, 42, 46, 51, 55, 64, 65, 68, 86, 90, 95, 103, 104, 129, 132, 148, 201, 204, 205, 208, 226, 240, 254, 255, 284, 330, 369, 383 e 397
<i>Cluster 4</i> (6 poços)	44, 77, 82, 138, 159 e 316
<i>Cluster 5</i> (20 poços)	48, 100, 111, 115, 118, 122, 123, 193, 195, 198, 280, 291, 302, 313, 328, 357, 365, 381, 385 e 389
<i>Cluster 6</i> (25 poços)	6, 33, 47, 53, 67, 85, 91, 126, 127, 128, 178, 187, 194, 197, 199, 217, 223, 225, 248, 295, 305, 307, 318, 352 e 384
<i>Cluster 7</i> (9 poços)	7, 37, 45, 114, 117, 119, 121, 283 e 333
<i>Cluster 8</i> (27 poços)	8, 16, 25, 34, 41, 66, 72, 75, 110, 162, 167, 182, 220, 224, 244, 259, 260, 265, 290, 292, 294, 296, 300, 329, 335, 349 e 388
<i>Cluster 9</i> (21 poços)	9, 23, 36, 125, 152, 153, 156, 157, 158, 164, 169, 170, 189, 245, 264, 286, 293, 301, 347, 380 e 390
<i>Cluster 10</i> (8 poços)	50, 59, 71, 84, 151, 214, 310 e 372
<i>Cluster 11</i> (11 poços)	57, 101, 133, 137, 140, 173, 174, 190, 200, 279 e 360
<i>Cluster 12</i> (9 poços)	69, 212, 231, 233, 275, 303, 334, 343 e 346
<i>Cluster 13</i> (22 poços)	13, 22, 31, 76, 80, 135, 139, 147, 160, 163, 206, 210, 215, 249, 253, 273, 285, 289, 319, 345, 361 e 366
<i>Cluster 14</i> (8 poços)	14, 149, 207, 232, 263, 269, 362 e 367
<i>Cluster 15</i> (20 poços)	15, 20, 70, 79, 87, 94, 96, 98, 105, 161, 165, 168, 202, 234, 258, 282, 320, 325, 355 e 377
<i>Cluster 16</i> (16 poços)	83, 93, 130, 141, 219, 230, 270, 272, 277, 306, 314, 341, 374, 386, 393 e 398
<i>Cluster 17</i> (8 poços)	97, 108, 166, 209, 241, 274, 308 e 326
<i>Cluster 18</i> (22 poços)	18, 19, 35, 56, 88, 99, 109, 143, 144, 145, 155, 196, 235, 239, 242, 251, 262, 281, 322, 337, 338 e 339
<i>Cluster 19</i> (12 poços)	106, 107, 113, 146, 176, 188, 250, 317, 340, 348, 356 e 378
<i>Cluster 20</i> (11 poços)	134, 150, 179, 203, 218, 237, 246, 257, 271, 276 e 402
<i>Cluster 21</i> (3 poços)	21, 124 e 222
<i>Cluster 22</i> (12 poços)	172, 297, 312, 321, 323, 350, 351, 382, 392, 396, 399 e 400
<i>Cluster 23</i> (8 poços)	175, 180, 181, 298, 363, 391, 394 e 401
<i>Cluster 24</i> (3 poços)	238, 252 e 344
<i>Cluster 25</i> (1 poço)	332
<i>Cluster 26</i> (4 poços)	353, 368, 375 e 379
<i>Cluster 27</i> (14 poços)	27, 32, 38, 58, 61, 62, 81, 92, 227, 228, 229, 266, 359 e 376
<i>Cluster 28</i> (1 poço)	354
<i>Cluster 29</i> (1 poço)	373
<i>Cluster 30</i> (1 poço)	403

Ao analisar a Tabela 5, observa-se que os *clusters* que possuem mais poços são os *clusters* 2, 3, 8, 1 e 6, cada um com 38, 37, 27, 25 e 25 poços respectivamente. Unindo as informações da Tabela 5 com o dendrograma da Figura 33, constata-se que os *clusters* 1 e 6 possuem uma distância de ligação relativamente pequena, o que demonstra que tais *clusters* são similares entre si. Do mesmo modo, ao observar o dendrograma, os *clusters* 2 e 3 também apresentam similaridades, porém um pouco menores do que as apresentadas entre os *clusters* 1 e 6. Já o *cluster* 8, mostra ter uma maior similaridade com os poços pertencentes ao *cluster* 27.

Por outro lado, os *clusters* 25, 28, 29 e 30 possuem apenas um poço cada. O *cluster* 25 possui similaridades com os poços dos *clusters* 4 e 10. Da mesma forma, o *cluster* 28 é semelhante ao *cluster* 22. Diferente destes dois casos, os *clusters* 29 e 30 possuem grandes dissimilaridades quando comparados aos demais, isso mostra que esses dois poços (373 e 403) são bastante diferentes dos outros, já que, além de se encontrarem em um *cluster* individual, ainda possuem uma grande distância de ligação em relação aos outros.

A Figura 34 mostra o comportamento das séries temporais dos níveis d'água dos poços que se encontram em cada *cluster*. No eixo das ordenadas estão disponíveis informações sobre os níveis d'água das séries temporais, e no eixo das abscissas, o tempo em anos, totalizando 44 anos, do período de 1965 a 2008.

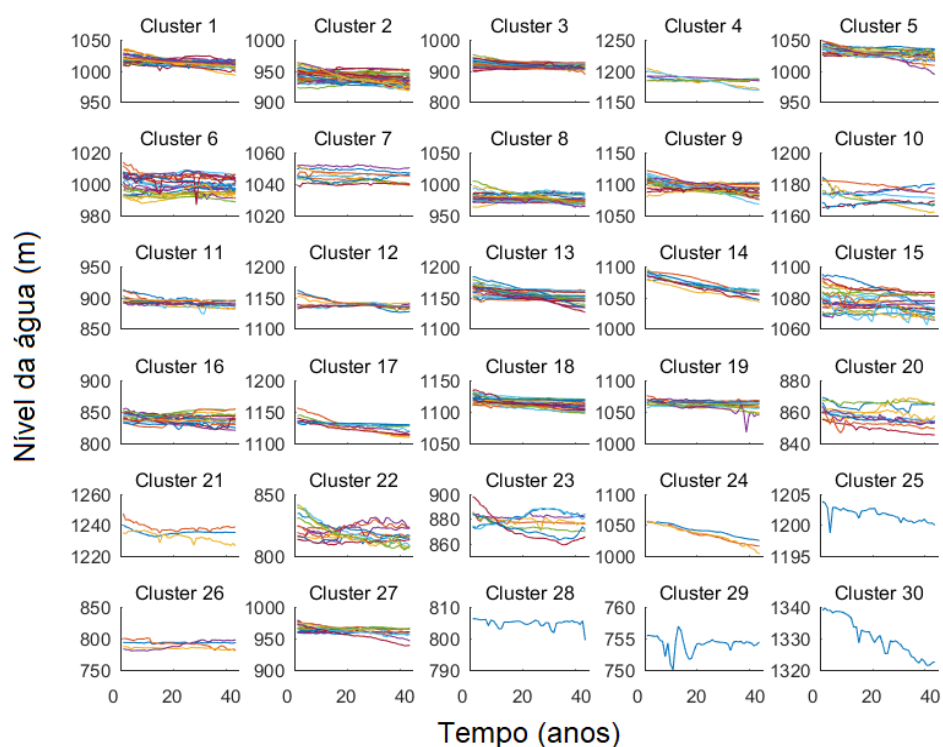


Figura 34 – Comportamento das séries temporais dos níveis d'água dos poços pertencentes a cada *cluster*.

O diagrama de caixa da Figura 35 traz informações sobre as diferenças nas flutuações do nível d'água de cada *cluster* mostrando, de modo mais claro do que na Figura 34, os limites máximos e mínimos em que o nível d'água dos poços de cada *cluster* está compreendido. A linha vermelha no interior de cada caixa indica o valor da mediana do *cluster*.

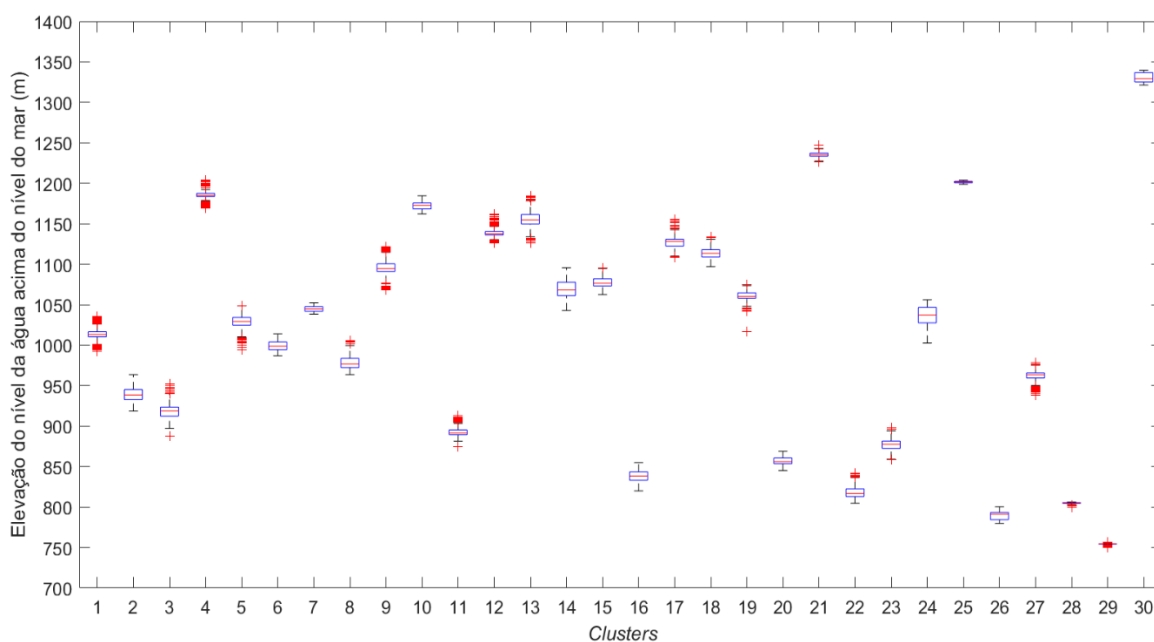


Figura 35 – Flutuações dos níveis das águas subterrâneas nos poços de diferentes *clusters*.

O *cluster* 1 possui uma variação na flutuação do nível d'água de 42,8 m, sendo 992,8 m o valor mínimo e 1035,6 m o máximo. Como discutido anteriormente, os *clusters* 1 e 6 são similares entre si, essa semelhança é justificada pois o *cluster* 6 possui um valor mínimo de nível d'água de 987 m e um valor máximo de 1014 m, o que demonstra que o intervalo da flutuação do nível d'água do *cluster* 6 encontra-se, em sua grande parte, sobreposto ao do *cluster* 1, com a diferença de que, no *cluster* 6, o valor mínimo do nível d'água é ligeiramente menor, aproximadamente 6 m.

O nível d'água do *cluster* 2 varia de 918,6 m, mínimo, a 963,5 m, máximo. O *cluster* que possui um intervalo de nível d'água que mais se aproxima dele é o *cluster* 3, com valor de nível d'água mínimo de 887,8 m e máximo de 952,1 m. Como demonstrado no dendrograma da Figura 33, esses dois *clusters* possuem uma dissimilaridade maior do que a encontrada no exemplo dos *clusters* 1 e 6 e isso é explicado pelo fato de que o nível d'água mínimo do *cluster* 3 é cerca de 30 m menor do que o nível d'água mínimo do *cluster* 2, expondo uma diferença maior entre esses dois *clusters*.

O *cluster* 4 possui níveis d'água entre 1169,1 e 1204 m, o intervalo de elevação do nível d'água no qual esse *cluster* se encontra é mais alto do que o da maioria dos demais *clusters*, com exceção dos *clusters* 21 e 30. O *cluster* que dispõe de níveis d'água mais semelhantes aos do *cluster* 4 é o 10, com um nível d'água mínimo de 1162,0 m e 1184,5 m o máximo, seguido pelo *cluster* 25 com intervalo de nível d'água entre 1198,7 e 1203,8 m. É possível visualizar a relação entre esses *clusters* pelo dendrograma da Figura 33.

O *cluster* 5 está compreendido em um intervalo de nível d'água (de 993,9 a 1048,6 m) que possui a maior quantidade de poços com níveis d'água próximos, um total de 82 poços. É possível observar tal fato ao traçar uma linha imaginária no dendrograma da Figura 33 a uma distância de ligação de 200, obtendo-se assim um *cluster* fruto da união dos *clusters* 1, 5, 6, 7 e 24. Os *clusters* que possuem uma maior semelhança com o *cluster* 5 são o 24 e o 7, com um nível d'água mínimo e máximo de 1002,6 e 1056 m e 1038,3 e 1052,4 m, respectivamente.

O *cluster* 8 detém um nível d'água mínimo de 963,6 m e máximo de 1005 m. Esse *cluster* se encontra em um intervalo de nível d'água um pouco mais baixo, no geral, do que os demais que possuem maiores similaridades com o *cluster* 5. O *cluster* que mais se assemelha ao 8 é o 27, com nível d'água mínimo de 938 m e máximo de 978,8 m. O *cluster* 9 possui uma variação na flutuação do nível d'água de 54,3 m, sendo 1067,8 m, o nível mínimo e 1122,1 m, o máximo. O outro *cluster* que tem níveis d'água semelhantes ao do 9 é o 18, com níveis d'água que vão desde 1097 a 1134,1 m.

O *cluster* 11 dispõe de nível d'água mínimo de 874,4 m e máximo de 913,3 m. O único *cluster* que compartilha de alguns valores de nível d'água com o *cluster* 11 é o 23, que possui nível d'água mínimo e máximo de 858,7 e 897,81 m, respectivamente. Já o *cluster* 12 tem uma variação de nível d'água que vai de 1126,4 a 1162,3 m, os *clusters* que apresentam maiores similaridades com o 12 são o 17 e o 13, com níveis que variam de 1108,9 a 1155 m e 1126,6 a 1184,1 m, nesta ordem.

O *cluster* 14 possui uma variação de nível d'água de 52,8 m, da qual 1042,9 m é a flutuação mínima e 1095,6 m a máxima. Como dito previamente, esse *cluster* apresenta uma grande similaridade com o *cluster* 15, cuja flutuação mínima do nível d'água é 1062,4 m e a máxima é 1095,2 m. É possível observar que os intervalos de níveis d'água desses dois *clusters* são bastante semelhantes entre si, com a diferença de que o *cluster* 14 possui um nível d'água mínimo cerca de 20 m mais baixo do que o do *cluster* 15. Já os níveis d'água máximos de ambos os *clusters* são praticamente iguais. Outro *cluster* que possui níveis d'água semelhantes com ambos os *clusters* 14 e 15 é o *cluster* 19, cujos níveis d'água variam

entre 1016,7 e 1074,9 m; tal *cluster* possui níveis d'água mais baixos do que os anteriores com uma diferença em torno de 26 m do nível d'água mínimo do *cluster* 14 e de 45 m do *cluster* 15. Desse modo, ele se mostra menos similar aos outros, como pode ser visualizado pelo dendrograma da Figura 33.

O *cluster* 16 dispõe de flutuações de nível d'água entre 819,9 e 854,8 m. Poucos *clusters* compartilham níveis d'água dentro desse intervalo, são eles os *clusters* 20 e 22. O *cluster* 20 possui poços com níveis d'água variando entre 845,1 e 868,8 m. Já o *cluster* 22 possui flutuações que vão desde 804,6 a 841,4 m. O *cluster* 22 possui maiores similaridades com o *cluster* 28, cujos níveis d'água variam entre 799,4 e 806,2 m, e ambos possuem um grau de semelhança um pouco menor com o *cluster* 26, cujos níveis d'água mínimo e máximo são 779,8 e 800,3 m, respectivamente.

O *cluster* 29, constituído pelo poço 373, é o que possui os menores valores de flutuação do nível d'água máximo e mínimo, quando comparado aos demais *clusters*. Isso explica por que esse *cluster* possui apenas um poço, já que nenhum dos outros 402 poços restantes apresentam valores de flutuação do nível d'água entre 750 e 757 m.

Em contrapartida, o *cluster* 30 é o que detém os níveis de água mínimo e máximo mais elevados, quando em contraste com os outros *clusters*. Tal fato justifica o motivo desse *cluster* possuir apenas o poço 403, pois esse é o único poço que possui flutuações de nível d'água entre 1321,4 m (mínima) e 1339,6 m (máxima).

O *cluster* que tem níveis d'água mais próximos do *cluster* 30 é o 21, com uma flutuação de nível d'água mínima de 1226,2 m e máxima de 1247,2 m. Ainda assim, há uma grande diferença de altura entre as flutuações de nível d'água desses dois *clusters*, sendo a diferença entre o nível mínimo do *cluster* 30 e o nível máximo no *cluster* 21 de cerca de 74,2 m. Dessa forma, embora o *cluster* 21 seja o que possui os níveis d'água mais próximos do *cluster* 30, eles ainda são bastante diferentes, como evidenciado pela distância de ligação no dendrograma da Figura 33.

O *cluster* 3 foi o que apresentou o maior intervalo entre o nível d'água mínimo e o máximo, visto que a diferença entre esses dois níveis é de 64,4 m. Portanto, esse é o *cluster* que possui poços com comportamentos mais heterogêneos entre si. Já o *cluster* 25, que possui apenas o poço 354, é o que possui a menor diferença entre o nível d'água máximo e o mínimo, com 5 m de diferença. Esse comportamento de haver uma diferença pequena entre o nível d'água máximo e mínimo também se repete nos outros *clusters* que possuem um poço único, a exemplo do *cluster* 28, com 6,87 m de diferença, e do *cluster* 29 com 6,9 m.

A Figura 36 mostra a distribuição espacial dos 403 poços separados em grupos estatisticamente definidos pela análise de *clusters*, no Aquífero Ogallala. Os poços pertencentes a um mesmo *cluster* são representados pela mesma cor. A divisão espacial dos *clusters* fornece informações sobre a localização geográfica e distâncias entre os membros do mesmo *cluster* e de diferentes *clusters*, demonstrando assim a heterogeneidade e conectividade do aquífero.

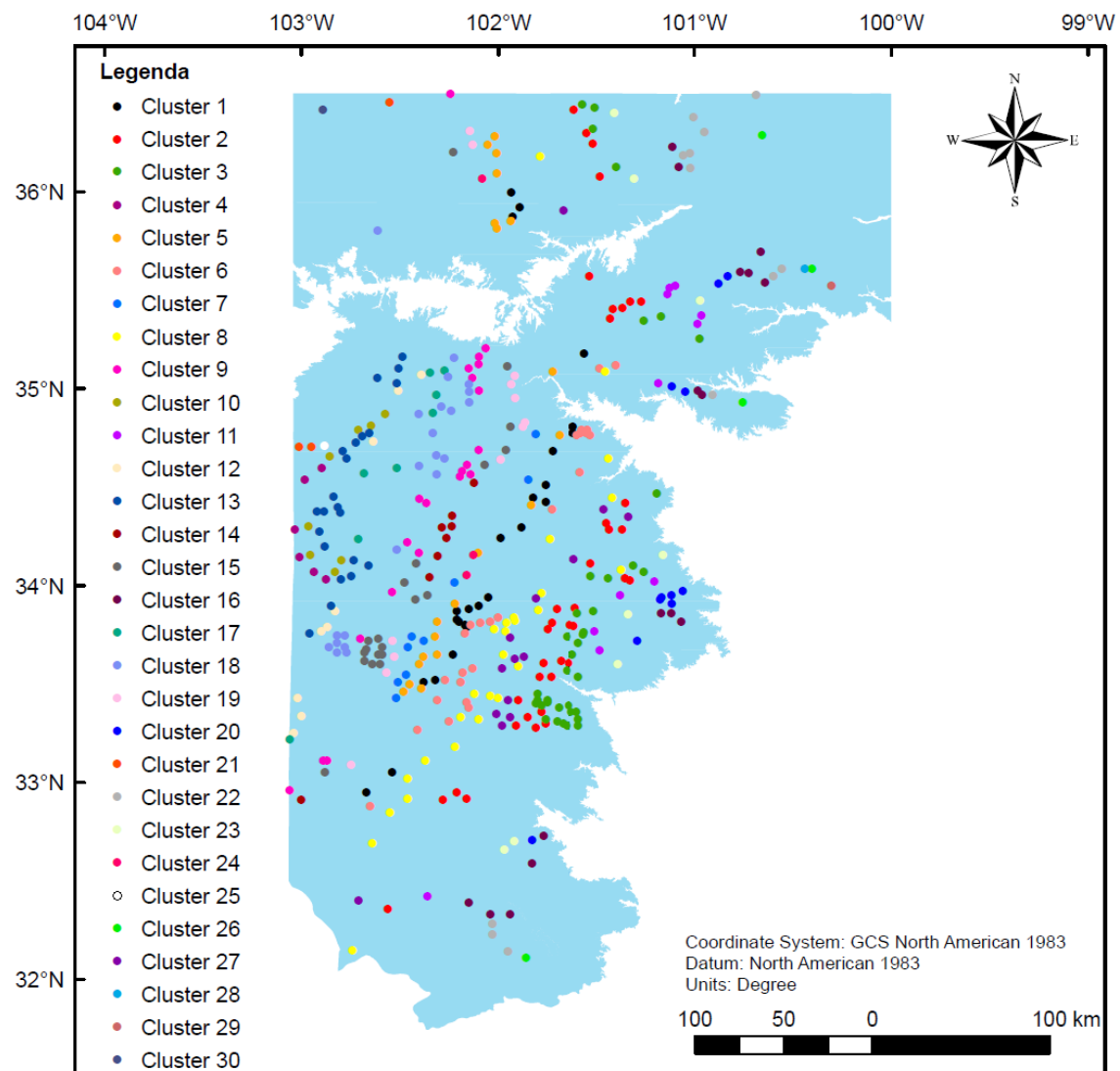


Figura 36 – Distribuição espacial dos poços, no Aquífero Ogallala, divididos por *clusters* resultantes da análise hierárquica de suas séries temporais.

A Figura 37 possui as mesmas informações da Figura 36, porém, na Figura 37 cada *cluster* é exibido de modo separado dos demais, tendo em vista facilitar a visualização do comportamento individual de cada um.



Figura 37 – Distribuição espacial dos poços, no Aquífero Ogallala, divididos por *clusters* resultantes da análise hierárquica de suas séries temporais.

Ao analisar a distribuição espacial dos poços divididos por *clusters* é possível observar que há um padrão peculiar nos *clusters* formados, sendo que a maioria dos *clusters* possui uma forma que lembra um arco, onde os poços pertencentes a cada *cluster* estão agrupados no formato de um arco que vai desde a parte sul até o norte da região do Aquífero Ogallala compreendida no estado do Texas.

Além disso, foi possível constatar que os *clusters* que apresentam uma grande similaridade em relação ao nível d'água nos poços, como evidenciado no dendrograma da Figura 33 e discutido anteriormente, também demonstraram se encontrar próximos uns aos outros em relação à sua distribuição espacial. Alguns exemplos são exibidos na Figura 38, que mostra a sobreposição espacial dos *clusters* 5, 7 e 24 (Fig. 38a), 14 e 15 (Fig. 38b), 12 e 17 (Fig. 38c) e 11 e 23 (Fig. 38d).

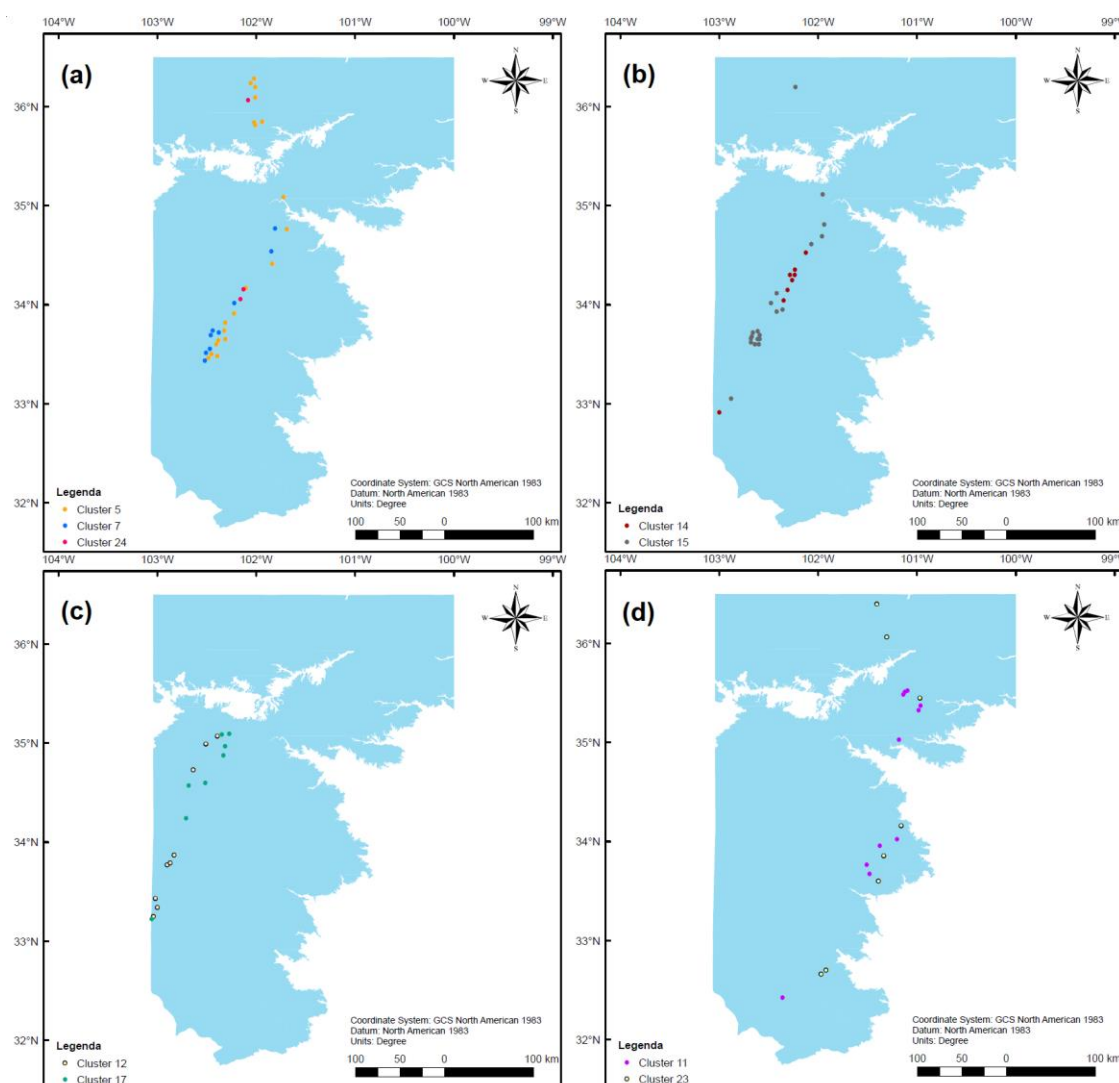


Figura 38 – Arranjo que mostra a sobreposição de alguns dos *clusters* que foram considerados similares pela análise do dendrograma da Figura 33.

O fato de que a grande maioria dos *clusters* possuem um comportamento que lembra o formato de um arco, que atravessa o aquífero de norte a sul, e de que *clusters* com níveis de água similares também se encontram próximos geograficamente, podem ser explicados pelo mesmo motivo. Ao analisar a distribuição espacial dos *clusters* com diferentes faixas de nível d'água foi verificado que o Aquífero Ogallala possui níveis de água subterrânea mais elevados na parte oeste e tais níveis vão diminuindo gradativamente ao seguir para o leste.

Assim, o *cluster* que possui o nível d'água mais elevado, *cluster* 30, encontra-se no extremo oeste da região estudada e ao seguir para o leste os níveis d'água vão caindo até chegar ao nível d'água mais baixo, encontrado no *cluster* 29, localizado no extremo leste do aquífero. Isso pode ser uma possível indicação da existência de faixas de elevação do nível d'água, explicitadas pelo formato de arco dos *clusters*, ao longo do aquífero.

6.3.1 Resultados das previsões dos níveis d'água em cada poço utilizando uma RNA por *cluster*, treinada com os dados de todos os poços do *cluster*

Para cada um dos 30 *clusters*, foram feitas 200 RNAs, treinadas com os dados de todos os poços pertencentes ao *cluster*. A Tabela 6 mostra os resultados dos índices estatísticos de cada RNA selecionada por *cluster*. É possível observar que os resultados dos índices estatísticos obtidos para cada *cluster* são satisfatórios. Com exceção do *cluster* 28, todos os demais possuem RNAs com coeficiente de correlação acima de 0,95 e NSE acima de 0,91.

Os resultados do *cluster* 28 podem ser explicados pelo fato de haver apenas 1 poço nesse *cluster*, o poço 354, que possui níveis d'água diferentes dos níveis dos demais poços. Além disso, como esperado, esse poço também obteve resultados semelhantes na análise dos resultados das RNAs criadas especificamente para cada poço.

Os resultados obtidos pelo MAPE foram considerados satisfatórios nas RNAs de todos os *clusters*, o *cluster* que obteve o maior valor do MAPE e, portanto, o pior resultado em relação aos demais foi o *cluster* 22, com um MAPE de 0,011. O RMSE também obteve bons valores, sendo que a RNA do *cluster* 30 foi a que teve o pior resultado quando comparada às outras, com um RMSE de 0,2889.

Após a escolha da RNA que atingiu os melhores resultados para a previsão do nível d'água de cada *cluster*, cada RNA selecionada por *cluster* foi utilizada para realizar a previsão do nível d'água subterrânea em cada poço pertencente ao *cluster*, de modo específico.

Tabela 6 – Resultados dos índices estatísticos das RNAs selecionadas para cada *cluster*

Clusters Índices	R	NSE	BIAS	MAPE	RMSE
<i>Cluster 1</i>	0,9883	0,9767	0,0000	0,0006	0,0006
<i>Cluster 2</i>	0,9921	0,9839	0,0001	0,0007	0,1014
<i>Cluster 3</i>	0,9931	0,9863	0,0000	0,0007	0,0039
<i>Cluster 4</i>	0,9969	0,9937	0,0000	0,0003	0,0045
<i>Cluster 5</i>	0,9904	0,9808	0,0000	0,0006	0,0271
<i>Cluster 6</i>	0,9859	0,9719	0,0000	0,0005	0,0314
<i>Cluster 7</i>	0,9921	0,9842	0,0000	0,0003	0,0224
<i>Cluster 8</i>	0,9900	0,9802	0,0000	0,0006	0,0122
<i>Cluster 9</i>	0,9959	0,9917	0,0000	0,0004	0,0204
<i>Cluster 10</i>	0,9939	0,9877	0,0001	0,0003	0,0669
<i>Cluster 11</i>	0,9824	0,9649	-0,0001	0,0008	0,0791
<i>Cluster 12</i>	0,9879	0,9758	0,0000	0,0005	0,0049
<i>Cluster 13</i>	0,9974	0,9949	0,0000	0,0004	0,0410
<i>Cluster 14</i>	0,9987	0,9973	0,0000	0,0004	0,0027
<i>Cluster 15</i>	0,9911	0,9823	0,0000	0,0005	0,0006
<i>Cluster 16</i>	0,9903	0,9806	-0,0001	0,0008	0,0673
<i>Cluster 17</i>	0,9971	0,9942	0,0000	0,0004	0,0275
<i>Cluster 18</i>	0,9956	0,9912	0,0000	0,0004	0,0346
<i>Cluster 19</i>	0,9927	0,9851	-0,0001	0,0005	0,1239
<i>Cluster 20</i>	0,9913	0,9826	0,0000	0,0006	0,0326
<i>Cluster 21</i>	0,9821	0,9636	0,0000	0,0004	0,0271
<i>Cluster 22</i>	0,9838	0,9675	0,0001	0,0011	0,1181
<i>Cluster 23</i>	0,9910	0,9820	0,0000	0,0007	0,0255
<i>Cluster 24</i>	0,9987	0,9975	0,0001	0,0005	0,0675
<i>Cluster 25</i>	0,9724	0,9435	0,0000	0,0002	0,0427
<i>Cluster 26</i>	0,9880	0,9760	0,0000	0,0008	0,0346
<i>Cluster 27</i>	0,9912	0,9826	0,0000	0,0006	0,0113
<i>Cluster 28</i>	0,8318	0,5472	-0,0001	0,0004	0,1093
<i>Cluster 29</i>	0,9585	0,9151	0,0001	0,0003	0,0677
<i>Cluster 30</i>	0,9913	0,9792	0,0002	0,0004	0,2889

As Tabelas 7 a 21 mostram os resultados dos índices estatísticos, para cada um dos poços pertencentes a cada *cluster*, obtidos por meio da avaliação da previsão dos níveis d'água nos poços com a RNA do *cluster*. Tal procedimento foi feito para verificar a eficiência das RNAs, treinadas com os dados de todos os poços presentes no *cluster*, na previsão do nível d'água de cada poço especificamente.

A Tabela 7 exhibe os resultados para os poços dos *clusters* 1 e 2. Para a RNA do *cluster* 1, todos os poços alcançaram valores de correlação acima de 0,7 e NSE acima de 0,5. Os poços 112, 131 e 213 foram os que tiveram o menor desempenho. O poço que obteve melhor resultado foi o 186. Para a RNA do *cluster* 2, com exceção do poço 74, todos os

demais apresentaram coeficiente de correlação maior do que 0,7. Em relação ao NSE, os poços 17, 73 e 74 foram os que tiveram menor performance, atingindo valores abaixo de 0,5.

Tabela 7 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 1 e 2 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 1						Cluster 2					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
1	0,9093	0,7954	0,0000	0,0004	0,0250	2	0,9914	0,9729	0,0002	0,0003	0,1638
11	0,9393	0,8678	0,0001	0,0005	0,0699	10	0,9967	0,9922	0,0001	0,0003	0,0823
49	0,9359	0,8708	-0,0001	0,0007	0,0818	17	0,7432	0,3657	0,0004	0,0007	0,3771
112	0,7719	0,5060	0,0001	0,0004	0,0922	28	0,9102	0,7490	0,0004	0,0005	0,3341
116	0,9446	0,8832	0,0001	0,0003	0,1091	40	0,9519	0,8960	0,0000	0,0002	0,0312
120	0,9869	0,9729	0,0000	0,0003	0,0059	43	0,9896	0,9768	0,0001	0,0004	0,1064
131	0,8586	0,7350	0,0000	0,0008	0,0261	52	0,9207	0,8259	0,0001	0,0006	0,0825
171	0,9668	0,9061	0,0002	0,0006	0,1600	54	0,9981	0,9959	-0,0002	0,0007	0,2301
183	0,9798	0,9585	-0,0001	0,0006	0,1368	60	0,9154	0,8045	0,0003	0,0009	0,3160
186	0,9979	0,9956	-0,0002	0,0007	0,2467	63	0,9970	0,9936	0,0000	0,0004	0,0081
213	0,8156	0,5130	0,0001	0,0003	0,0778	73	0,7929	0,4939	0,0005	0,0008	0,4702
243	0,9966	0,9911	0,0000	0,0001	0,0471	74	0,5046	-0,3090	0,0000	0,0005	0,0467
261	0,9470	0,8825	0,0001	0,0003	0,1039	78	0,9456	0,8787	0,0000	0,0004	0,0440
267	0,9725	0,9336	0,0003	0,0007	0,3349	89	0,9918	0,9827	0,0000	0,0007	0,0230
278	0,9418	0,8587	0,0000	0,0007	0,0018	102	0,9514	0,8661	0,0003	0,0005	0,2899
287	0,9847	0,9691	0,0001	0,0009	0,0519	136	0,9777	0,9503	0,0003	0,0012	0,2868
299	0,9956	0,9909	-0,0002	0,0007	0,1857	142	0,9871	0,9731	0,0000	0,0007	0,0111
309	0,9431	0,8825	0,0002	0,0008	0,1825	154	0,9493	0,8736	0,0001	0,0002	0,1015
331	0,9401	0,8808	0,0002	0,0008	0,1619	177	0,9874	0,9743	0,0001	0,0009	0,0677
342	0,9871	0,9742	-0,0001	0,0008	0,1306	184	0,8965	0,7856	0,0003	0,0011	0,2364
358	0,9900	0,9789	-0,0003	0,0008	0,2705	185	0,8691	0,7170	0,0001	0,0006	0,1315
364	0,9549	0,9110	-0,0001	0,0009	0,1495	191	0,9976	0,9773	0,0003	0,0006	0,3040
370	0,9756	0,9468	0,0001	0,0003	0,1246	192	0,9265	0,8422	0,0002	0,0009	0,1945
371	0,9929	0,9856	0,0000	0,0003	0,0183	211	0,9978	0,9950	-0,0001	0,0008	0,0928
387	0,9794	0,9508	-0,0003	0,0006	0,2726	216	0,8928	0,7162	0,0004	0,0007	0,3500
						221	0,9663	0,9221	0,0005	0,0013	0,4614
						236	0,9947	0,9887	-0,0001	0,0006	0,0469
						247	0,9850	0,9689	-0,0003	0,0009	0,2505
						256	0,9936	0,9859	0,0001	0,0005	0,0863
						268	0,8030	0,5827	0,0002	0,0007	0,1675
						288	0,9720	0,9422	-0,0002	0,0011	0,1575
						304	0,9898	0,9796	0,0000	0,0009	0,0311
						311	0,9631	0,9241	0,0001	0,0011	0,1260
						315	0,9887	0,9759	-0,0002	0,0009	0,1679
						324	0,9547	0,9102	-0,0001	0,0009	0,1366
						327	0,9743	0,9434	0,0001	0,0010	0,1062
						336	0,9694	0,9389	-0,0001	0,0011	0,0747
						395	0,9729	0,9464	0,0000	0,0014	0,0393

Na Tabela 8 são apresentados os resultados para os poços dos *clusters* 3 e 4. A RNA do *cluster* 3 atingiu valores de correlação acima de 0,7 para 34 dos 37 poços pertencentes a este *cluster*. Em relação ao NSE, 8 poços tiveram resultados abaixo de 0,5, foram eles os

poços 86, 226, 29, 205, 208, 330, 68 e 132, na ordem do menor valor de NSE para o maior. A RNA do *cluster* 4 obteve resultados de correlação e NSE acima de 0,7 e 0,5, respectivamente, em todos os poços com exceção do poço 77.

Tabela 8 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 3 e 4 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 3						Cluster 4					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
3	0,9951	0,9894	-0,0002	0,0005	0,1553	44	0,9687	0,9316	0,0001	0,0003	0,1303
4	0,9684	0,9367	0,0001	0,0005	0,0664	77	0,6755	0,4326	0,0000	0,0004	0,0424
5	0,9716	0,9405	0,0001	0,0004	0,1196	82	0,9972	0,9942	-0,0001	0,0004	0,1234
12	0,9667	0,9328	0,0001	0,0004	0,0717	138	0,9986	0,9952	-0,0001	0,0001	0,0745
24	0,9954	0,9906	-0,0001	0,0010	0,0691	159	0,8123	0,6281	0,0001	0,0003	0,0919
26	0,9926	0,9846	-0,0001	0,0003	0,0573	316	0,9993	0,9984	-0,0001	0,0003	0,1109
29	0,6913	0,2592	0,0001	0,0006	0,0885						
30	0,9054	0,7820	0,0003	0,0006	0,3031						
39	0,9677	0,9314	-0,0002	0,0008	0,2237						
42	0,9941	0,9865	-0,0002	0,0006	0,1848						
46	0,9187	0,7865	0,0003	0,0009	0,2465						
51	0,9924	0,9841	-0,0002	0,0007	0,1812						
55	0,9940	0,9879	-0,0001	0,0005	0,0933						
64	0,8988	0,7585	0,0002	0,0007	0,2263						
65	0,9937	0,9856	-0,0004	0,0009	0,3375						
68	0,8090	0,6225	0,0001	0,0005	0,0880						
86	0,3997	-0,2049	0,0000	0,0006	0,0199						
90	0,8972	0,7972	-0,0001	0,0007	0,0503						
95	0,9406	0,8751	0,0000	0,0005	0,0397						
103	0,9598	0,9061	0,0002	0,0004	0,1724						
104	0,9179	0,8198	0,0001	0,0007	0,0938						
129	0,9878	0,9758	0,0000	0,0008	0,0031						
132	0,8097	0,4036	0,0003	0,0004	0,2571						
148	0,9918	0,9822	-0,0003	0,0009	0,2510						
201	0,9386	0,8711	0,0002	0,0010	0,1892						
204	0,9431	0,8825	0,0002	0,0006	0,1914						
205	0,7241	0,4546	0,0004	0,0011	0,3358						
208	0,7411	0,4482	0,0001	0,0004	0,0825						
226	0,5917	-0,0808	0,0002	0,0003	0,1670						
240	0,9953	0,9896	-0,0001	0,0007	0,0746						
254	0,9525	0,9020	0,0000	0,0007	0,0300						
255	0,9859	0,9677	-0,0002	0,0009	0,2175						
284	0,8621	0,7424	-0,0001	0,0010	0,0955						
330	0,7718	0,4381	0,0002	0,0004	0,1701						
369	0,9842	0,9682	0,0000	0,0011	0,0247						
383	0,9893	0,9773	-0,0003	0,0009	0,3150						
397	0,9914	0,9786	-0,0005	0,0008	0,4829						

Os resultados obtidos pelas RNAs dos *clusters* 5 e 6 são mostrados na Tabela 9. As RNAs dos *clusters* 5 e 6 não atingiram valores de correlação e NSE superiores a 0,7 e 0,5,

nesta ordem, em apenas um poço dentre todos os poços pertencentes a cada *cluster*. Foram eles, o poço 291, para a RNA do *cluster* 5, e o poço 305, para a RNA do *cluster* 6.

Tabela 9 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 5 e 6 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 5						Cluster 6					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
48	0,9715	0,9353	0,0001	0,0003	0,0638	6	0,9291	0,8490	0,0000	0,0004	0,0038
100	0,9885	0,9742	-0,0001	0,0005	0,0524	33	0,8365	0,6031	-0,0001	0,0003	0,0761
111	0,9832	0,9495	0,0003	0,0005	0,3142	47	0,9744	0,9490	0,0000	0,0003	0,0148
115	0,9496	0,8949	0,0001	0,0006	0,1155	53	0,9038	0,8069	-0,0001	0,0005	0,0842
118	0,9449	0,8802	0,0001	0,0003	0,0899	67	0,9264	0,8223	-0,0001	0,0003	0,0666
122	0,9566	0,8888	0,0001	0,0003	0,1469	85	0,9854	0,9685	0,0000	0,0002	0,0178
123	0,9803	0,9520	0,0002	0,0003	0,1660	91	0,9773	0,9447	0,0000	0,0003	0,0344
193	0,9278	0,8479	0,0001	0,0003	0,1367	126	0,9766	0,9427	-0,0001	0,0004	0,0744
195	0,9983	0,9961	-0,0002	0,0006	0,2066	127	0,9906	0,9806	0,0000	0,0003	0,0091
198	0,9843	0,9686	0,0000	0,0003	0,0310	128	0,9825	0,9540	-0,0001	0,0002	0,0733
280	0,9901	0,9793	-0,0001	0,0010	0,0981	178	0,9337	0,8682	0,0001	0,0007	0,0612
291	0,3297	-0,2527	0,0001	0,0004	0,1037	187	0,9561	0,8976	-0,0001	0,0004	0,0885
302	0,9907	0,9715	0,0001	0,0004	0,0628	194	0,8997	0,7042	0,0002	0,0004	0,2318
313	0,9635	0,9264	-0,0001	0,0009	0,1341	197	0,9509	0,8839	0,0002	0,0004	0,2137
328	0,9650	0,9309	0,0000	0,0008	0,0013	199	0,9595	0,9158	0,0001	0,0005	0,0738
357	0,9359	0,8748	-0,0002	0,0008	0,1964	217	0,9576	0,8868	0,0002	0,0004	0,1697
365	0,8975	0,7973	-0,0005	0,0012	0,5165	223	0,9646	0,9279	0,0000	0,0003	0,0279
381	0,9980	0,9954	-0,0004	0,0007	0,3627	225	0,9628	0,9092	-0,0004	0,0007	0,4434
385	0,9320	0,8670	-0,0001	0,0011	0,1225	248	0,9732	0,9402	-0,0001	0,0002	0,1066
389	0,9884	0,9717	-0,0001	0,0008	0,0558	295	0,8249	0,6019	0,0000	0,0003	0,0277
						305	0,4819	0,1983	-0,0006	0,0014	0,5565
						307	0,9244	0,8482	-0,0002	0,0009	0,2454
						318	0,9605	0,9091	0,0002	0,0005	0,1978
						352	0,9405	0,8765	-0,0001	0,0007	0,0877
						384	0,8982	0,7652	0,0002	0,0005	0,1900

A Tabela 10 apresenta os resultados das RNAs dos *clusters* 7 e 8. A RNA do *cluster* 7 alcançou valores de correlação superiores a 0,7 para 8 dos 9 poços presentes no *cluster*, o único poço que não teve um bom resultado de correlação foi o 121. Em relação ao NSE, dois poços (119 e 121) tiveram valores abaixo de 0,5.

Já a RNA do *cluster* 8 não obteve resultados satisfatórios de correlação para 3 dos 27 poços pertencentes ao *cluster*, são eles os poços 294, 265 e 34, em ordem crescente de valores de correlação. O NSE teve resultados abaixo de 0,5 para 4 poços, incluindo os mesmos que obtiveram resultados de correlação insatisfatórios, somado ao poço 182.

Tabela 10 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 7 e 8 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 7						Cluster 8					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
7	0,9684	0,9346	-0,0001	0,0003	0,0732	8	0,9519	0,9037	0,0000	0,0005	0,0185
37	0,9827	0,9584	0,0000	0,0002	0,0084	16	0,9414	0,7370	0,0000	0,0002	0,0369
45	0,9933	0,9837	-0,0002	0,0003	0,2414	25	0,9766	0,9524	0,0001	0,0006	0,1027
114	0,9234	0,8510	0,0000	0,0003	0,0048	34	0,6738	0,4304	0,0000	0,0002	0,0093
117	0,9320	0,8644	-0,0001	0,0004	0,0713	41	0,9677	0,9303	0,0001	0,0002	0,0901
119	0,7066	0,4184	0,0000	0,0003	0,0522	66	0,9716	0,9417	0,0001	0,0002	0,0545
121	0,6706	0,2599	0,0001	0,0003	0,1279	72	0,9865	0,9527	0,0001	0,0003	0,0755
283	0,9509	0,9036	0,0000	0,0004	0,0501	75	0,9725	0,9445	0,0000	0,0005	0,0019
333	0,8053	0,6400	0,0000	0,0003	0,0121	110	0,9877	0,9632	0,0000	0,0003	0,0446
						162	0,7896	0,5609	-0,0001	0,0004	0,0636
						167	0,8397	0,6318	0,0002	0,0003	0,1599
						182	0,7771	0,4534	0,0000	0,0006	0,0156
						220	0,9026	0,8132	-0,0001	0,0010	0,0618
						224	0,9381	0,8792	0,0000	0,0009	0,0030
						244	0,9399	0,8653	0,0003	0,0006	0,2634
						259	0,9742	0,9457	-0,0003	0,0008	0,2648
						260	0,9977	0,9941	-0,0003	0,0006	0,3359
						265	0,2214	-0,5226	0,0002	0,0007	0,2023
						290	0,7755	0,5893	0,0000	0,0007	0,0366
						292	0,9128	0,8050	-0,0001	0,0007	0,1309
						294	-0,0120	-0,3325	0,0003	0,0010	0,2792
						296	0,7654	0,5799	0,0001	0,0006	0,0529
						300	0,9402	0,8822	-0,0001	0,0005	0,0764
						329	0,9877	0,9752	0,0000	0,0005	0,0341
						335	0,9811	0,9562	-0,0004	0,0009	0,4261
						349	0,9969	0,9921	-0,0004	0,0008	0,4074
						388	0,8048	0,6294	0,0002	0,0015	0,1566

Na Tabela 11 são encontrados os resultados obtidos pelas RNAs dos *clusters* 9 e 10. A RNA do *cluster* 9 alcançou resultados de correlação e NSE acima de 0,7 e 0,5, nesta ordem, para todos os poços, com exceção do poço 390. A RNA do *cluster* 10 obteve correlação e NSE adequados para todos os poços.

A Tabela 12 exhibe os valores alcançados pelas RNAs dos *clusters* 11 e 12. A RNA do *cluster* 11 atingiu valores de correlação e NSE acima de 0,7 e 0,5, respectivamente, para 9 dos 11 poços do *cluster*. Os poços que não tiveram bons resultados de correlação e NSE foram o 279 e o 360. Já a RNA do *cluster* 12 teve valores de correlação e NSE acima de 0,7 e 0,5, na devida ordem, para todos os poços.

Tabela 11 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 9 e 10 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 9						Cluster 10					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
9	0,9996	0,9990	-0,0001	0,0002	0,0977	50	0,9863	0,9703	0,0001	0,0003	0,1409
23	0,9991	0,9980	-0,0001	0,0005	0,1062	59	0,9930	0,9860	0,0000	0,0003	0,0371
36	0,9382	0,8526	0,0001	0,0002	0,0925	71	0,9947	0,9891	-0,0001	0,0003	0,0821
125	0,9942	0,9882	0,0000	0,0003	0,0424	84	0,9042	0,7680	0,0002	0,0004	0,1886
152	0,9871	0,9668	0,0001	0,0004	0,1275	151	0,9801	0,9510	0,0000	0,0003	0,0148
153	0,9988	0,9975	0,0000	0,0003	0,0133	214	0,9374	0,8658	0,0000	0,0002	0,0507
156	0,9624	0,9239	0,0001	0,0003	0,0755	310	0,8829	0,7420	0,0001	0,0004	0,1313
157	0,9530	0,9070	0,0001	0,0005	0,0566	372	0,8226	0,6754	0,0000	0,0003	0,0326
158	0,9963	0,9926	0,0000	0,0002	0,0207						
164	0,8241	0,6300	0,0002	0,0007	0,2615						
169	0,9933	0,9818	0,0000	0,0003	0,0364						
170	0,9969	0,9933	-0,0001	0,0003	0,0663						
189	0,9994	0,9987	-0,0001	0,0005	0,1485						
245	0,9591	0,9197	0,0000	0,0008	0,0174						
264	0,9130	0,8209	0,0001	0,0004	0,1600						
286	0,9715	0,9427	0,0000	0,0002	0,0470						
293	0,9457	0,8784	0,0001	0,0006	0,0545						
301	0,9931	0,9857	0,0000	0,0007	0,0304						
347	0,9854	0,9698	0,0001	0,0006	0,0847						
380	0,9896	0,9775	0,0000	0,0003	0,0216						
390	0,5105	0,2260	-0,0001	0,0012	0,1170						

Tabela 12 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 11 e 12 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 11						Cluster 12					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
57	0,9906	0,9799	-0,0002	0,0011	0,1812	69	0,9867	0,9726	-0,0001	0,0006	0,1582
101	0,9932	0,9848	-0,0004	0,0009	0,3708	212	0,8664	0,7339	0,0001	0,0004	0,1520
133	0,9447	0,8659	-0,0003	0,0007	0,2830	231	0,9620	0,9041	0,0002	0,0004	0,1977
137	0,8243	0,6473	0,0001	0,0008	0,1024	233	0,7807	0,5465	0,0001	0,0006	0,1230
140	0,9205	0,8339	-0,0001	0,0004	0,0964	275	0,9370	0,8733	-0,0002	0,0007	0,2065
173	0,9117	0,8250	-0,0002	0,0010	0,2210	303	0,9676	0,9361	0,0000	0,0006	0,0416
174	0,7784	0,5744	-0,0001	0,0006	0,0924	334	0,8607	0,7363	0,0000	0,0006	0,0057
190	0,8634	0,7253	0,0003	0,0008	0,2575	343	0,9983	0,9965	0,0000	0,0003	0,0274
200	0,9826	0,9651	0,0000	0,0006	0,0422	346	0,9987	0,9973	-0,0001	0,0002	0,0571
279	0,5446	0,1311	0,0002	0,0006	0,1580						
360	0,6627	0,4001	-0,0001	0,0008	0,0578						

Na Tabela 13 são apresentados os resultados adquiridos pelas RNAs dos *clusters* 13 e 14. A RNA do *cluster* 13 obteve valores de correlação superiores a 0,7 para 21 dos 22 poços do *cluster*, o poço 13 foi o único com correlação abaixo de 0,7. No que diz respeito ao NSE, dois poços (13 e 160) tiveram valores abaixo de 0,5. A RNA do *cluster* 14 alcançou bons resultados em todos os poços presentes no *cluster*.

Tabela 13 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 13 e 14 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 13						Cluster 14					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
13	0,5999	0,2464	0,0000	0,0004	0,0156	14	0,9979	0,9953	0,0001	0,0003	0,1014
22	0,9984	0,9963	-0,0001	0,0004	0,1692	149	0,9995	0,9989	0,0000	0,0004	0,0182
31	0,9966	0,9908	0,0000	0,0003	0,0435	207	0,9990	0,9981	0,0000	0,0004	0,0310
76	0,9989	0,9975	-0,0001	0,0004	0,1107	232	0,9979	0,9956	0,0001	0,0005	0,0814
80	0,9982	0,9964	-0,0001	0,0006	0,1229	263	0,9976	0,9951	0,0000	0,0005	0,0004
135	0,8745	0,7222	0,0000	0,0002	0,0509	269	0,9943	0,9884	-0,0001	0,0006	0,0780
139	0,9991	0,9976	-0,0004	0,0005	0,4087	362	0,9989	0,9978	0,0000	0,0004	0,0530
147	0,9870	0,9718	-0,0002	0,0007	0,1781	367	0,9986	0,9970	0,0000	0,0005	0,0204
160	0,7055	0,3714	0,0000	0,0003	0,0065						
163	0,9845	0,9669	0,0001	0,0002	0,0635						
206	0,9695	0,9385	0,0000	0,0005	0,0534						
210	0,9964	0,9926	0,0000	0,0004	0,0102						
215	0,9831	0,9597	0,0001	0,0002	0,1501						
249	0,9923	0,9839	0,0000	0,0004	0,0529						
253	0,9991	0,9979	-0,0001	0,0005	0,1696						
273	0,8132	0,5790	0,0001	0,0006	0,1014						
285	0,9919	0,9805	-0,0001	0,0004	0,0801						
289	0,9963	0,9925	0,0000	0,0005	0,0518						
319	0,9990	0,9979	-0,0001	0,0002	0,0780						
345	0,9956	0,9870	0,0001	0,0002	0,1741						
361	0,9839	0,9676	0,0000	0,0003	0,0485						
366	0,9811	0,9622	0,0001	0,0003	0,0664						

A Tabela 14 mostra os resultados das RNAs dos *clusters* 15 e 16. A RNA do *cluster* 15 obteve valores de correlação superiores a 0,7 para 19 dos 20 poços pertencentes a esse *cluster*, sendo o poço 15 o único que não alcançou bons resultados. No que concerne ao NSE, dois poços (15 e 325) apresentaram valores abaixo de 0,5. A RNA do *cluster* 16 obteve valor de correlação menor do que 0,7 apenas para o poço 230. Os poços 230 e 274 foram os únicos com valores de NSE abaixo de 0,5, para a RNA do *cluster* 16.

Na Tabela 15 são encontrados os resultados das RNAs dos *clusters* 17 e 18. Para ambas as RNAs, todos os poços obtiveram resultados satisfatórios, com valores de correlação e NSE acima de 0,7 e 0,5, respectivamente.

Tabela 14 –Resultados dos índices estatísticos das RNAs feitas para os *clusters* 15 e 16 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 15						Cluster 16					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
15	0,5934	0,0371	0,0000	0,0003	0,0115	83	0,8617	0,6949	-0,0002	0,0007	0,1527
20	0,9809	0,9618	0,0000	0,0005	0,0280	93	0,9836	0,9585	-0,0003	0,0009	0,2473
70	0,9493	0,8987	0,0000	0,0005	0,0090	130	0,9171	0,8361	0,0001	0,0009	0,0515
79	0,9866	0,9583	0,0001	0,0002	0,0960	141	0,9807	0,9595	-0,0002	0,0009	0,1650
87	0,9397	0,8799	0,0001	0,0003	0,0889	219	0,9881	0,9749	0,0001	0,0008	0,1211
94	0,9768	0,9538	0,0000	0,0003	0,0411	230	0,3931	-0,2207	-0,0001	0,0007	0,0866
96	0,9762	0,9303	0,0001	0,0002	0,1573	270	0,9301	0,8581	-0,0001	0,0008	0,0759
98	0,9993	0,9970	-0,0002	0,0004	0,2483	272	0,9698	0,9375	-0,0002	0,0009	0,1964
105	0,9858	0,9709	-0,0001	0,0006	0,0938	277	0,9393	0,8728	0,0001	0,0007	0,0588
161	0,9935	0,9859	-0,0002	0,0006	0,2425	306	0,7960	0,6262	-0,0001	0,0005	0,0525
165	0,9648	0,8912	0,0003	0,0004	0,3410	314	0,9972	0,9930	-0,0003	0,0007	0,2821
168	0,9538	0,9083	0,0001	0,0006	0,0939	341	0,8955	0,7850	-0,0002	0,0010	0,1428
202	0,9738	0,9482	0,0000	0,0006	0,0119	374	0,7194	0,4105	0,0001	0,0005	0,0871
234	0,9856	0,9704	0,0001	0,0003	0,0571	386	0,9190	0,8311	0,0000	0,0013	0,0082
258	0,9818	0,9609	0,0000	0,0003	0,0034	393	0,9121	0,8231	-0,0004	0,0012	0,3002
282	0,7833	0,5932	0,0000	0,0006	0,0085	398	0,9915	0,9813	0,0004	0,0009	0,3108
320	0,9371	0,8705	0,0000	0,0005	0,0184						
325	0,7577	0,4549	0,0000	0,0002	0,0433						
355	0,9087	0,8192	0,0001	0,0005	0,0646						
377	0,8976	0,7996	-0,0003	0,0015	0,3537						

Tabela 15– Resultados dos índices estatísticos das RNAs feitas para os *clusters* 17 e 18 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 17						Cluster 18					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
97	0,9602	0,9185	0,0001	0,0004	0,1098	18	0,9906	0,9807	0,0000	0,0004	0,0490
108	0,9985	0,9969	-0,0001	0,0004	0,1571	19	0,9734	0,9454	0,0000	0,0004	0,0421
166	0,9982	0,9963	0,0000	0,0004	0,0038	35	0,9535	0,9000	0,0001	0,0003	0,1156
209	0,9937	0,9870	0,0001	0,0005	0,1108	56	0,9637	0,9264	0,0001	0,0004	0,0798
241	0,9947	0,9889	0,0001	0,0004	0,1089	88	0,9726	0,9434	-0,0001	0,0005	0,1026
274	0,9902	0,9806	0,0000	0,0004	0,0051	99	0,9970	0,9893	0,0002	0,0002	0,2182
308	0,9975	0,9944	0,0000	0,0003	0,0337	109	0,9990	0,9958	-0,0001	0,0002	0,1052
326	0,9037	0,7827	0,0001	0,0002	0,0893	143	0,9925	0,9827	-0,0001	0,0004	0,1012
						144	0,9326	0,8611	0,0000	0,0004	0,0282
						145	0,9972	0,9923	-0,0001	0,0003	0,0699
						155	0,9995	0,9987	-0,0001	0,0001	0,0951
						196	0,9180	0,8336	0,0002	0,0005	0,1750
						235	0,9978	0,9955	0,0000	0,0003	0,0445
						239	0,9964	0,9924	-0,0002	0,0006	0,2134
						242	0,9161	0,8267	0,0000	0,0005	0,0474
						251	0,9893	0,9779	-0,0001	0,0005	0,0759
						262	0,9983	0,9955	0,0000	0,0001	0,0409
						281	0,9875	0,9724	-0,0002	0,0007	0,2380
						322	0,9975	0,9924	-0,0002	0,0004	0,2671
						337	0,9914	0,9788	-0,0002	0,0005	0,2253
						338	0,9361	0,8612	0,0001	0,0003	0,1256
						339	0,7760	0,5524	0,0000	0,0003	0,0092

Na Tabela 16 são apresentados os resultados obtidos pelas RNAs dos *clusters* 19 e 20. A RNA do *cluster* 19 alcançou valores de correlação superiores a 0,7 para 10 dos 12 poços pertencentes ao *cluster*. Em relação ao NSE, a RNA obteve valores menores do que 0,5 para três poços (317, 176 e 106). A RNA do *cluster* 20 teve resultados de correlação e NSE superiores a 0,7 e 0,5, na devida ordem, para todos os poços.

Tabela 16 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 19 e 20 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 19						Cluster 20					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
106	0,7039	0,3787	-0,0001	0,0003	0,0791	134	0,9090	0,8042	0,0000	0,0007	0,0047
107	0,9056	0,8183	0,0000	0,0003	0,0452	150	0,9895	0,9790	0,0000	0,0003	0,0044
113	0,9963	0,9883	0,0000	0,0004	0,0419	179	0,9815	0,9615	0,0001	0,0008	0,1044
146	0,9838	0,9667	-0,0001	0,0006	0,1499	203	0,9451	0,8888	-0,0001	0,0005	0,0865
176	0,6678	0,3738	0,0000	0,0006	0,0148	218	0,9077	0,8193	0,0000	0,0006	0,0029
188	0,9701	0,9362	0,0000	0,0004	0,0303	237	0,9731	0,9465	0,0000	0,0004	0,0159
250	0,9909	0,9801	-0,0001	0,0003	0,1412	246	0,9868	0,9654	-0,0002	0,0004	0,1662
317	0,3171	-0,3886	-0,0001	0,0004	0,0992	257	0,9904	0,9807	0,0000	0,0004	0,0354
340	0,9479	0,8981	0,0000	0,0005	0,0131	271	0,9469	0,8512	-0,0001	0,0005	0,0489
348	0,9395	0,8574	-0,0002	0,0005	0,1815	276	0,9323	0,8650	-0,0001	0,0009	0,0628
356	0,9928	0,9826	-0,0005	0,0009	0,5412	402	0,9282	0,8488	-0,0002	0,0009	0,1754
378	0,9899	0,9782	-0,0003	0,0009	0,3382						

Os resultados atingidos pelas RNAs dos *clusters* 21 e 22 são exibidos na Tabela 17. A RNA do *cluster* 21 obteve valores considerados satisfatórios para todos os poços. A RNA do *cluster* 22 teve resultados de correlação acima de 0,7 para 11 dos 12 poços do *cluster*, o único poço que não atingiu valores adequados foi o poço 172. Quanto ao NSE, dois poços tiveram resultados inferiores a 0,5, são eles os poços 172 e 351.

Tabela 17 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 21 e 22 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 21						Cluster 22					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
21	0,9705	0,8867	0,0001	0,0003	0,1580	172	0,6068	0,2423	0,0001	0,0012	0,1049
124	0,9748	0,9490	0,0001	0,0004	0,0633	297	0,7929	0,5699	0,0002	0,0013	0,1676
222	0,9599	0,9098	-0,0001	0,0006	0,1610	312	0,8558	0,6629	0,0003	0,0005	0,2538
						321	0,9026	0,7872	0,0003	0,0013	0,2072
						323	0,9921	0,9805	0,0002	0,0009	0,1846
						350	0,9927	0,9847	-0,0003	0,0011	0,2209
						351	0,7982	0,4588	0,0004	0,0007	0,3359
						382	0,9833	0,9660	-0,0001	0,0010	0,1022
						392	0,9144	0,8346	0,0001	0,0011	0,0876
						396	0,9885	0,9764	-0,0002	0,0013	0,1456
						399	0,9664	0,9241	0,0007	0,0014	0,5357
						400	0,9937	0,9874	-0,0001	0,0010	0,1081

A Tabela 18 mostra os resultados obtidos pelas RNAs dos *clusters* 23 e 24. As RNAs de ambos os *clusters* atingiram valores de correlação e NSE acima de 0,7 e 0,5, respectivamente, para todos os poços.

Na Tabela 19 estão os resultados para as RNAs dos *clusters* 25 e 26. O *cluster* 25 possui apenas 1 poço e sua RNA obteve índices estatísticos com resultados adequados. A RNA do *cluster* 26 não teve valores satisfatórios apenas para o poço 353.

Tabela 18 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 23 e 24 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 23						Cluster 24					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
175	0,9767	0,9491	0,0003	0,0009	0,3074	238	0,9980	0,9955	0,0002	0,0005	0,2388
180	0,7959	0,5497	0,0000	0,0005	0,0428	252	0,9984	0,9967	0,0000	0,0005	0,0155
181	0,8960	0,7984	-0,0001	0,0008	0,0585	344	0,9993	0,9986	0,0000	0,0004	0,0420
298	0,8361	0,6649	0,0001	0,0006	0,0452						
363	0,9816	0,9620	0,0001	0,0005	0,0957						
391	0,9811	0,9606	0,0002	0,0008	0,1762						
394	0,9962	0,9914	-0,0004	0,0008	0,3542						
401	0,9928	0,9848	-0,0001	0,0007	0,0901						

Tabela 19 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 25 e 26 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 25						Cluster 26					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
332	0,9724	0,9435	0,0000	0,0002	0,0427	353	0,3308	-0,1803	-0,0002	0,0007	0,1885
						368	0,9885	0,9767	0,0001	0,0008	0,0831
						375	0,8662	0,7266	0,0000	0,0010	0,0296
						379	0,9919	0,9833	-0,0001	0,0008	0,0570

A Tabela 20 apresenta os valores encontrados para as RNAs dos *clusters* 27 e 28. A RNA do *cluster* 27 alcançou resultados de correlação e de NSE acima de 0,7 e 0,5, na devida ordem, para 13 dos 14 poços do *cluster*. O único poço que não teve resultados adequados foi o poço 229. O *cluster* 28, assim como o *cluster* 25, possui apenas um poço e sua RNA atingiu valores aceitáveis.

Na Tabela 21 estão os resultados encontrados para as RNAs dos *clusters* 29 e 30. Esses dois *clusters* possuem só um poço cada, o que significa que a RNA de cada *cluster* nada mais é do que a RNA apenas do poço individual pertencente ao *cluster*. Para ambos os *clusters* as RNAs alcançaram valores satisfatórios.

Tabela 20 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 27 e 28 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 27						Cluster 28					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
27	0,9083	0,8140	0,0001	0,0006	0,1386	354	0,8318	0,5472	-0,0001	0,0004	0,1093
32	0,8977	0,7700	0,0002	0,0003	0,1491						
38	0,8571	0,7035	0,0000	0,0005	0,0392						
58	0,9978	0,9943	-0,0003	0,0007	0,2850						
61	0,8793	0,7436	0,0002	0,0006	0,1876						
62	0,9289	0,8624	0,0000	0,0009	0,0386						
81	0,9120	0,8230	0,0001	0,0004	0,0660						
92	0,9277	0,8516	0,0001	0,0006	0,1219						
227	0,9769	0,9507	0,0002	0,0007	0,1937						
228	0,9804	0,9600	-0,0002	0,0009	0,1801						
229	0,2497	-0,5391	0,0000	0,0007	0,0068						
266	0,7610	0,5750	0,0000	0,0005	0,0197						
359	0,9901	0,9779	-0,0003	0,0006	0,3155						
376	0,9983	0,9957	-0,0003	0,0006	0,3141						

Tabela 21 – Resultados dos índices estatísticos das RNAs feitas para os *clusters* 29 e 30 aplicadas aos dados da série temporal de cada poço do *cluster*

Cluster 29						Cluster 30					
Poço	R	NSE	BIAS	MAPE	RMSE	Poço	R	NSE	BIAS	MAPE	RMSE
373	0,9585	0,9151	0,0001	0,0003	0,0677	403	0,9913	0,9792	0,0002	0,0004	0,2889

Em relação aos índices estatísticos MAPE e RMSE, as RNAs de todos os *clusters* tiveram resultados considerados bons para esses dois índices. O valor mais alto identificado para o MAPE, entre os poços de todos os *clusters*, e, portanto, o pior resultado, foi 0,0015, esse valor foi encontrado no poço 388 que pertence ao *cluster* 8. No que concerne ao RMSE, o maior resultado foi obtido pela RNA do *cluster* 6 para o poço 305, com um valor de 0,5565.

6.3.2 Comparação entre os resultados obtidos pelas RNAs feitas de modo específico para cada poço com os das RNAs feitas por *cluster*, treinadas com os dados de todos os poços do *cluster*

As Figuras 39 a 42 comparam os resultados obtidos pelas RNAs de cada *cluster*, treinadas com os dados de todos os poços do *cluster*, com as RNAs feitas para cada poço especificamente. Nessas Figuras, para os resultados das RNAs por *clusters*, os poços, mesmo pertencendo a *clusters* diferentes, são apresentados em ordem crescente, de modo a facilitar a contraposição dos resultados com os das RNAs específicas de cada poço.

Em todas as Figuras, a cor azul representa os resultados obtidos pelas RNAs individuais e a cor vermelha os resultados das RNAs dos *clusters*. As Figuras 39 e 40

mostram os resultados de correlação e NSE, respectivamente. Já as Figuras 41 e 42 exibem os valores do MAPE e do RMSE, nesta ordem.

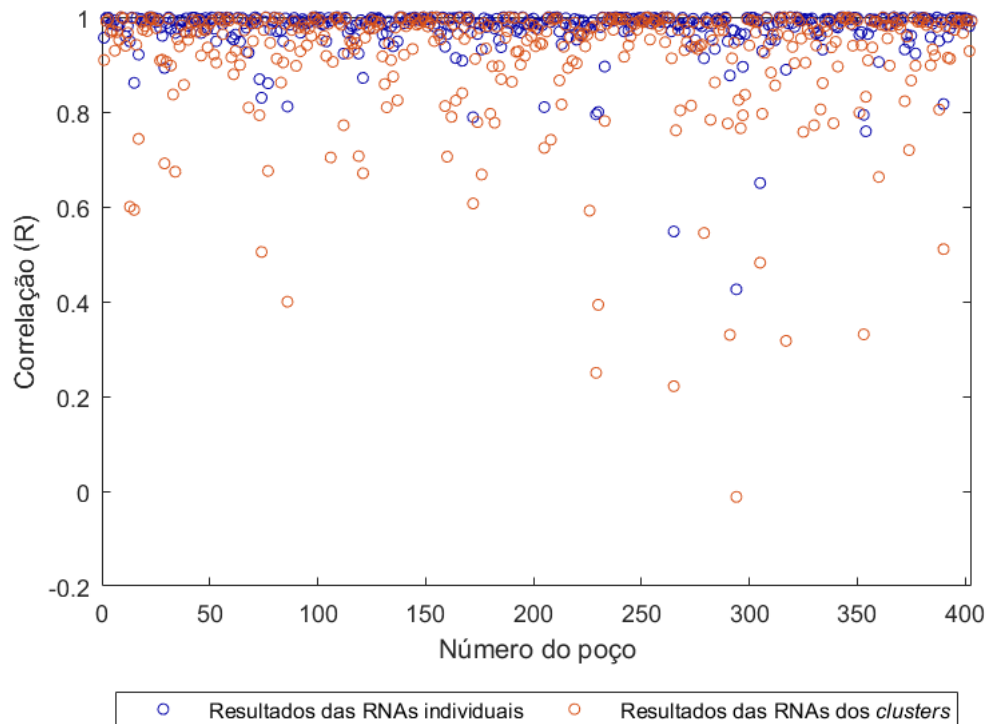


Figura 39 – Comparação dos resultados de correlação obtidos pelas RNAs individuais de cada poço e pelas RNAs dos *clusters*.

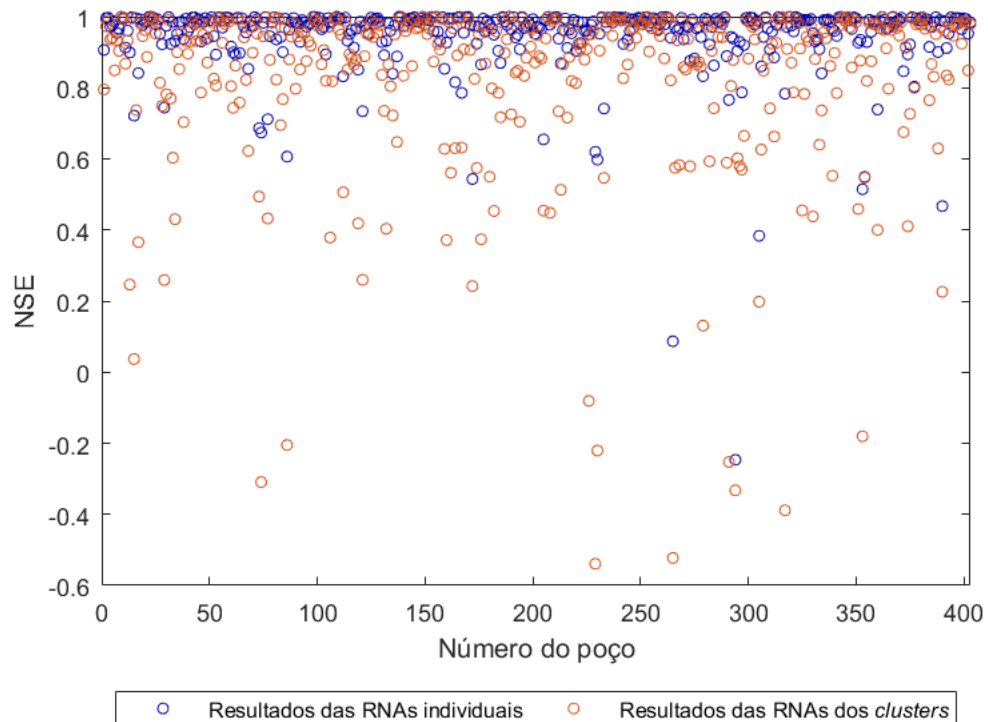


Figura 40 – Comparação dos resultados de NSE obtidos pelas RNAs individuais de cada poço e pelas RNAs dos *clusters*.

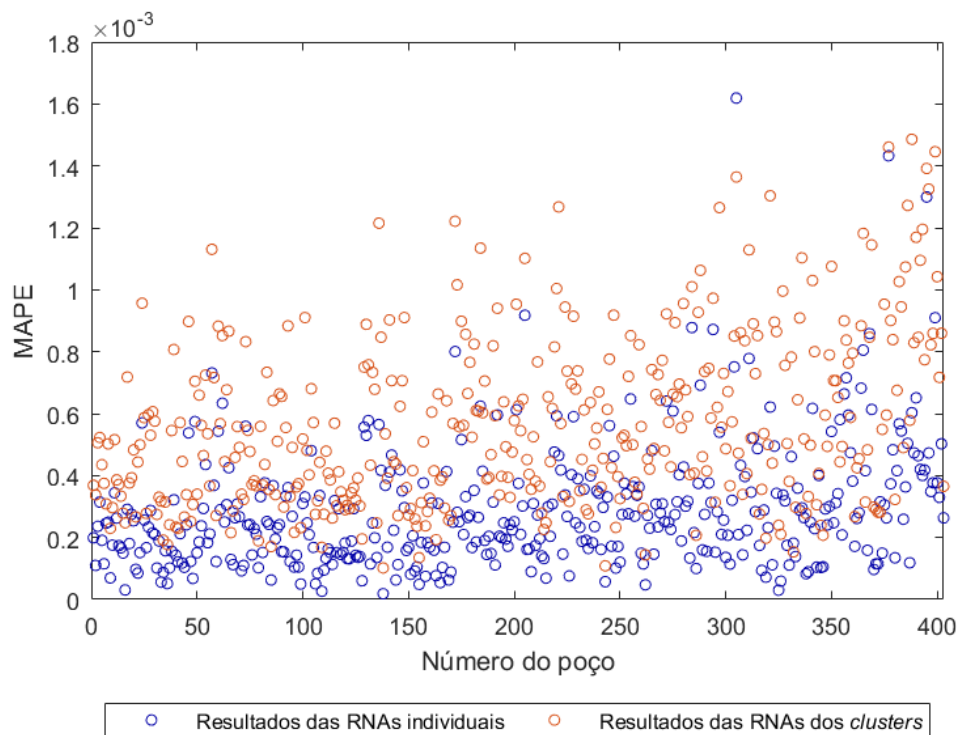


Figura 41 – Comparação dos resultados de MAPE obtidos pelas RNAs individuais de cada poço e pelas RNAs dos *clusters*.

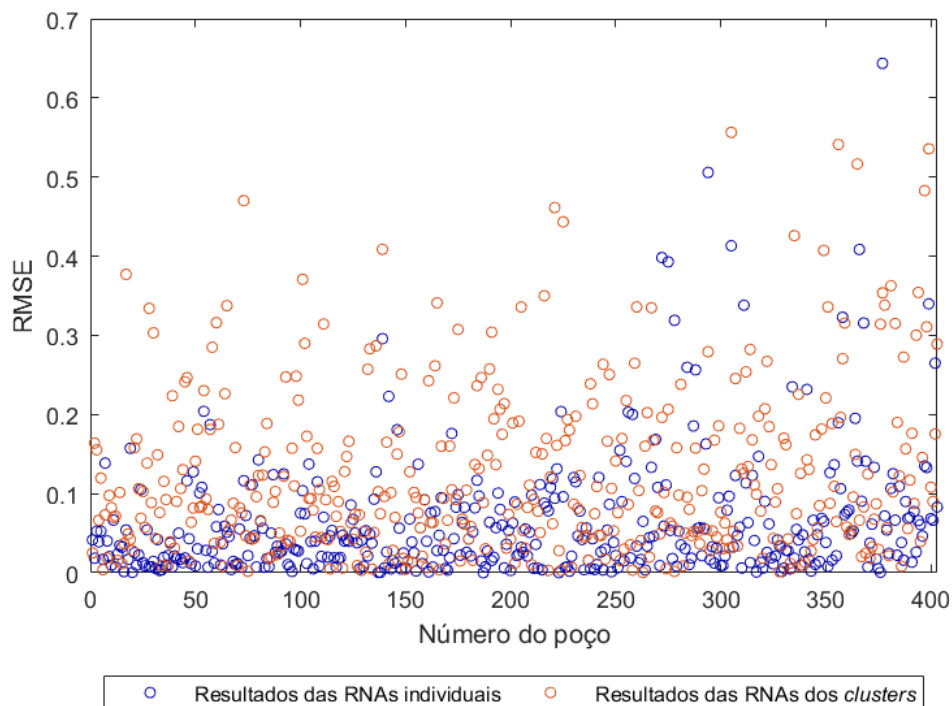


Figura 42 – Comparação dos resultados de RMSE obtidos pelas RNAs individuais de cada poço e pelas RNAs dos *clusters*.

Ao observar as Figuras 41 a 44 é possível notar que, no geral, os resultados dos índices estatísticos obtidos pelas RNAs individuais são melhores do que os das RNAs dos *clusters*. Para a correlação e o NSE, isso pode ser observado pelo fato de que as bolinhas

vermelhas, que simbolizam as respostas das RNAs dos *clusters* para cada poço, encontram-se, na maior parte dos casos, em níveis mais baixos do que os das bolinhas azuis, indicando piores adequações. Tanto para a correlação quanto para o NSE, 99,26% dos poços alcançaram resultados melhores com as RNAs individuais.

No tocante ao MAPE, a grande maioria dos resultados das RNAs individuais, cerca de 98,5%, foram melhores do que os das RNAs por *clusters*. Tal fato pode ser visualizado na Figura 41 pela maior quantidade de bolinhas azuis mais próximas de zero do que de vermelhas. No que se refere ao RMSE, aproximadamente 72% dos poços tiveram resultados melhores com as RNAs individuais. A explicação para resultados melhores nas RNAs dos poços individuais pode ser atribuída ao fato da sua maior especificidade em relação aos dados de aprendizagem do que as RNAs dos *clusters*, que utilizam os dados de todos os poços presentes no *cluster*.

Embora as RNAs individuais tenham se saído melhores do que as RNAs feitas por *clusters*, é importante ressaltar que os resultados das RNAs por *clusters* também foram satisfatórios para a grande maioria dos poços. Além do mais, as RNAs por *clusters* possuem a vantagem de reduzir o volume de cálculos e de RNAs necessárias para a previsão dos níveis d'água, quando comparadas às RNAs individuais de poços não agrupados (MIRZAVAND; GHAZAVI, 2014). Para prever os níveis d'água dos 403 poços, com as RNAs específicas para cada poço, são necessárias 403 RNAs. Já para prever os níveis d'água dos mesmos poços, agrupados por *clusters*, foram utilizadas 30 RNAs, uma redução de cerca de 93% na quantidade de RNAs.

6.3.3 Resultados das previsões dos níveis d'água em cada poço utilizando uma RNA por *cluster*, treinada com os dados de alguns poços do *cluster*

Este é um caso parecido com o do tópico anterior, porém, para esse caso as RNAs feitas para cada *cluster* foram treinadas com os dados de aproximadamente 80% dos poços pertencentes ao *cluster*, de modo a verificar a capacidade de previsão dos níveis d'água subterrânea das RNAs por *clusters*, mesmo quando essa RNA não utilizasse, em sua aprendizagem, os dados de todos os poços do *cluster*.

A Tabela 22 exibe a quantidade de poços que foram retirados de cada *cluster*, por volta de 20%, e quais foram esses poços. Os poços retirados foram os que possuíam as maiores numerações de identificação do poço em cada *cluster*, entretanto, tal retirada se configura como sendo de maneira aleatória, já que esses poços se encontram de modo randômico ao longo do Aquífero Ogallala. O *cluster* 2 e o *cluster* 3 foram os que tiveram a

maior quantidade de poços retirados, sendo 8 e 7 poços, respectivamente. Para os *clusters* que possuem apenas um poço (*clusters* 25, 28, 29 e 30), nada foi modificado.

Tabela 22 – Poços que foram retirados, por *cluster*, antes da realização da aprendizagem da RNA de cada *cluster*

<i>Clusters</i>	Total de poços	Quantidade de poços retirados	Poços retirados
<i>Cluster</i> 1	25	5	358, 364, 370, 371 e 387
<i>Cluster</i> 2	38	8	288, 304, 311, 315, 324, 327, 336 e 395
<i>Cluster</i> 3	37	7	254, 255, 284, 330, 369, 383 e 397
<i>Cluster</i> 4	6	1	316
<i>Cluster</i> 5	20	4	365, 381, 385 e 389
<i>Cluster</i> 6	25	5	305, 307, 318, 352 e 384
<i>Cluster</i> 7	9	2	283 e 333
<i>Cluster</i> 8	27	5	300, 329, 335, 349 e 388
<i>Cluster</i> 9	21	4	301, 347, 380 e 390
<i>Cluster</i> 10	8	2	310 e 372
<i>Cluster</i> 11	11	2	279 e 360
<i>Cluster</i> 12	9	2	343 e 346
<i>Cluster</i> 13	22	4	319, 345, 361 e 366
<i>Cluster</i> 14	8	2	362 e 367
<i>Cluster</i> 15	20	4	320, 325, 355 e 377
<i>Cluster</i> 16	16	3	386, 393 e 398
<i>Cluster</i> 17	8	2	308 e 326
<i>Cluster</i> 18	22	4	322, 337, 338 e 339
<i>Cluster</i> 19	12	2	356 e 378
<i>Cluster</i> 20	11	2	276 e 402
<i>Cluster</i> 21	3	1	222
<i>Cluster</i> 22	12	2	399 e 400
<i>Cluster</i> 23	8	2	394 e 401
<i>Cluster</i> 24	3	1	344
<i>Cluster</i> 25	1	0	Não se aplica
<i>Cluster</i> 26	4	1	379
<i>Cluster</i> 27	14	3	266, 359 e 376
<i>Cluster</i> 28	1	0	Não se aplica
<i>Cluster</i> 29	1	0	Não se aplica
<i>Cluster</i> 30	1	0	Não se aplica

A Tabela 23 mostra os resultados dos índices estatísticos de cada RNA selecionada por *cluster*. Nota-se que os valores alcançados pelas RNAs de cada *cluster* são adequados. O coeficiente de correlação e o NSE tiveram resultados acima de 0,9562 e 0,9110, nesta ordem, para todos os *clusters* analisados. O maior valor encontrado para o MAPE e para o RMSE, e, portanto, pior desempenho da RNA para esses índices, foi 0,0010 e 0,1589, na

devida ordem, ambos no *cluster* 22. Para este caso, não foram feitas análises nos *clusters* 25, 28, 29 e 30, tendo em vista que eles não sofreram nenhuma alteração.

Tabela 23 – Resultados dos índices estatísticos das RNAs selecionadas para cada *cluster*

Clusters Índices	R	NSE	BIAS	MAPE	RMSE
Cluster 1	0,9900	0,9798	-0,0001	0,0006	0,0922
Cluster 2	0,9955	0,9910	0,0000	0,0006	0,0035
Cluster 3	0,9950	0,9901	0,0000	0,0006	0,0001
Cluster 4	0,9949	0,9891	-0,0001	0,0003	0,1463
Cluster 5	0,9927	0,9854	0,0000	0,0005	0,0107
Cluster 6	0,9940	0,9880	0,0000	0,0003	0,0051
Cluster 7	0,9947	0,9893	0,0000	0,0003	0,0202
Cluster 8	0,9917	0,9834	0,0000	0,0005	0,0264
Cluster 9	0,9975	0,9949	0,0000	0,0004	0,0222
Cluster 10	0,9943	0,9885	0,0000	0,0003	0,0493
Cluster 11	0,9845	0,9686	-0,0001	0,0008	0,1135
Cluster 12	0,9752	0,9510	0,0000	0,0005	0,0044
Cluster 13	0,9975	0,9950	0,0000	0,0004	0,0002
Cluster 14	0,9987	0,9974	0,0000	0,0004	0,0159
Cluster 15	0,9953	0,9905	0,0000	0,0004	0,0185
Cluster 16	0,9928	0,9857	-0,0001	0,0007	0,0452
Cluster 17	0,9970	0,9937	0,0001	0,0004	0,1104
Cluster 18	0,9956	0,9913	0,0000	0,0004	0,0096
Cluster 19	0,9932	0,9865	0,0000	0,0004	0,0158
Cluster 20	0,9930	0,9860	0,0000	0,0005	0,0002
Cluster 21	0,9891	0,9778	0,0000	0,0002	0,0615
Cluster 22	0,9839	0,9675	0,0002	0,0010	0,1589
Cluster 23	0,9851	0,9699	0,0000	0,0006	0,0177
Cluster 24	0,9993	0,9985	0,0000	0,0004	0,0203
Cluster 25	-	-	-	-	-
Cluster 26	0,9890	0,9774	0,0001	0,0008	0,0949
Cluster 27	0,9897	0,9795	0,0000	0,0006	0,0289
Cluster 28	-	-	-	-	-
Cluster 29	-	-	-	-	-
Cluster 30	-	-	-	-	-

As Figuras 43 a 68 fazem uma comparação entre os resultados obtidos pelas RNAs dos *clusters* que foram treinadas com os dados de todos os poços com as que foram treinadas com os dados de apenas 80% dos poços presentes no *cluster*. Desse modo, é possível perceber a influência que a retirada dos poços de cada *cluster* teve sobre os resultados de previsão dos níveis d'água.

Em todas as figuras, as bolinhas azuis representam os resultados, de cada índice estatístico, obtidos para cada poço por meio das RNAs de cada *cluster*, treinadas com os

dados de todos os poços pertencentes ao *cluster*. Já as bolinhas vermelhas são os resultados obtidos para cada poço pelas RNAs de cada *cluster*, treinadas com 80% dos poços pertencentes ao *cluster*. A distância vertical entre as bolinhas vermelhas e azuis de cada poço é o que deve ser observado para comparar os resultados das duas RNAs.

Para o *cluster* 1, Figura 43, foi possível observar que os resultados dos índices estatísticos, para a maioria dos poços, foram parecidos tanto para a RNA treinada com os dados de todos os poços quanto para a RNA treinada com 80% deles. Em relação aos índices correlação e NSE, para os poços que foram retirados, apenas dois deles apresentaram mudanças consideráveis no resultado, são eles os poços 364 e 387. Os dois possuem coeficiente de correlação e NSE acima de 0,90 para a RNA treinada com os dados de todos os poços, porém, ao serem retirados, passaram a apresentar valores de 0,82 e 0,75, para a correlação e de 0,64 e 0,48 para o NSE, respectivamente.

Do mesmo modo, a RNA treinada com 80% dos poços também teve menor performance para o índice MAPE nos poços 364 e 387. Quanto ao RMSE, a maior diferença entre os resultados foi para o poço 358, que também é um dos que foram excluídos, com um valor de 0,70. Os outros poços que foram retirados, 358, 370 e 371 continuaram com bons resultados.

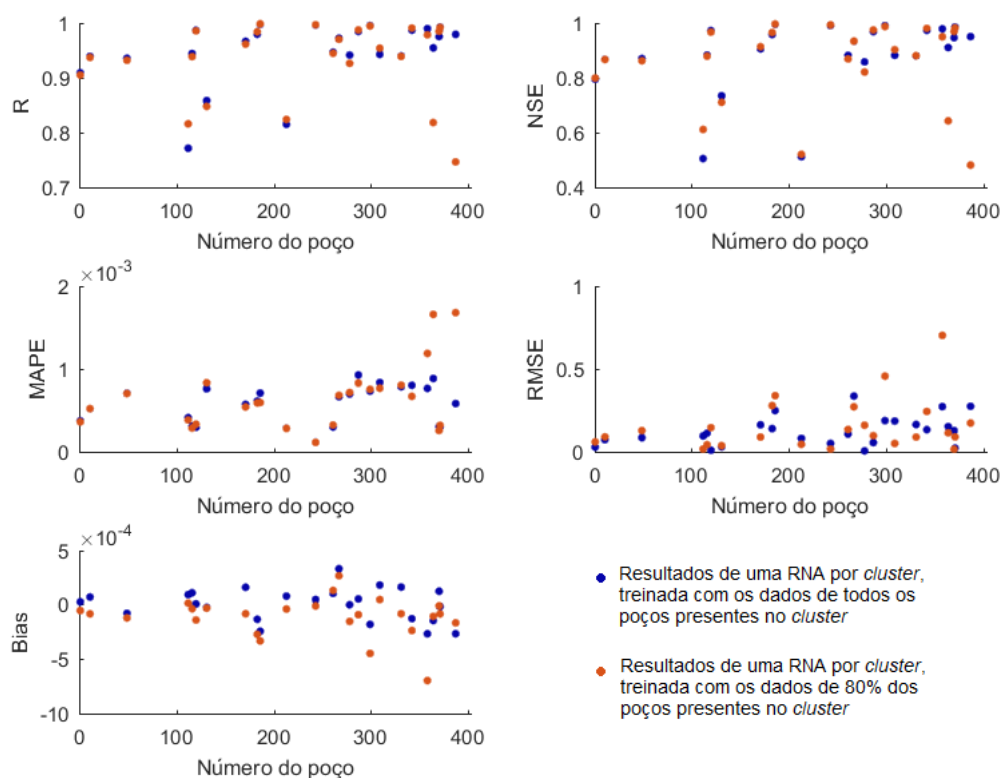


Figura 43 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 1.

A Figura 44 exibe a comparação dos resultados para o *cluster* 2. Para este caso, o único poço que teve resultados discrepantes entre as duas RNAs, para todos os índices estatísticos, foi o 395. Assim, a RNA treinada com 80% dos poços do *cluster* teve uma performance menor em apenas um poço, entre os oito que foram retirados para esse *cluster*.

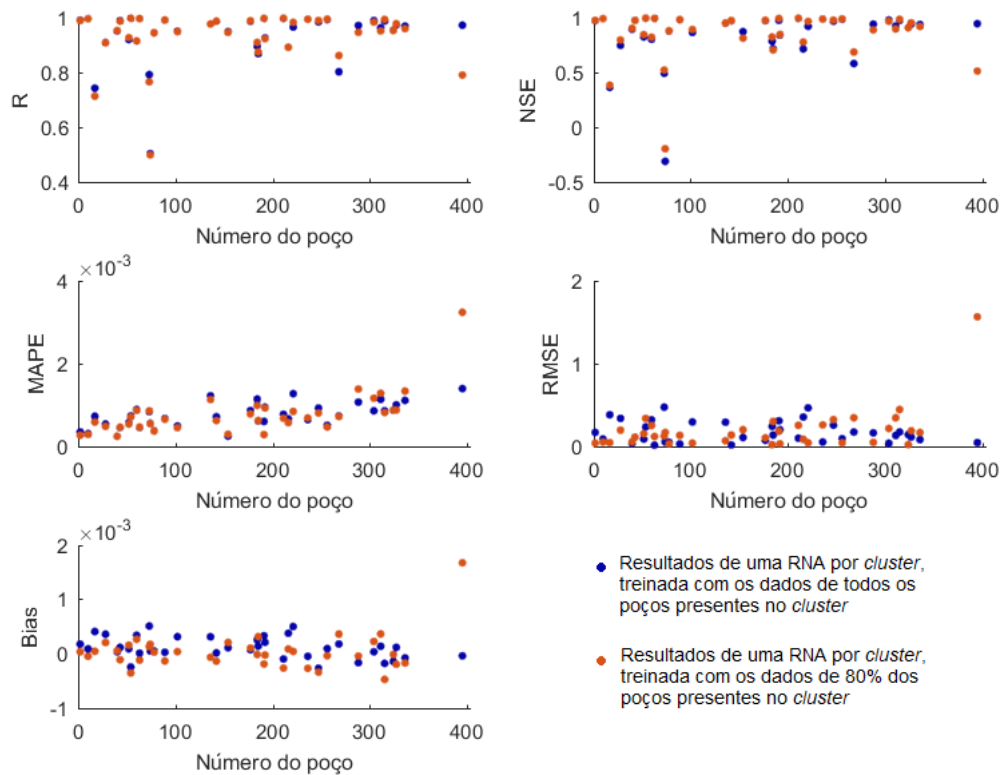


Figura 44 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 2.

Na Figura 45 são comparados os resultados das RNAs para o *cluster* 3. No que diz respeito à correlação, NSE e MAPE, o poço 284 foi o único, entre os sete poços excluídos, que apresentou resultados com grandes diferenças entre a RNA treinada com os dados de todos os poços e a RNA treinada com os dados de 80% deles. Quanto ao RMSE, os poços 255, 284, 369 e 383 foram os que, entre os que foram retirados, tiveram maiores alterações nos resultados.

Para o *cluster* 4, Figura 46, o poço 316 foi o único que foi retirado e que apresentou variações significativas em seus resultados. Os parâmetros que tiveram maiores alterações para o poço 316 foram o MAPE, o RMSE e o Bias. O RMSE foi o que exibiu a maior variação entre todos, passando de um valor de 0,1109, na RNA treinada com todos os poços, para 1,134, na RNA treinada sem esse poço.

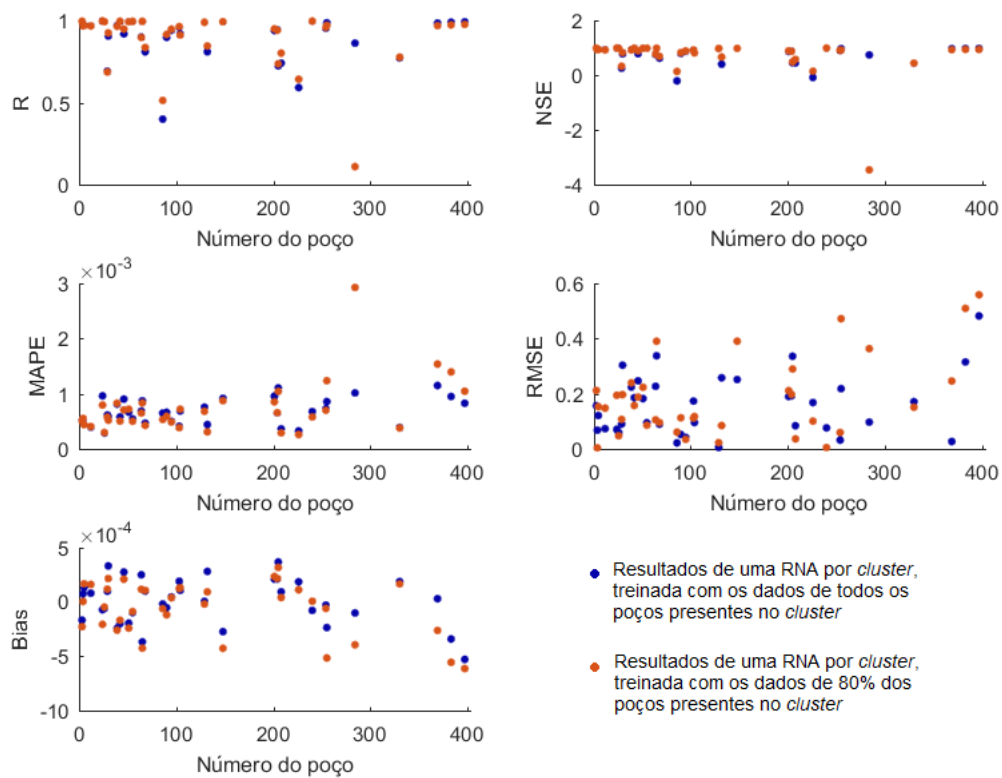


Figura 45 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 3.

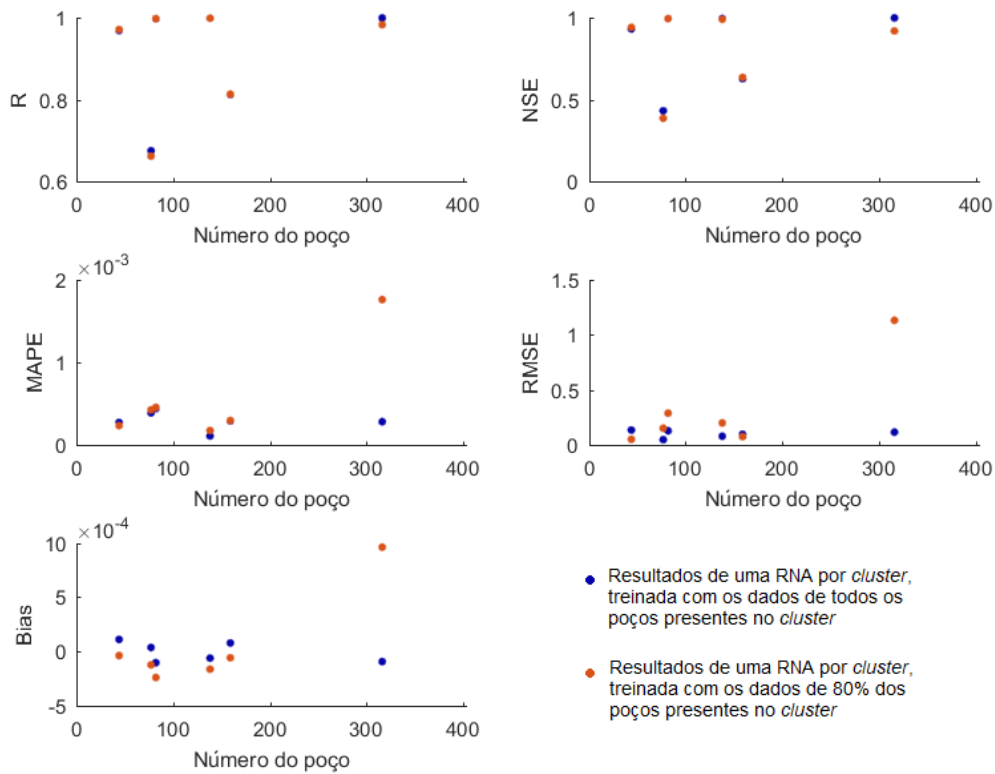


Figura 46 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 4.

A Figura 47 mostra a comparação entre as RNAs para o *cluster* 5. No tocante à correlação, o poço 365 foi o que mostrou maiores mudanças em seus resultados, passando de um valor de 0,8975, na RNA treinada com todos os poços do *cluster*, para 0,6205, na RNA treinada com 80% dos poços. Em relação ao NSE e ao MAPE, a RNA treinada com 80% dos poços obteve uma menor performance para os poços 365 e 381, cujos dados não participaram do aprendizado, entretanto, a mesma teve um resultado de NSE melhor para o poço 291, cujos dados permaneceram na fase de aprendizagem. Para o RMSE, o resultado que mais diferiu foi o do poço 381, com um valor de 4,59 para a RNA treinada com 80% dos poços. Os poços 385 e 389, que também foram retirados para o teste, permaneceram com bons resultados.

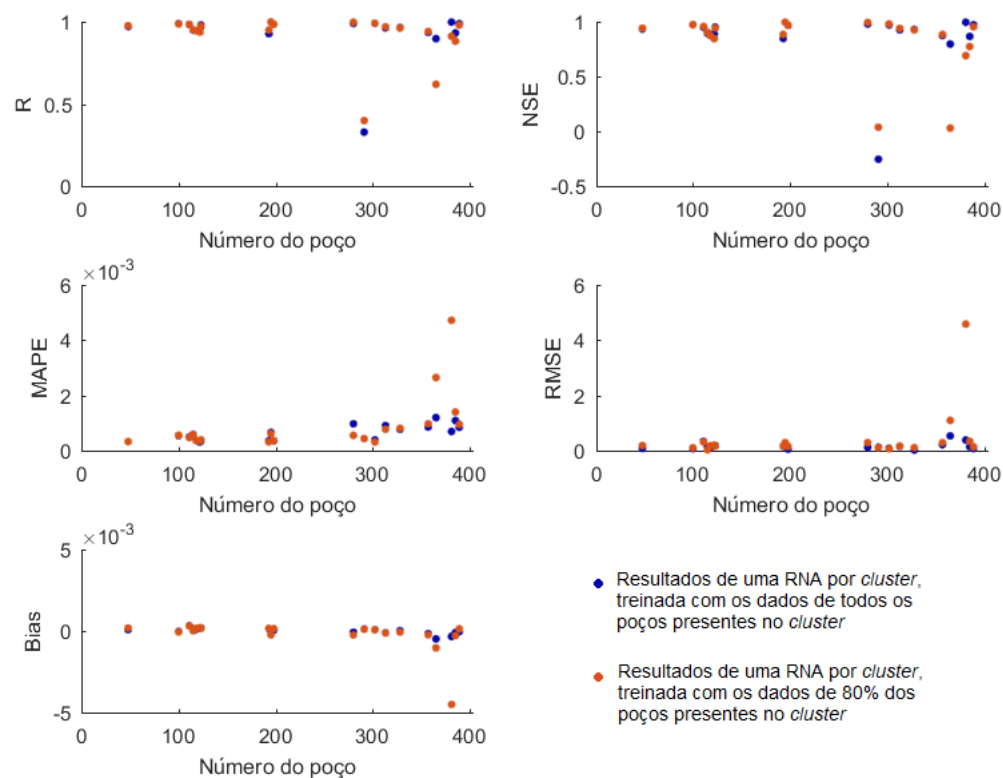


Figura 47 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 5.

A Figura 48 exibe a comparação dos resultados das RNAs para o *cluster* 6. A RNA treinada com 80 % dos poços teve menor desempenho para os índices correlação, NSE, MAPE e Bias do poço 305, para os demais poços os resultados de ambas as RNAs foram semelhantes. Assim, os resultados de quatro dos cinco poços retirados foram satisfatórios.

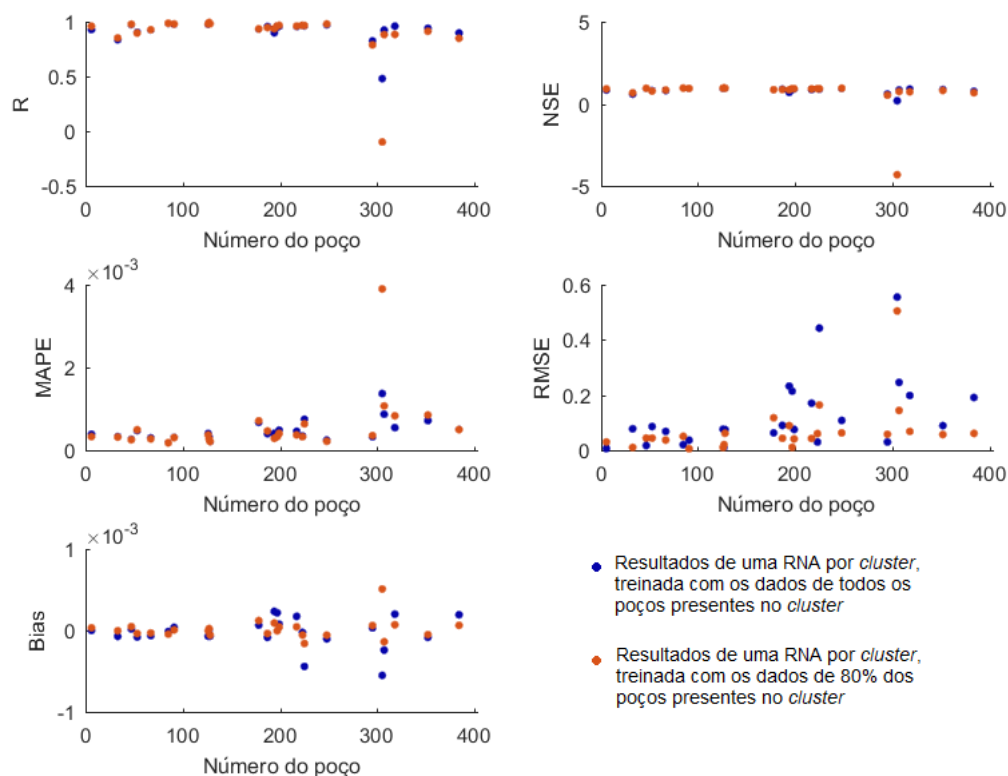


Figura 48 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 6.

A Figura 49 mostra a comparação dos resultados para as RNAs do *cluster* 7. Com relação à correlação e ao NSE, a RNA treinada com 80% dos poços teve uma menor performance para o poço 333. Já para os outros índices estatísticos, os resultados tiveram diferenças significativas para os poços 283 e 333, esses dois poços tiveram os dados removidos na etapa de aprendizagem. Porém, mesmo alcançando uma menor performance para alguns índices, o poço 283 possui resultados considerados adequados.

Ao analisar a comparação dos resultados das RNAs do *cluster* 8, Figura 50, nota-se que, para a correlação e o NSE, a RNA treinada com 80% dos poços teve um menor desempenho para o poço 388. No que diz respeito ao MAPE, os poços 349 e 388 obtiveram resultados mais distantes de zero para a RNA treinada com 80% dos poços. Quanto ao RMSE os resultados que apresentaram maior diferença entre as RNAs foram os dos poços 335, 349 e 388. Para o poço 335 houve uma menor performance para a RNA treinada com 80% dos poços, entretanto, para os poços 349 e 388, a RNA treinada com 80% dos poços teve um melhor resultado. Assim, dentre poços que foram retirados da fase de aprendizagem (300, 329, 335, 349 e 388), os poços 300 e 329 foram os que tiveram resultados mais semelhantes, para todos os índices estatísticos, com os da RNA treinada com os dados de todos os poços.

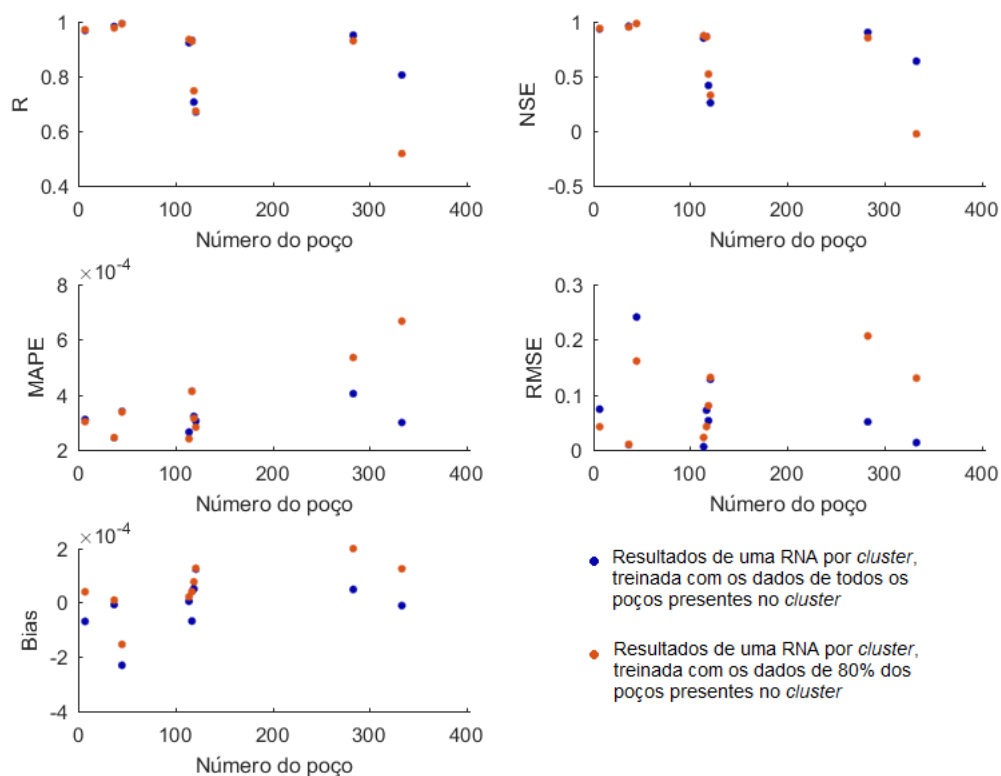


Figura 49 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 7.

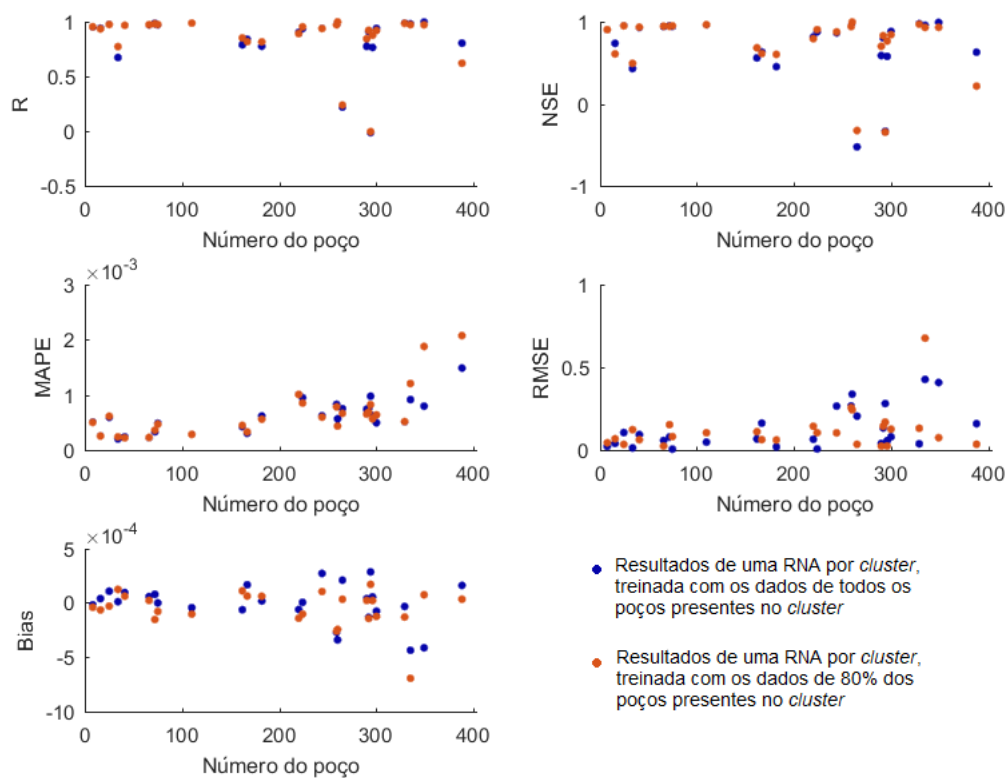


Figura 50 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 8.

Na Figura 51 são exibidos a comparação dos resultados para as RNAs do *cluster* 9. Dos quatro poços que foram retirados, o poço 390 foi o que demonstrou uma maior diferença, para todos os índices, entre os valores encontrados pela RNA treinada com todos os poços e a treinada com 80% deles. Os demais poços, 301, 347 e 380, tiveram bons resultados, entre eles, o poço 380 foi o único que também apresentou uma diferença considerável entre os valores das RNAs para o índice RMSE.

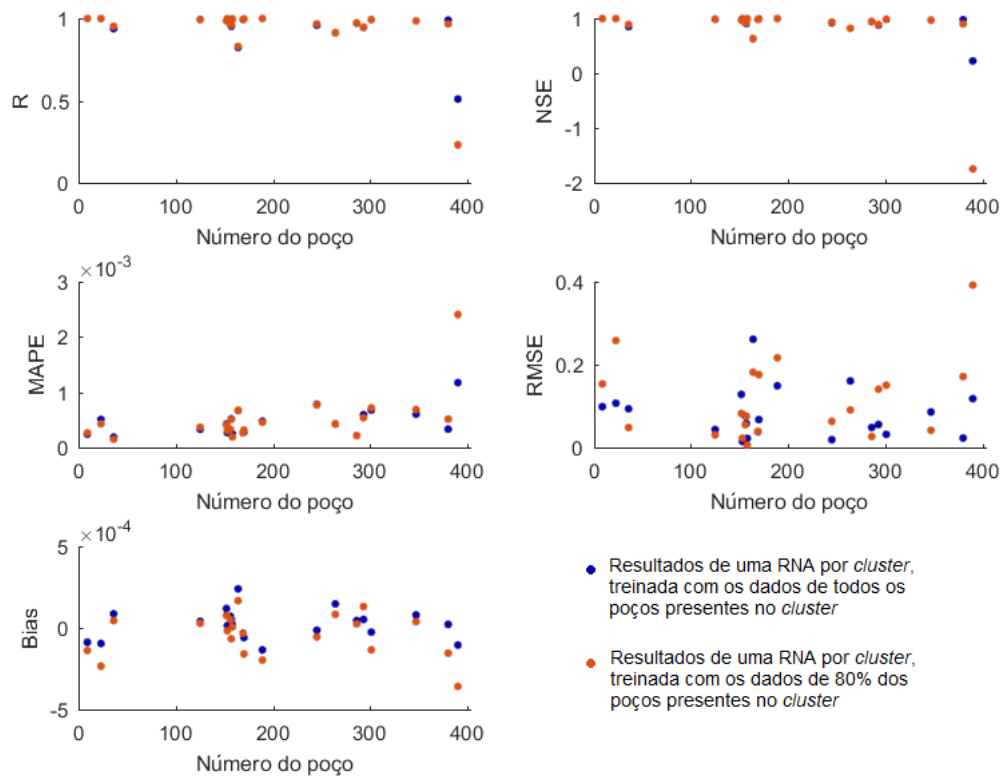


Figura 51 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 9.

Para o *cluster* 10, Figura 52, é possível observar que os dois poços que foram retirados, o poço 310 e o 372, são os que apresentam, para a maior parte dos índices estatísticos, as maiores diferenças entre os resultados das duas RNAs. Do mesmo modo, no *cluster* 11, Figura 53, os poços 279 e 360, cujos dados foram retirados para a RNA treinada com 80% dos poços do *cluster*, foram os que tiveram diferenças mais expressivas entre os resultados das RNAs.

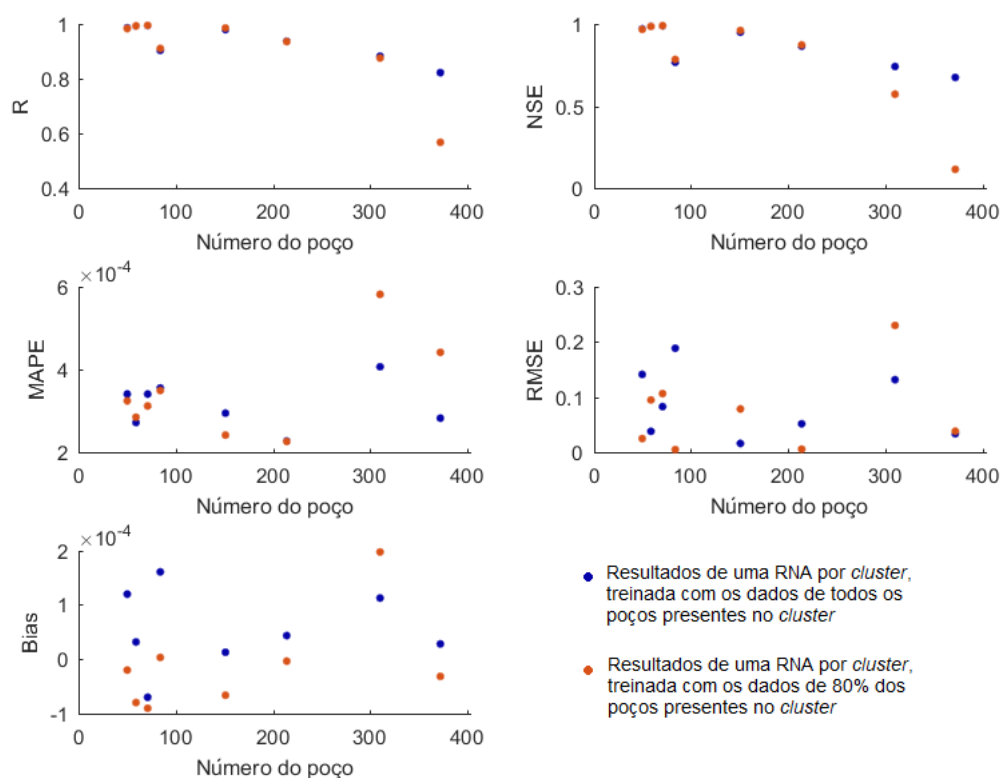


Figura 52 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 10.

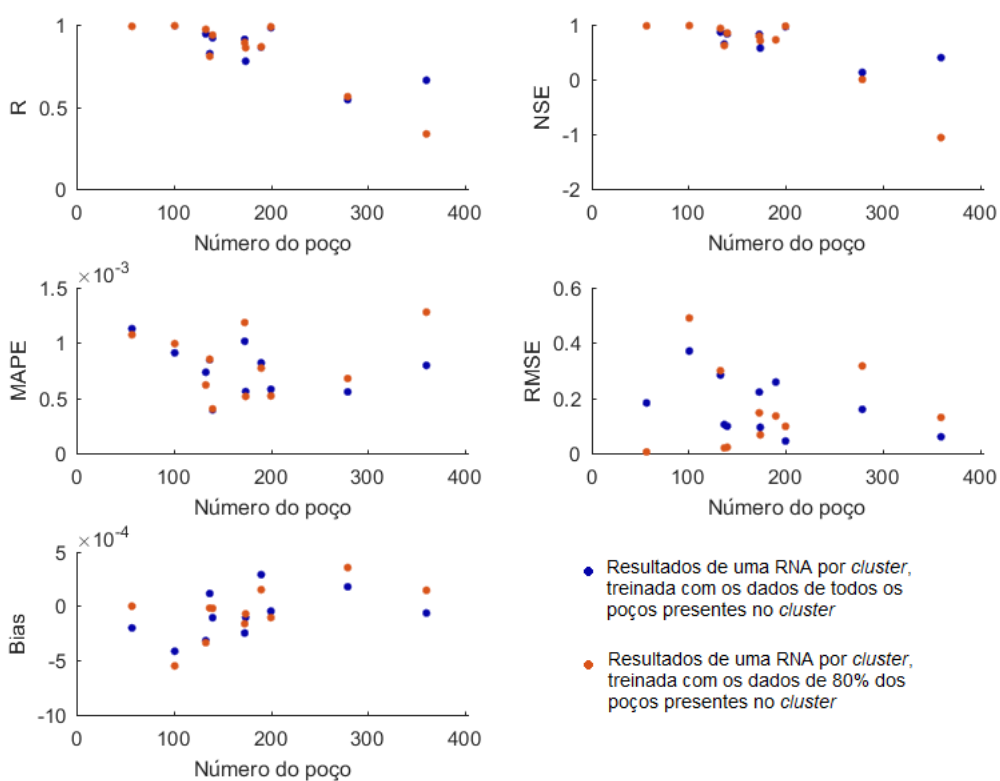


Figura 53 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 11.

A Figura 54 compara os resultados para as RNAs do *cluster* 12. Os poços 343 e 346 sofreram a influência de terem sido retirados e obtiveram um menor desempenho para a RNA treinada com 80% dos poços. Porém, mesmo assim, os resultados alcançados pela RNA, treinada sem esses dois poços, ainda são considerados adequados para ambos.

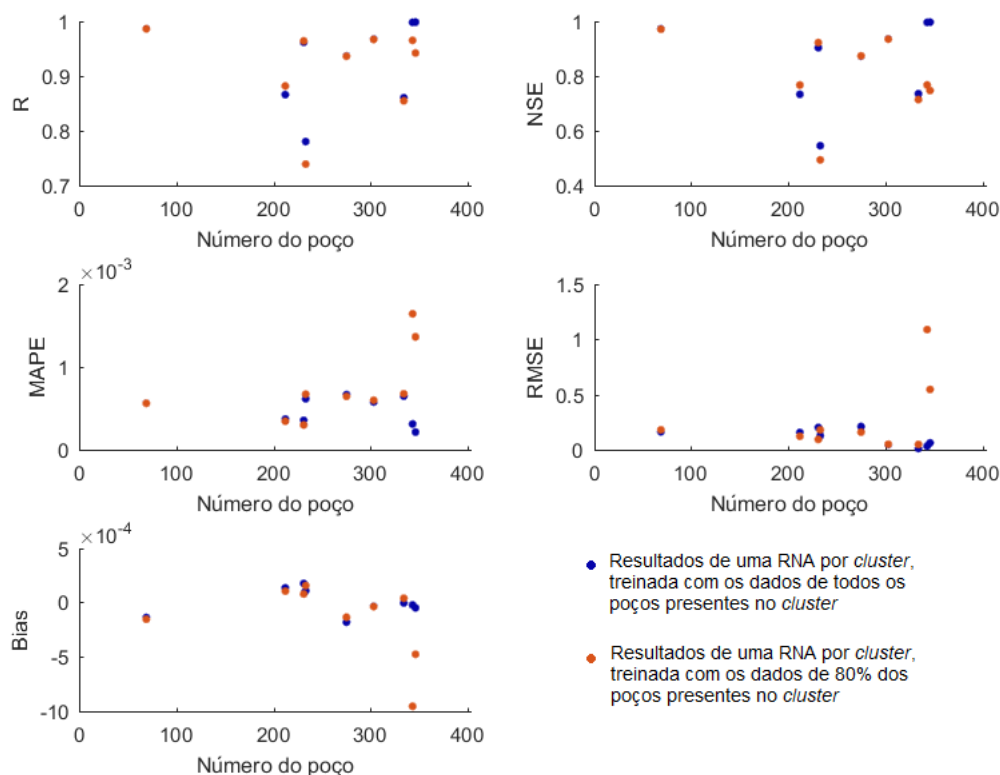


Figura 54 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 12.

Para o *cluster* 13, Figura 55, foram retirados quatro poços, dentre eles, o poço que teve os seus resultados mais prejudicados em virtude disso foi o 345, pois apresentou valores de correlação e NSE mais baixos do que os obtidos pela RNA treinada com os dados de todos os poços e valor de MAPE mais alto. Os outros três poços (319, 361 e 366) obtiveram resultados semelhantes para ambas as RNAs.

A Figura 56 mostra a comparação entre os resultados das RNAs para o *cluster* 14. Para este caso, dois poços foram removidos (poços 362 e 367). Ao analisar os resultados, constata-se que as alterações de valores entre os índices estatísticos das duas RNAs são mínimas. O índice estatístico que apresentou maiores variações foi o RMSE.

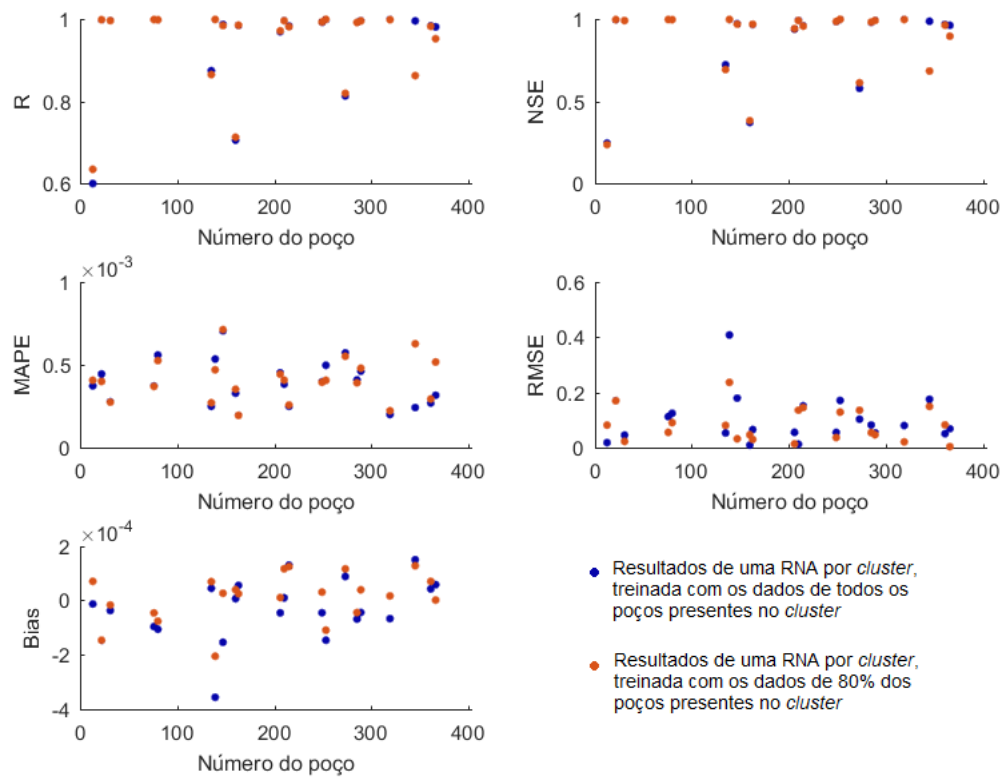


Figura 55 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 13.

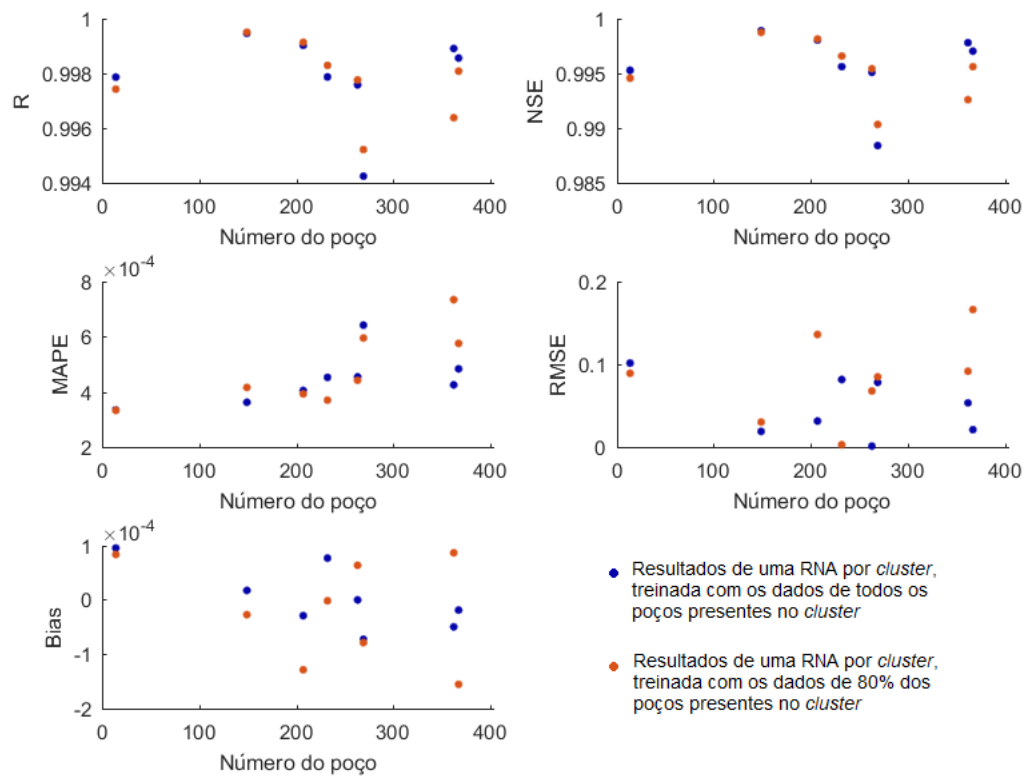


Figura 56 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 14.

A Figura 57 exibe os resultados das RNAs para o *cluster* 15. Para esse *cluster* foram removidos quatro poços. O poço 377 foi um dos que foram retirados e que, portanto, apresentou grandes variações entres os resultados das duas RNAs. Os demais poços permaneceram com valores semelhantes.

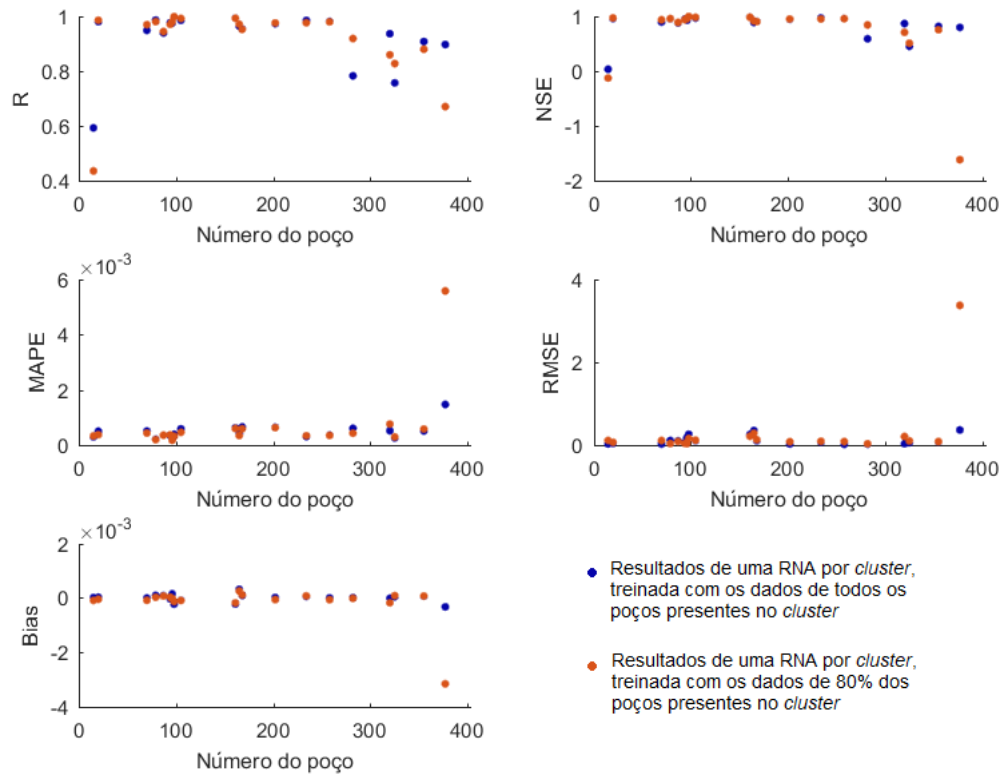


Figura 57 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 15.

Na Figura 58 são apresentadas as comparações entres as RNAs para o *cluster* 16. Dentre os três poços que foram retirados, dois deles tiveram um menor desempenho para a RNA treinada com os dados de 80% dos poços presentes no *cluster*, foram eles os poços 386 e 393. Ambos tiveram grandes mudanças nos resultados entre as duas RNAs, já o poço 398, cujos dados também foram removidos, continuou com resultados aceitáveis.

Para o *cluster* 17, Figura 59, o poço 326 foi o único entres os dois poços retirados que apresentou variações significativas entre os resultados dos índices estatísticos obtidos para a RNA treina com 80% dos poços e a RNA treinada com todos os poços. Assim, esse poço obteve um menor desempenho para a RNA em que seus dados foram removidos.

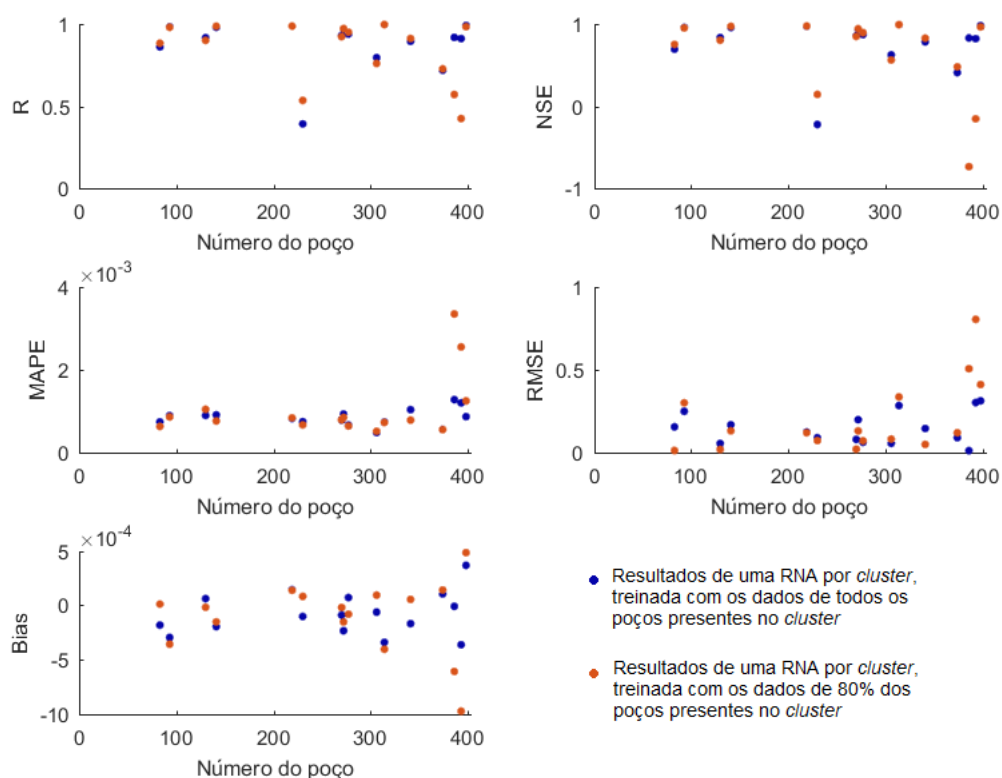


Figura 58 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 16.

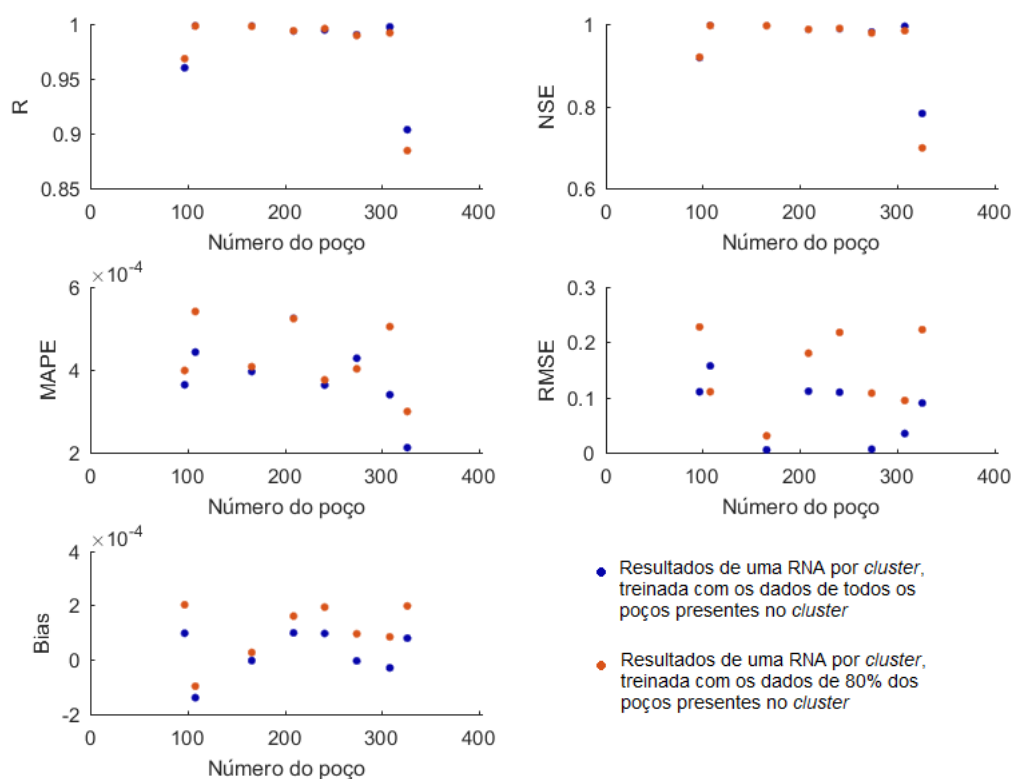


Figura 59 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 17.

A Figura 60 mostra a comparação dos resultados entre as RNAs para o *cluster* 18. Para esse *cluster* foram retirados quatro poços, pela figura constata-se que a RNA treinada com 80% dos poços obteve uma menor performance no poço 339, que foi um dos que tiveram seus dados removidos. Os demais poços apresentaram resultados similares.

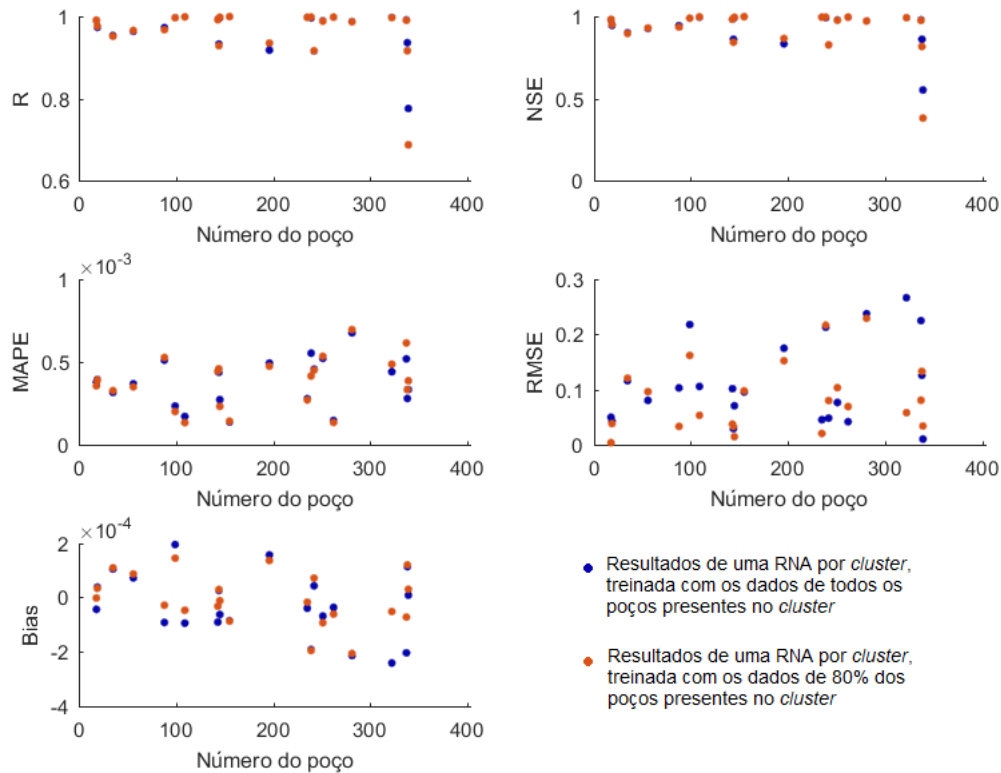


Figura 60 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 18.

A comparação dos resultados das RNAs para o *cluster* 19 é exposta na Figura 61. Para este *cluster*, foram retirados dois poços (356 e 378), desses, o poço 356 foi o que teve maiores modificações nos resultados, a RNA treinada com 80% dos poços apresentou uma performance bem menor para este poço, quando comparada à RNA treinada com todos os poços. Para os demais poços do *cluster*, incluindo o 378, os resultados foram semelhantes para as duas RNAs.

Na Figura 62 é mostrada a comparação dos resultados para as RNAs do *cluster* 20. Para o índice correlação, a RNA treinada com 80% dos poços apresentou um menor desempenho para o poço 276. Quanto ao NSE e ao MAPE, ambos os poços removidos, 276 e 402, tiveram uma performance mais baixa para a RNA treinada com 80% dos poços. Para os índices Bias e RMSE, os valores encontrados pelas duas RNAs foram semelhantes.

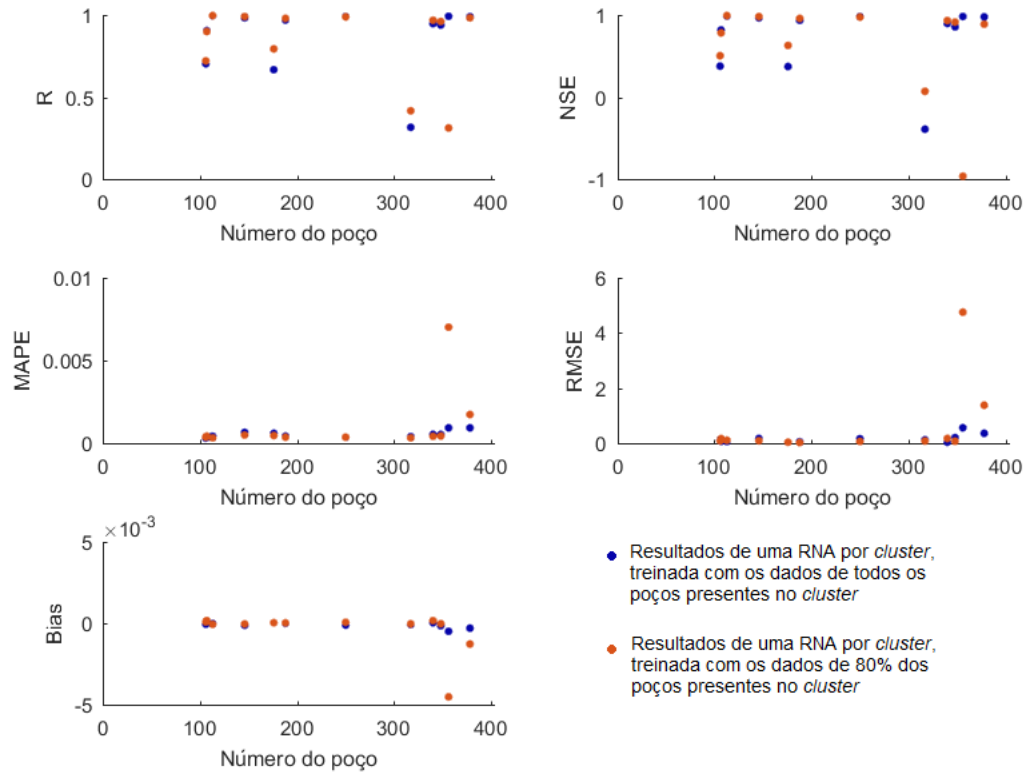


Figura 61 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 19.

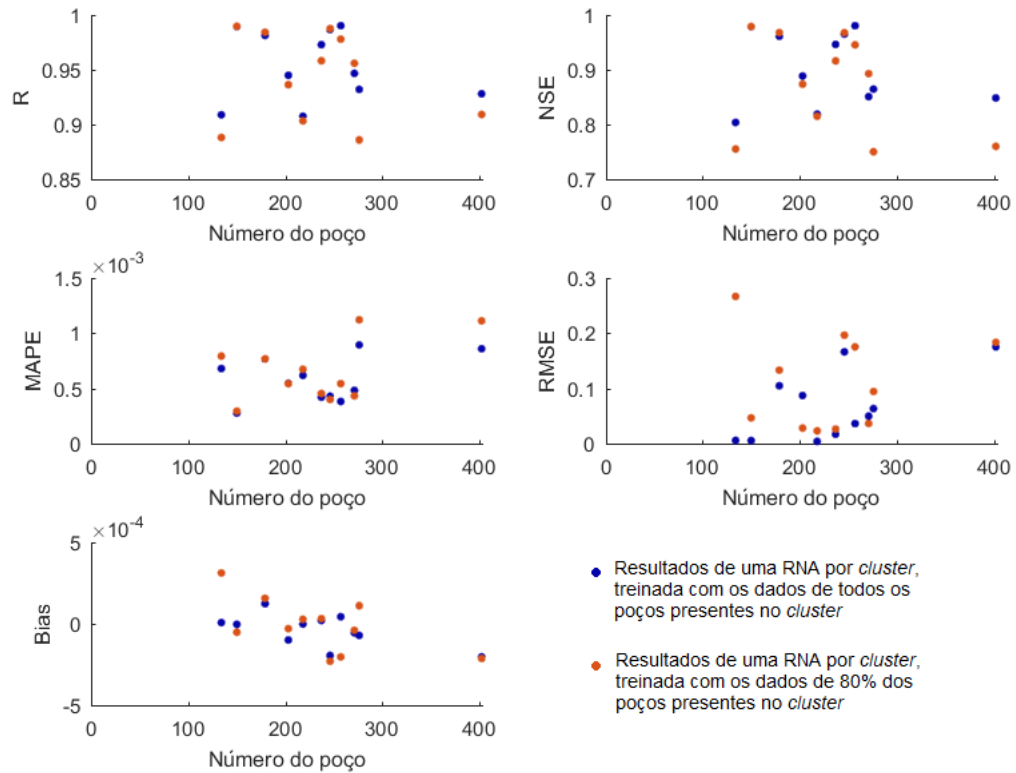


Figura 62 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 20.

Para o *cluster* 21, Figura 63, apenas o poço 222 foi removido. Ao observar a comparação dos resultados obtidos entre a RNA treinada com todos os poços e a RNA em que o poço 222 foi retirado, percebe-se que esta última teve um desempenho bem menor, havendo, assim, uma grande alteração entre os resultados dessas duas RNAs para esse poço.

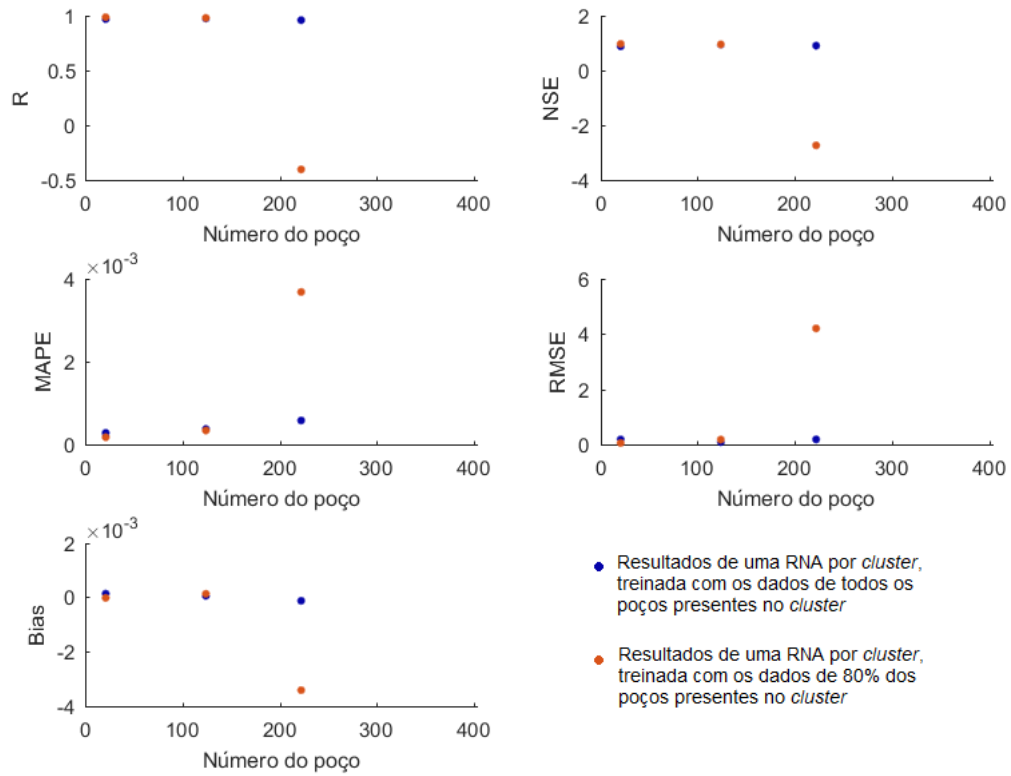


Figura 63 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 21.

Para o *cluster* 22, a comparação dos resultados das duas RNAs é apresentado na Figura 64. Os poços que permaneceram na fase de aprendizado da RNA treinada com 80% dos poços mantiveram valores semelhantes com os da RNA treinada com todos os poços. Contudo, os dois poços que foram removidos (399 e 400) apresentaram diferenças significativas para alguns dos índices estatísticos. O poço 399 teve resultados de NSE, MAPE e RMSE piores para a RNA treinada com 80% dos poços, e o mesmo ocorreu com o índice MAPE para o poço 400.

A Figura 65 exhibe a comparação dos resultados das RNAs para o *cluster* 23. Ao observar a figura, percebe-se que os poços 394 e 401, cujos dados foram retirados do treinamento da RNA treinada com 80% dos poços, apresentaram piores resultados quando comparados aos valores alcançados pela RNA treinada com todos os poços.

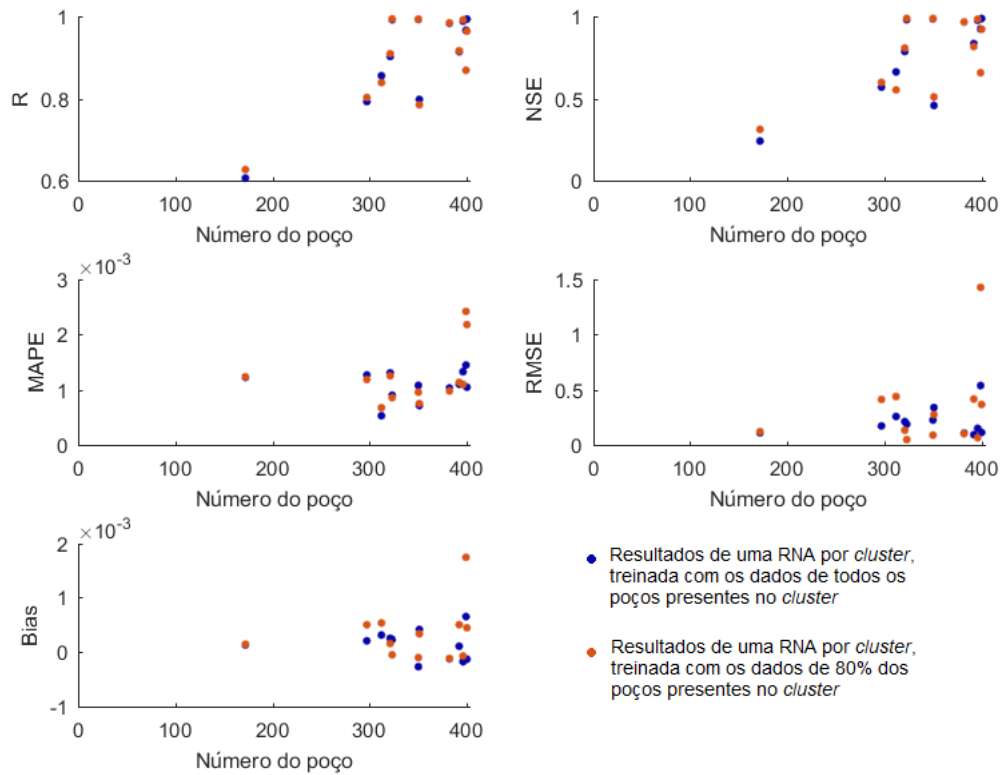


Figura 64 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 22.

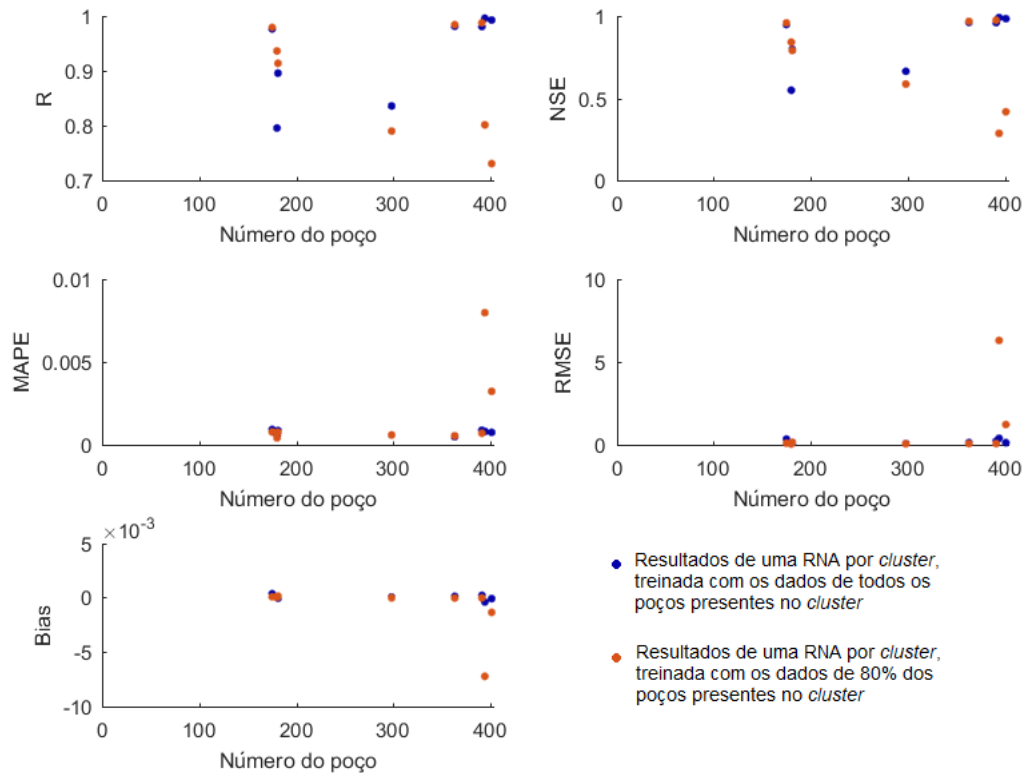


Figura 65 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 23.

Na Figura 66 são apresentados os resultados para ambas as RNAs do *cluster* 24. Por possuir apenas três poços, só o poço 344 foi removido da RNA treinada com 80% dos poços. Ao analisar os resultados, verifica-se que o poço 344, mesmo tendo seus dados retirados da fase de aprendizagem da RNA treinada com 80% dos poços, adquiriu valores bem próximos dos obtidos pela RNA treinada com seus dados. O índice estatístico que mais diferiu entre as duas RNAs foi o RMSE.

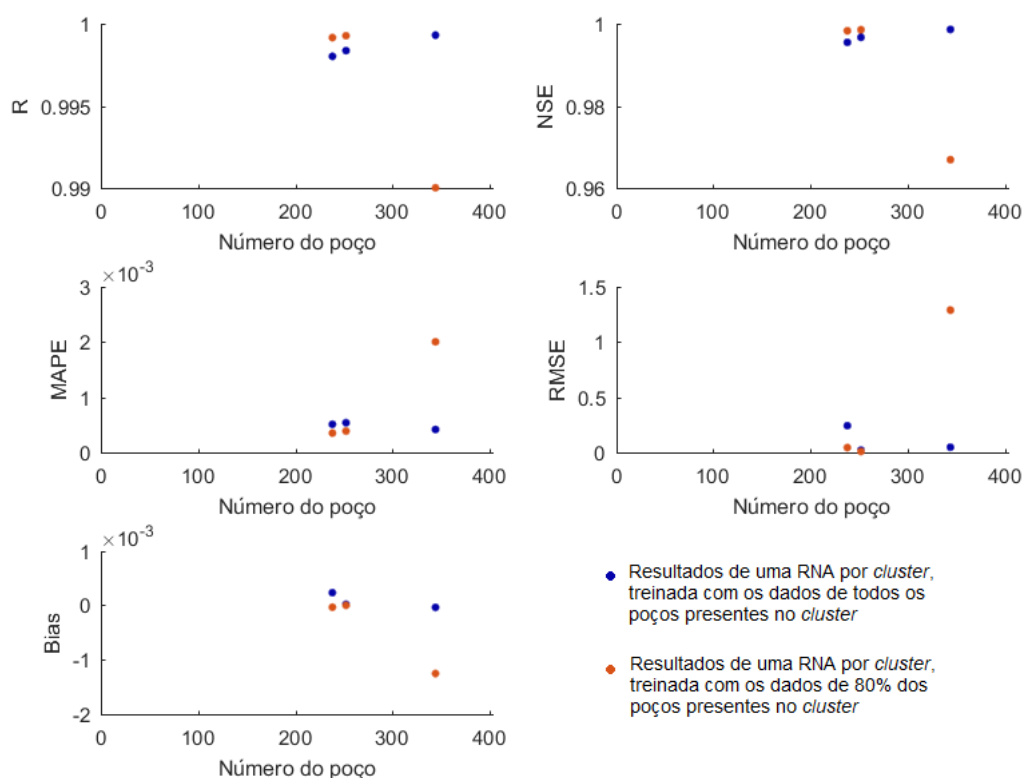


Figura 66 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 24.

A comparação entre os resultados das RNAs do *cluster* 26 é mostrado na Figura 67. Esse *cluster* possui quatro poços e o único que foi removido da fase de aprendizagem da RNA treinada com 80% dos poços foi o poço 379. Na figura é possível observar que o poço 379 obteve bons valores para todos os índices estatísticos.

A Figura 68 apresenta a comparação dos resultados entre as RNAs do *cluster* 27. Para esse *cluster*, três poços foram retirados antes do aprendizado da RNA treinada com 80% dos poços, são eles os poços 266, 359 e 376. Para essa RNA, o poço 266 obteve valores bem mais baixos de correlação e NSE, já o poço 359, teve mudanças significativas nos valores de MAPE e RMSE, enquanto o poço 376 apresentou variações consideráveis apenas no valor do MAPE.

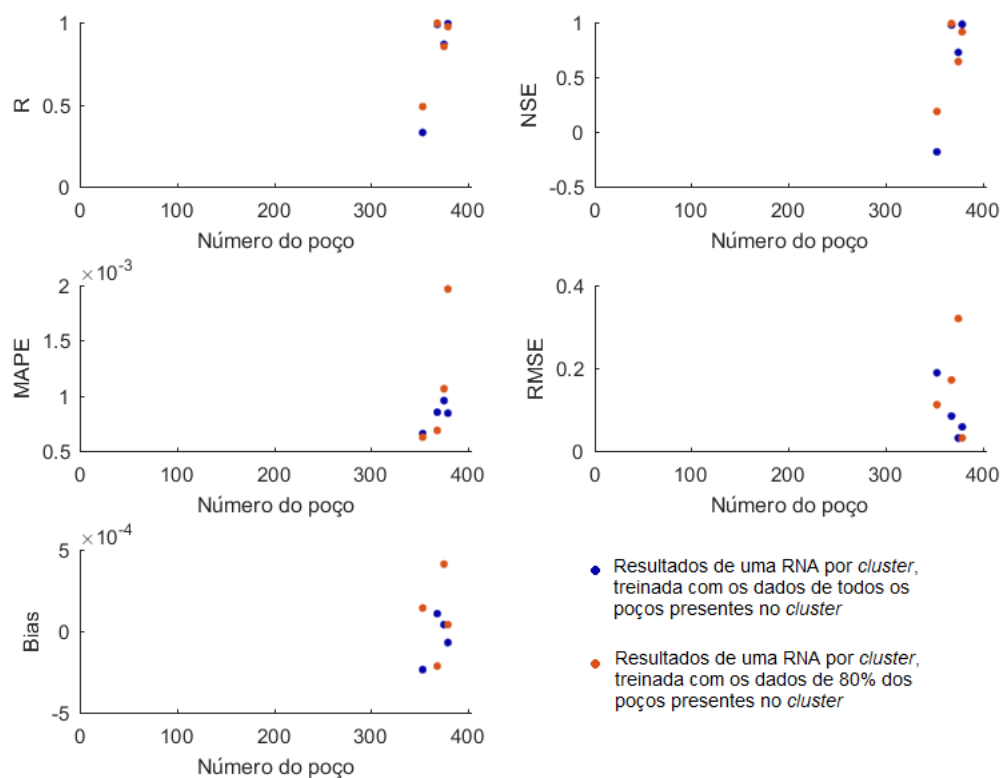


Figura 67 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 26.

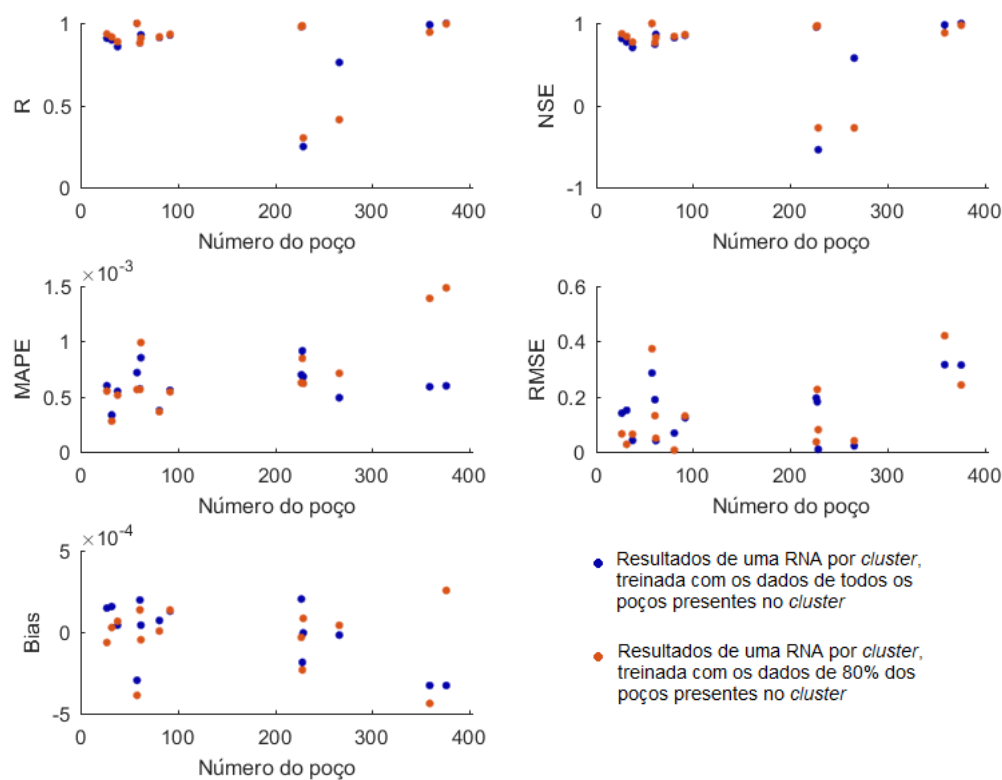


Figura 68 – Comparação dos resultados obtidos pelas RNAs dos *clusters*, treinadas com todos os poços e treinadas com 80% dos poços, para o *cluster* 27.

No geral, pode-se concluir que a remoção de alguns poços de cada *cluster*, antes da fase de aprendizagem da RNA, influencia de modo negativo, na maioria dos casos, as respostas da RNA para as previsões dos níveis d'água nos poços removidos. Já os poços que permanecem para a fase de aprendizagem são pouco influenciados e as alterações em seus valores nas respostas alcançadas pela RNA, treinada com 80% dos poços, deve-se mais ao modo aleatório com que a RNA faz o treinamento dos dados.

Ao todo, foram removidos 80 poços, somando os poços retirados de cada *cluster* para a fase de aprendizagem das RNAs dos *clusters*, treinadas com 80% dos poços presentes em cada *cluster*. Do total de 80 poços, 60 obtiveram resultados considerados bons (poços que possuem simultaneamente correlação acima de 0,7 e NSE acima de 0,5), o que mostra que 75% dos poços que foram removidos tiveram bons resultados com a RNA treinada apenas com os outros poços do *cluster*. Isso indica que há uma relação satisfatória entre os valores de níveis d'água dos poços de cada *cluster*.

A partir do teste que foi feito para observar se as RNAs dos *clusters*, treinadas sem alguns poços, também eram capazes de prever os níveis d'água dos poços que estavam faltando e obtendo para esse teste um resultado positivo, é possível dizer que as RNAs feitas para cada *cluster* também podem ser capazes de prever os níveis d'água de outros poços presentes no Aquífero Ogallala e que porventura venham a ser classificados como fazendo parte de um determinado *cluster*. Assim, cria-se uma ferramenta que pode ser utilizada para aplicações em outros poços que ainda não possuem RNAs específicas para eles ou que não possuem dados suficientes para fazer uma aprendizagem adequada da sua própria RNA.

6.4 Resultados das previsões dos níveis d'água utilizando uma única RNA para todos os poços

Também foi desenvolvida uma única RNA com o intuito de fazer previsões dos níveis d'água subterrânea em qualquer um dos 403 poços estudados no Aquífero Ogallala. A entrada da RNA, para este caso, são os dados dos níveis d'água subterrânea de quatro anos antecedentes, para o poço desejado. A saída do modelo da RNA é o nível da água subterrânea previsto para um ano à frente.

Depois de terem sido feitas 200 RNAs, treinadas com os dados de todos os 403 poços analisados no Aquífero Ogallala, a RNA que obteve melhores índices estatísticos foi escolhida para realizar a previsão dos níveis d'água em qualquer um dos poços estudados. Após isso, a RNA selecionada foi aplicada em cada poço de modo individual a fim de

verificar a sua capacidade de previsão em cada um dos 403 poços. A Figura 69 mostra os resultados dos índices estatísticos, alcançados com o uso dessa RNA, para cada poço.

Ao analisar os resultados da RNA para a correlação, foi verificado que 91,81% dos poços tiveram valores acima de 0,7, ou seja, 33 dos 403 poços não alcançaram valores de correlação considerados adequados. Em referência ao NSE, 49 poços obtiveram resultados menores do que 0,5, por volta de 12% do total. O poço 356 foi o que apresentou os piores resultados tanto para o MAPE quanto para o RMSE, com valores de 0,0025 e 0,9476, na devida ordem.

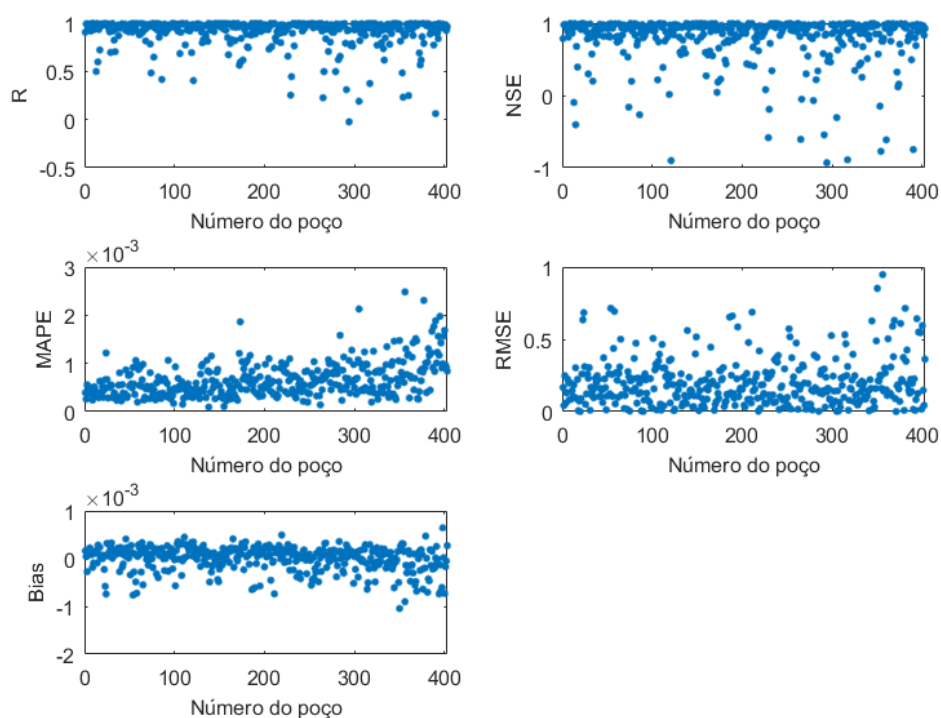


Figura 69 – Resultados dos índices estatísticos utilizando a RNA treinada com os dados de todos os poços.

Os resultados obtidos pela RNA única, para previsão dos níveis d'água em todos os poços estudados no Aquífero Ogallala, mostraram-se adequados para a maioria dos poços. Assim, a RNA treinada com os dados de todos os poços é uma alternativa prática, já que pode ser utilizada para qualquer um dos poços que foram analisados neste estudo, para a realização de previsões de níveis d'água subterrânea.

As Figuras 70 e 71 comparam os resultados obtidos pelas RNAs desenvolvidas de modo específico para cada poço e pelas RNAs feitas para cada *cluster*, treinadas com os dados de todos os poços do *cluster*, com os resultados obtidos pela RNA única desenvolvida para todos os poços, respectivamente.

Nas duas figuras, a cor vermelha representa os resultados de cada poço obtidos pela RNA feita para todos os poços. A cor azul, na Figura 70, simboliza os resultados das RNAs feitas especificamente para cada poço, e a cor preta, Figura 71, ilustra os valores alcançados pela RNA por *cluster*, treinada com os dados de todos os poços presentes no *cluster*.

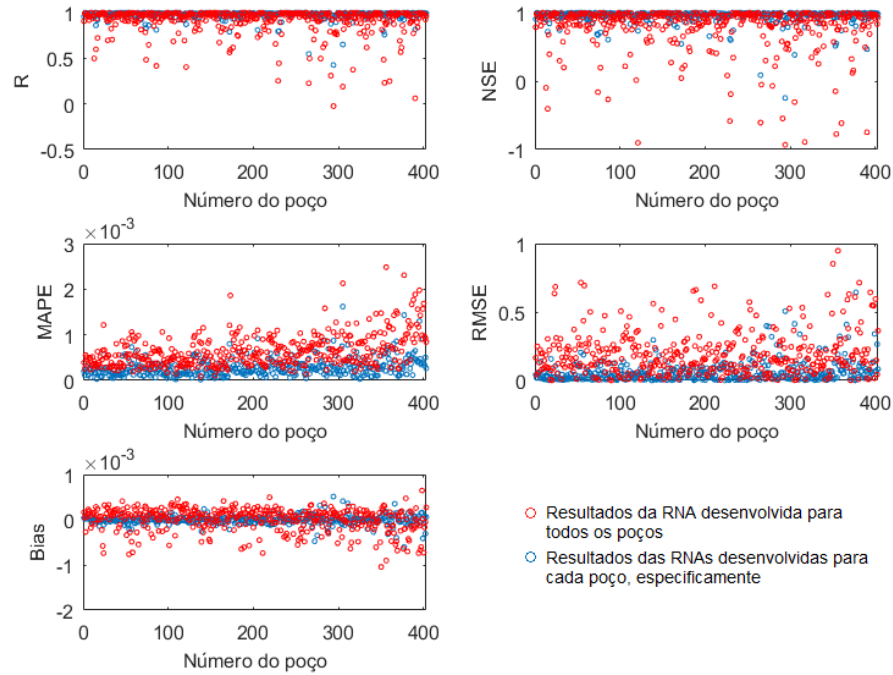


Figura 70 – Comparação dos resultados obtidos pela RNA desenvolvida para todos os poços e pelas RNAs específicas de cada poço.

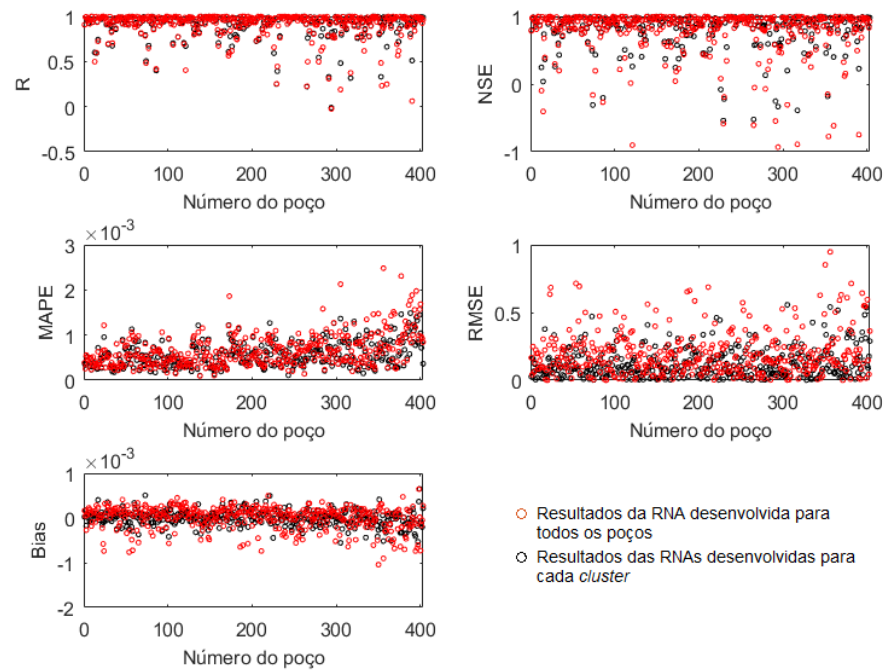


Figura 71 – Comparação dos resultados obtidos pela RNA desenvolvida para todos os poços e pelas RNAs feitas para cada *cluster*, treinadas com os dados de todos os poços do *cluster*.

Ao analisar a Figura 70, repara-se que os resultados obtidos pelas RNAs específicas de cada poço foram, na maioria dos casos, melhores do que os resultados da RNA desenvolvida para todos os poços. Isto pode ser verificado pelo fato de que as bolinhas vermelhas para a correlação e para o NSE se encontram com valores menores do que os vistos nas bolinhas azuis. Para os índices correlação e NSE, todos os resultados obtidos pela RNA específica de cada poço foram melhores do que os da RNA desenvolvida para todos os poços.

Já os resultados do MAPE e do RMSE, alcançados pela RNA desenvolvida para todos os poços, estão localizados em níveis mais altos, no geral, do que os resultados das RNAs específicas de cada poço, indicando piores adequações da RNA desenvolvida para todos os poços. Em relação ao índice MAPE, o único poço que teve resultado melhor com a RNA desenvolvida para todos os poços foi o 258. No que diz respeito ao RMSE, apenas 65 dos 403 poços obtiveram resultados melhores para a previsão do nível d'água ao utilizar a RNA desenvolvida para todos os poços.

A Figura 71 compara os resultados da RNA desenvolvida para todos os poços com os resultados das RNAs por *cluster*, treinadas com os dados de todos os poços presentes no *cluster*. Para os índices correlação e NSE, 56% e 72,7% dos poços, respectivamente, tiveram resultados mais altos com as RNAs desenvolvidas para cada *cluster*. Quanto ao MAPE e ao RMSE, 70,22% e 65,76% dos poços, nesta ordem, alcançaram os melhores resultados com as RNAs desenvolvidas por *clusters*. Assim, é possível concluir que, na maioria dos casos, a RNA desenvolvida por *clusters* teve uma melhor performance do que a RNA desenvolvida para todos os poços.

7 CONCLUSÕES E RECOMENDAÇÕES

O Aquífero Ogallala destaca-se por sua importância para o abastecimento de água agrícola e pública, tendo sustentado o desenvolvimento econômico dos EUA por mais de 80 anos. Porém, assim como vários outros aquíferos do mundo, o Ogallala vem sofrendo com os impactos da demanda insustentável sobre os seus recursos hídricos. Nesse contexto, a previsão dos níveis de água subterrânea aparece como um importante fator que pode ajudar a garantir o uso sustentável dos recursos hídricos subterrâneos.

Neste estudo, foram desenvolvidos três modelos de RNA para realizar a previsão dos níveis de água subterrânea em poços localizados na porção do Aquífero Ogallala que se

localiza no estado do Texas. O primeiro modelo consistiu em desenvolver 403 RNAs, uma para cada poço analisado. A entrada das RNAs para esse modelo são os dados anuais, de nível d'água nos poços, dos quatro anos anteriores ao horizonte de previsão desejado (de um a dez anos à frente). Desse modo, a saída da RNA é o nível d'água para um ou até dez anos à frente.

Para desenvolver o segundo modelo de RNAs, foi empregada a análise hierárquica de agrupamentos, de modo a dividir os 403 poços analisados em *clusters* com semelhantes flutuações dos níveis d'água. Assim, foram selecionados trinta *clusters* e para cada um deles foi desenvolvida uma RNA. A RNA desse modelo utilizou durante a sua fase de aprendizagem os dados de todos os poços pertencentes ao seu correspondente *cluster*. A camada de entrada, para cada uma das trinta RNAs desse caso, consiste nos dados anuais, de níveis d'água no poço, para quatro anos precedentes e a saída é a previsão do nível d'água para um ano à frente.

A partir do modelo acima, foi feito um teste para avaliar a capacidade de previsão das RNAs por *clusters*, treinadas sem parte dos poços do *cluster*. Assim, foram feitas trinta RNAs, uma para cada *cluster*, porém, antes da fase de aprendizagem das RNAs, foram retirados cerca de 20% dos poços que faziam parte de cada *cluster*.

O último modelo desenvolvido, definiu uma única RNA para realizar as previsões dos níveis d'água em todos os 403 poços analisados no Aquífero Ogallala. Desse modo, a RNA foi treinada com os dados pertencentes a todos os poços estudados. A entrada dessa RNA são os dados anuais, de nível d'água no poço desejado, de quatro anos precedentes e a saída é a previsão para um ano à frente.

O desempenho, de todos os modelos de RNA desenvolvidos nesse trabalho, foi avaliado a partir de índices estatísticos. Ao comparar os resultados obtidos entre as RNAs desenvolvidas de modo específico para cada poço, as RNAs desenvolvidas por *clusters*, treinadas com os dados de todos os poços presentes no *cluster*, e a RNA única, desenvolvida para todos os poços em conjunto, observou-se que a última alternativa foi a que apresentou, no geral, o menor desempenho para a previsão dos níveis d'água em poços, refletido pelos valores dos índices estatísticos. A avaliação dos resultados de previsão demonstrou que a RNA desenvolvida de modo específico para cada poço foi a que obteve, na maior parte dos casos, melhores previsões para os níveis d'água de cada poço.

Assim, conclui-se que o modelo que possui 403 RNAs, uma para cada poço, foi o que obteve as melhores previsões dos níveis d'água quando comparado aos demais. Entretanto, esse modelo possui como desvantagem o fato de ser muito específico e necessitar

de muitas RNAs. Dessa forma, as RNAs desenvolvidas por *clusters*, podem ser uma alternativa adequada para realizar as previsões, pois obtiveram, no geral, bom desempenho.

Além disso, ao analisar a capacidade de previsão das RNAs por *clusters*, treinadas sem parte dos poços presentes no *cluster*, observou-se que tais RNAs podem ser capazes de realizar previsões de modo satisfatório, nos poços de seu próprio *cluster*, que, a princípio, não foram utilizados durante a fase de aprendizagem da RNA. Isso indica que há a possibilidade das RNAs por *clusters* serem capazes de prever o nível d'água de outros poços do Aquífero Ogallala que, porventura, possuam flutuações de níveis d'água semelhantes aos intervalos de flutuações de algum dos trinta *clusters*.

Já a RNA feita para todos os poços do Aquífero Ogallala, embora tenha obtido um desempenho menor do que o dos modelos anteriores, também se mostrou capaz de, no geral, alcançar resultados considerados satisfatórios. Essa RNA possui como vantagem a praticidade, por ser uma única RNA que atende a todos os poços.

As sugestões que podem ser feitas para possíveis futuros trabalhos, quanto aos modelos de RNAs aqui apresentados, seria modificar a arquitetura das RNAs por *clusters*, como por exemplo a quantidade de neurônios na camada escondida e de nós na camada de entrada. De modo a observar se uma arquitetura diferente para a RNA de cada *cluster* seria mais adequada ou se uma arquitetura diferente da que foi apresentada nesse trabalho, porém, igual para todas as RNAs de cada *cluster*, obteria resultados melhores do que os que foram apresentados.

Sobre a análise de *cluster* realizada com os 403 poços do Aquífero Ogallala, é possível concluir que a análise hierárquica de agrupamentos utilizando a distância euclidiana como métrica de dissimilaridade e o *average linkage* como método de agrupamento foi capaz de criar *clusters* bem definidos tanto em função da flutuação do nível d'água, quanto espacialmente. Os poços foram agrupados de maneira satisfatória de acordo com a flutuação dos níveis d'água, formando *clusters* que possuem poços com grandes similaridades de nível d'água. Além disso, ao analisar a disposição espacial dos *clusters* formados, foi detectado que os *clusters* também possuem um padrão espacial bem definido. Tais fatos evidenciam que a análise de *cluster* empregada foi satisfatória para as finalidades deste estudo.

Uma recomendação que pode ser feita quanto à análise de *cluster*, para ser realizada em trabalhos futuros, seria testar se uma quantidade menor de *clusters* (por exemplo 20 *clusters*) poderia obter resultados igualmente bons ou melhores, reduzindo assim a quantidade de *clusters* e, conseqüentemente, de RNAs que seriam utilizadas.

De acordo com este estudo, pode-se concluir que os modelos de RNAs aqui apresentados são instrumentos eficientes, de modo geral, para a previsão dos níveis de água subterrânea em poços do Aquífero Ogallala. Assim, as ferramentas discutidas nesse trabalho podem ajudar e fornecer orientação aos tomadores de decisão para uma melhor gestão dos recursos hídricos subterrâneos do Aquífero Ogallala.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABU-ALNAEEM, M. F. et al. **Assessment of groundwater salinity and quality in Gaza coastal aquifer, Gaza Strip, Palestine: An integrated statistical, geostatistical and hydrogeochemical approaches study.** Science of the Total Environment, v. 615, p. 972–989, 2018.
- ADAMOWSKI, JAN; CHAN, HIU FUNG. **A wavelet neural network conjunction model for groundwater level forecasting.** Journal of Hydrology, v. 407, n. Junho, p. 28-40, 2011.
- ADHIKARI, P. et al. **Simulating future climate change impacts on seed cotton yield in the Texas High Plains using the CSM-CROPGRO-Cotton model.** Agricultural Water Management, v. 164, p. 317–330, 2016.
- AESCHBACH-HERTIG, W.; GLEESON, T. **Regional strategies for the accelerating global problem of groundwater depletion.** Nature Geoscience, v. 5, n. 12, p. 853–861, 2012.
- AGHABOZORGI, S.; SEYED SHIRKHORSHIDI, A.; YING WAH, T. **Time-series clustering - A decade review.** Information Systems, v. 53, p. 16–38, 2015.
- AHMAD, I. et al. **Artificial neural network modeling for annual peak flows: A case study.** Applied Ecology and Environmental Research, v. 17, n. 3, p. 6917–6935, 2019.
- ALI, M. et al. **Clustering and Classification for Time Series Data in Visual Analytics: A Survey.** IEEE Access, v. 7, p. 181314–181338, 2019.
- ASCE. **Artificial neural networks in hydrology I. Preliminary concepts.** J. Hydrol. Eng. 5 (2), 115–123. 2000a.
- ASCE Task Committee. **Artificial neural networks in hydrology II: hydrologic applications.** Journal of Hydrologic Engineering, 5 (2), 124–132. 2000b
- BERMUDEZ-EDO, M.; BARNAGHI, P.; MOESSNER, K. **Analysing real world data streams with spatio-temporal correlations: Entropy vs. Pearson correlation.** Automation in Construction, v. 88, n. December 2017, p. 87–100, 2018.
- BLOOMFIELD, J. P. et al. **Regional analysis of groundwater droughts using hydrograph classification.** Hydrology and Earth System Sciences, v. 19, n. 10, p. 4327–4344, 2015.
- BUDUMA, N. **Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms.** 1ª edição. United States of America: Editora O'Reilly Media, 2017.
- CANO, A. et al. **Current knowledge and future research directions to link soil health and water conservation in the Ogallala Aquifer region.** Geoderma, v. 328, n. April, p. 109–118, 2018.

CHAUDHURI, S.; ALE, S. **Long term (1960-2010) trends in groundwater contamination and salinization in the Ogallala aquifer in Texas.** Journal of Hydrology, v. 513, p. 376–390, 2014a.

CHAUDHURI, S.; ALE, S. **Long-term (1930-2010) trends in groundwater levels in Texas: Influences of soils, landcover and water use.** Science of the Total Environment, v. 490, p. 379–390, 2014b.

CHEN, Y. et al. **Assessment of alternative agricultural land use options for extending the availability of the Ogallala Aquifer in the Northern High Plains of Texas.** Hydrology, v. 5, n. 4, 2018.

CHRISTOFOLETTI, A. **Modelagem de sistemas ambientais.** 1ª edição. São Paulo: Editora Blucher, 1999.

COLLISCHONN, W.; DORNELLES, F. **Hidrologia para engenharia e ciências ambientais:** 2. ed. Porto Alegre: Associação Brasileira de Recursos Hídricos (ABRH), 336p, 2015.

COPPIN, Ben. **Inteligência Artificial;** tradução e revisão técnica Jorge Duarte Pires Valério. – Rio de Janeiro: LTC, 2010.

DALIAKOPOULOS, I. N.; COULIBALY, P.; TSANIS, I. K. **Groundwater level forecasting using artificial neural networks.** Journal of Hydrology, v. 309, n. 1–4, p. 229–240, 2005.

DE BRITO NETO, R. T. **Avaliação da disponibilidade de água em aquíferos por meio de análises espaço-temporais.** Dissertação (Mestrado em Engenharia Civil e Ambiental) – Centro de Tecnologia, Universidade Federal da Paraíba, João Pessoa, 2012.

DE BRITO NETO, R. T. et al. **Spatial and temporal water-level variations in the Texas portion of the Ogallala Aquifer.** Natural Hazards, v. 80, n. 1, p. 351–365, 2016.

DELNAZ, A.; RAKHSHANDEHROO, G.; NIKOO, M. R. **Confined Aquifer's Hydraulic Parameters Estimation by a Generalized Regression Neural Network.** Iranian Journal of Science and Technology - Transactions of Civil Engineering, v. 44, n. 1, p. 259–269, 2020.

FABBROCINO, S. et al. **Cluster analysis for groundwater classification in multi-aquifer systems based on a novel correlation index.** Journal of Geochemical Exploration, v. 204, n. December 2018, p. 90–111, 2019.

FACELI, K. et al. **Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina.** – Rio de Janeiro: LTC, 2011.

FAUSETT, L. **Fundamentals of neural networks.** Prentice Hall, Englewood Cliffs, N.J., 1994.

GOWDA, P. et al. **Featured Series Introduction: Optimizing Ogallala Aquifer Water Use to Sustain Food Systems.** Journal of the American Water Resources Association, v. 55, n. 1, p. 3–5, 2019.

GUTENTAG, E. D. et al. **Geohydrology of the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming** (USGS, USA, groundwater). US Geological Survey Professional Paper, v. 1400 B, 1984.

GUZMAN, S. M.; PAZ, J. O.; TAGERT, M. L. M. **The Use of NARX Neural Networks to Forecast Daily Groundwater Levels.** Water Resources Management, v. 31, n. 5, p. 1591–1603, 2017.

HANNAH, D. M. et al. **An approach to hydrograph classification.** Hydrological Processes, v. 14, n. 2, p. 317–338, 2000.

HAYKIN, S. **Redes neurais: princípios e prática**; tradução Paulo Martins Engel: 2. ed. – Dados eletrônicos. - Porto Alegre: Editora Bookman, 2007.

HE, L.; AGARD, B.; TRÉPANIER, M. **A classification of public transit users with smart card data based on time series distance metrics and a hierarchical clustering method.** Transportmetrica A: Transport Science, v. 16, n. 1, p. 56–75, 2020.

JEIHOUNI, E. et al. **Potential impacts of climate change on groundwater level through hybrid soft-computing methods: a case study—Shabestar Plain, Iran.** Environmental Monitoring and Assessment, v. 191, n. 10, 2019.

JEUNG, M. et al. **Evaluation of random forest and regression tree methods for estimation of mass first flush ratio in urban catchments.** Journal of Hydrology, v. 575, n. May, p. 1099–1110, 2019.

JIANG, J. JIA et al. **An efficient algorithm for WBAF estimation based on linear interpolation and its estimation error.** Applied Acoustics, v. 142, p. 44–52, 2018.

KOTHARI, K. et al. **Simulation of efficient irrigation management strategies for grain sorghum production over different climate variability classes.** Agricultural Systems, v. 170, n. January, p. 49–62, 2019.

KOUZIOKAS, G. N.; CHATZIGEORGIOU, A.; PERAKIS, K. **Multilayer Feed Forward Models in Groundwater Level Forecasting Using Meteorological Data in Public Management.** Water Resources Management, v. 32, n. 15, p. 5041–5052, 2018.

KRAMER, S. J.; SIEGEL, D. A. **How Can Phytoplankton Pigments Be Best Used to Characterize Surface Ocean Phytoplankton Groups for Ocean Color Remote Sensing Algorithms?** Journal of Geophysical Research: Oceans, v. 124, n. 11, p. 7557–7574, 2019.

KUMAR, S.; TOSHNIWAL, D. **Analysis of hourly road accident counts using hierarchical clustering and cophenetic correlation coefficient (CPCC).** Journal of Big Data, v. 3, n. 1, p. 1–11, 2016.

LALLAHEM, S. et al. **On the use of neural networks to evaluate groundwater levels in fractured media.** Journal of Hydrology, v. 307, n. 1–4, p. 92–111, 2005.

LI, H. et al. **Ground water level prediction for the arid oasis of Northwest China based on the artificial bee colony algorithm and a back-propagation neural network with double hidden layers.** Water (Switzerland), v. 11, n. 4, p. 1–20, 2019.

LIU, Z.; BARAHONA, M. **Graph-based data clustering via multiscale community detection.** Applied Network Science, v. 5, n. 1, 2020.

MACHIWAL, D. et al. **A review of GIS-integrated statistical techniques for groundwater quality evaluation and protection.** Environmental Earth Sciences, v. 77, n. 19, p. 1–30, 2018.

MACKAY, J. D. et al. **Seasonal forecasting of groundwater levels in principal aquifers of the United Kingdom.** Journal of Hydrology, v. 530, p. 815–828, 2015.

MAIER, H. R. et al. **Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions.** Environmental Modelling and Software, v. 25, n. 8, p. 891–909, 2010.

MAREK, G. W. et al. **Assessing planting date effects on seasonal water use of full- and short-season maize using SWAT in the southern Ogallala Aquifer region.** Irrigation Science, v. 38, n. 1, p. 77–87, 2020.

MATLAB, 2017. **The MathWorks Inc.**, Natick, MA.

MCGREE, S.; SCHREIDER, S.; KULESHOV, Y. **Trends and variability in droughts in the Pacific islands and Northeast Australia.** Journal of Climate, v. 29, n. 23, p. 8377–8397, 2016.

MELAKU, N. D.; WANG, J. **A modified SWAT module for estimating groundwater table at Lethbridge and Barons, Alberta, Canada.** Journal of Hydrology, v. 575, n. May, p. 420–431, 2019.

MIRZAVAND, M.; GHAZAVI, R. **A Stochastic Modelling Technique for Groundwater Level Forecasting in an Arid Environment Using Time Series Methods.** Water Resources Management, v. 29, n. 4, p. 1315–1328, 2014.

MITCHELL-MCCALLISTER, D. et al. **Maximizing profits via irrigation timing for capacity-constrained cotton production.** Agricultural Water Management, v. 229, n. November 2019, p. 105932, 2020.

MODALA, N. R. et al. **Climate change projections for the Texas High Plains and Rolling Plains.** Theoretical and Applied Climatology, v. 129, n. 1–2, p. 263–280, 2017.

MOGHADDAM, H. K. et al. **Developing comparative mathematic models, BN and ANN for forecasting of groundwater levels.** Groundwater for Sustainable Development, v. 9, n. January, p. 100237, 2019.

MOHANTY, S. et al. **Comparative evaluation of numerical model and artificial neural network for simulating groundwater flow in Kathajodi-Surua Inter-basin of Odisha, India.** Journal of Hydrology, v. 495, p. 38–51, 2013.

MOJID, M. A.; HOSSAIN, A. B. M. Z.; ASHRAF, M. A. **Artificial neural network model to predict transport parameters of reactive solutes from basic soil properties.** Environmental Pollution, v. 255, p. 113355, 2019.

MUKHERJEE, A.; RAMACHANDRAN, P. **Prediction of GWL with the help of GRACE TWS for unevenly spaced time series data in India: Analysis of comparative performances of SVR, ANN and LRM.** Journal of Hydrology, v. 558, n. August 2002, p. 647–658, 2018.

NATARAJAN, N.; SUDHEER, C. **Groundwater level forecasting using soft computing techniques.** Neural Computing and Applications, v. 5, 2019.

NAYAK, P. C.; SATYAJI RAO, Y. R.; SUDHEER, K. P. **Groundwater level forecasting in a shallow aquifer using artificial neural network approach.** Water Resources Management, v. 20, n. 1, p. 77–90, 2006.

NOURANI, V. et al. **Applications of hybrid wavelet-Artificial Intelligence models in hydrology: A review.** Journal of Hydrology, v. 514, p. 358–377, 2014.

NOURANI, V.; ALAMI, M. T.; VOUSOUGHI, F. D. **Wavelet-entropy data pre-processing approach for ANN-based groundwater level modeling.** Journal of Hydrology, v. 524, p. 255–269, 2015.

NOURANI, V.; ALAMI, M. T.; VOUSOUGHI, F. D. **Hybrid of SOM-clustering method and wavelet-ANFIS approach to model and infill missing groundwater level data.** Journal of Hydrologic Engineering, v. 21, n. 9, p. 1–19, 2016.

PATHAK, A. A.; DODAMANI, B. M. **Trend Analysis of Groundwater Levels and Assessment of Regional Groundwater Drought: Ghataprabha River Basin, India.** Natural Resources Research, v. 28, n. 3, p. 631–643, 2019.

PATTERSON, J.; GIBSON, A. **Deep Learning: A Practitioner's Approach.** 1. ed. Sebastopol: Editora O'Reilly Media, Inc., 2017.

RAJAEI, T.; EBRAHIMI, H.; NOURANI, V. **A review of the artificial intelligence methods in groundwater level modeling.** Journal of Hydrology, v. 572, n. January, p. 336–351, 2019.

RAMESH, R. et al. **Challenges Calibrating Hydrology for Groundwater-Fed Wetlands: a Headwater Wetland Case Study.** Environmental Modeling and Assessment, 2020.

ROSHNI, T. **Neural network modeling for groundwater-level forecasting in coastal aquifers.** Neural Computing and Applications, v. 0123456789, 2020.

RUSSEL, S.; NORVING, P. **Inteligência artificial;** tradução Regina Célia Simille: 3. ed. – Rio de Janeiro: Elsevier, 2013.

SANTOS, C. A. G.; SILVA, G. B. L. DA. **Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models.** Hydrological Sciences Journal, v. 59, n. 2, p. 312–324, 2014.

SANTOS, C. A. G. et al. **Cluster analysis applied to spatiotemporal variability of monthly precipitation over Paraíba state using tropical rainfall measuring mission (TRMM) data.** Remote Sensing, v. 11, n. 6, 2019.

STEWART, D. R.; ALLEN, A. J. **Peak groundwater depletion in the High Plains Aquifer, projections from 1930 to 2110.** Agricultural Water Management, v. 170, p. 36–48, 2016.

STOUT, J. E. **Seasonal water-level perturbations beneath the high plains of the Llano Estacado.** Journal of Hydrology: Regional Studies, v. 18, n. April, p. 1–14, 2018.

SUBBA RAO, N.; CHAUDHARY, M. **Hydrogeochemical processes regulating the spatial distribution of groundwater contamination, using pollution index of groundwater (PIG) and hierarchical cluster analysis (HCA): A case study.** Groundwater for Sustainable Development, v. 9, n. April, p. 100238, 2019.

SUN, Y. et al. Technical note: **Application of artificial neural networks in groundwater table forecasting-a case study in a Singapore swamp forest.** Hydrology and Earth System Sciences, v. 20, n. 4, p. 1405–1412, 2016.

TOROK, S. J. **The development of a high quality historical temperature database for Australia.** Ph.D. thesis, University of Melbourne, 298 pp., 1996.

TUCCI, C. E. M. **Modelos Hidrológicos.** 2ª Edição. Porto Alegre: UFRGS, 2005. 678p.

TUCCI, C. E. M.; SILVEIRA, A. L. L.; **Hidrologia: ciência e aplicação.** Porto Alegre: UFRGS/ABRH, 3ed., 2004, 943p.

TWDB (2016). **Texas Aquifers Study: Groundwater Quantity, Quality, Flow, and Contributions to Surface Water**, 31 de dez. de 2016. Disponível em: < http://www.twdb.texas.gov/groundwater/docs/studies/TexasAquifersStudy_2016.pdf#page=125>. Acesso em: 16 de fev. de 2020

TWDB (2018) **Groundwater Database Reports.** <http://www.twdb.texas.gov/groundwater/index.asp>. Acesso em: 12 de agosto de 2018.

WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data Mining: practical machine learning tools and techniques.** 3ª edição. United States: Elsevier, 2011.

WUNSCH, A.; LIESCH, T.; BRODA, S. **Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX).** Journal of Hydrology, 2018.

YANG, J. et al. **Using cluster analysis for understanding spatial and temporal patterns and controlling factors of groundwater geochemistry in a regional aquifer.** Journal of Hydrology, v. 583, n. December 2019, p. 124594, 2020.

YOON, H. et al. **A method to improve the stability and accuracy of ANN- and SVM-based time series models for long-term groundwater level predictions.** Computers and Geosciences, v. 90, p. 144–155, 2016.

ZHANG, F. et al. **Application of ANN to predict the apparent viscosity of waxy crude oil.** Fuel, v. 254, n. May, p. 115669, 2019.

APÊNDICE A

Tabela 24 – Variação máxima do nível d'água de cada poço a partir do nível d'água mínimo e máximo

Poço	NA mínimo	NA máximo	Variação máxima do NA	Poço	NA mínimo	NA máximo	Variação máxima do NA
1	1008,9	1012,0	3,2	221	921,2	938,0	16,9
2	938,1	945,4	7,3	222	1226,2	1236,4	10,1
3	908,6	927,3	18,7	223	988,0	994,2	6,2
4	919,3	926,0	6,7	224	970,6	984,2	13,6
5	924,0	930,6	6,6	225	990,4	1001,6	11,2
6	992,4	997,7	5,3	226	913,2	916,4	3,2
7	1047,1	1050,5	3,4	227	962,5	976,0	13,5
8	967,6	973,6	6,1	228	955,8	974,7	18,9
9	1076,6	1103,5	26,9	229	957,4	960,2	2,8
10	936,5	951,0	14,6	230	841,6	844,3	2,7
11	1009,0	1014,8	5,8	231	1139,4	1152,8	13,3
12	924,1	929,4	5,3	232	1055,9	1086,8	30,9
13	1145,7	1151,1	5,3	233	1133,6	1140,9	7,3
14	1060,3	1079,8	19,4	234	1083,0	1092,9	9,9
15	1068,7	1070,0	1,3	235	1108,4	1129,1	20,7
16	970,4	973,0	2,6	236	928,6	954,7	26,1
17	927,1	932,7	5,6	237	853,5	860,7	7,3
18	1107,1	1116,8	9,7	238	1024,6	1055,0	30,4
19	1113,0	1120,3	7,2	239	1101,5	1134,1	32,6
20	1081,1	1091,2	10,2	240	913,5	942,7	29,2
21	1232,2	1239,9	7,7	241	1125,2	1144,3	19,1
22	1136,3	1171,1	34,8	242	1115,6	1120,7	5,1
23	1075,4	1122,1	46,7	243	1012,2	1018,4	6,2
24	913,2	952,1	39,0	244	975,0	983,2	8,3
25	963,6	975,4	11,8	245	1079,5	1093,2	13,7
26	915,4	923,4	8,0	246	845,1	854,8	9,8
27	955,5	961,6	6,2	247	924,6	950,5	25,9
28	934,9	940,8	5,9	248	988,8	993,2	4,4
29	924,4	928,1	3,7	249	1144,5	1165,5	21,0
30	920,1	924,4	4,3	250	1056,9	1068,8	11,9
31	1152,7	1168,0	15,3	251	1097,0	1110,1	13,2
32	964,8	967,8	3,0	252	1015,6	1055,4	39,8
33	993,4	997,3	3,9	253	1134,0	1184,1	50,2
34	984,3	986,9	2,5	254	914,4	923,1	8,8
35	1106,5	1110,9	4,4	255	906,4	927,9	21,5
36	1095,2	1098,8	3,5	256	935,8	955,4	19,7
37	1040,6	1045,0	4,4	257	851,9	865,5	13,6
38	960,5	966,0	5,5	258	1069,9	1077,9	8,0
39	903,5	925,0	21,5	259	968,2	980,8	12,7
40	943,9	946,9	3,0	260	970,7	997,1	26,4

Poço	NA mínimo	NA máximo	Variação máxima do NA	Poço	NA mínimo	NA máximo	Variação máxima do NA
41	971,3	974,7	3,4	261	1015,7	1020,9	5,3
42	907,2	926,1	19,0	262	1109,9	1117,4	7,4
43	927,9	940,6	12,7	263	1060,3	1095,6	35,3
44	1184,5	1191,0	6,5	264	1099,9	1106,3	6,4
45	1039,2	1051,9	12,6	265	974,2	979,1	4,9
46	897,7	909,1	11,4	266	959,5	967,6	8,1
47	991,9	996,1	4,2	267	1012,8	1024,2	11,4
48	1028,1	1034,9	6,8	268	944,4	953,3	8,9
49	1011,1	1018,8	7,7	269	1061,8	1088,4	26,6
50	1172,7	1184,5	11,8	270	837,7	850,2	12,5
51	908,7	929,0	20,3	271	846,3	859,0	12,7
52	942,4	948,5	6,1	272	833,9	846,5	12,6
53	995,6	999,2	3,6	273	1158,6	1163,2	4,6
54	921,0	959,2	38,2	274	1118,8	1132,5	13,6
55	916,9	935,3	18,5	275	1129,1	1138,5	9,4
56	1110,1	1115,6	5,4	276	853,3	862,8	9,5
57	886,3	913,3	27,0	277	822,8	831,4	8,6
58	945,6	978,8	33,2	278	1010,9	1023,1	12,2
59	1174,0	1182,4	8,4	279	893,0	897,7	4,7
60	935,0	943,8	8,8	280	1012,9	1048,4	35,6
61	960,7	968,2	7,5	281	1104,6	1130,5	25,9
62	957,0	967,4	10,5	282	1072,7	1084,0	11,3
63	933,4	954,3	20,9	283	1039,6	1046,3	6,7
64	911,7	919,5	7,8	284	897,9	919,1	21,2
65	904,6	931,9	27,3	285	1152,1	1173,0	20,9
66	969,8	972,7	2,9	286	1099,1	1104,4	5,3
67	993,6	996,4	2,7	287	1003,0	1034,6	31,6
68	905,7	910,0	4,2	288	930,1	952,8	22,7
69	1126,4	1138,6	12,3	289	1144,2	1168,7	24,4
70	1073,2	1079,2	6,0	290	980,8	989,1	8,3
71	1162,0	1183,3	21,3	291	1024,0	1027,6	3,6
72	976,1	983,4	7,3	292	968,2	975,1	6,8
73	928,4	934,8	6,4	293	1080,1	1090,2	10,1
74	945,5	949,4	3,9	294	978,9	989,6	10,7
75	972,1	978,2	6,1	295	998,7	1002,7	4,0
76	1133,9	1168,3	34,4	296	978,4	987,3	8,9
77	1182,5	1186,0	3,5	297	814,2	820,3	6,1
78	942,5	947,0	4,4	298	879,7	884,7	5,0
79	1072,6	1076,4	3,8	299	1004,4	1035,7	31,3
80	1139,4	1179,6	40,2	300	974,3	982,2	7,9
81	958,2	961,6	3,4	301	1086,6	1109,5	22,9
82	1170,1	1204,0	33,9	302	1022,3	1033,9	11,7
83	835,9	842,1	6,2	303	1132,3	1142,2	9,9
84	1173,5	1177,7	4,2	304	930,6	963,5	33,0
85	997,8	1003,7	5,9	305	987,0	1006,3	19,3

Poço	NA mínimo	NA máximo	Variação máxima do NA	Poço	NA mínimo	NA máximo	Variação máxima do NA
86	902,9	906,9	4,0	306	834,4	837,0	2,6
87	1081,3	1087,4	6,1	307	995,5	1007,9	12,4
88	1106,1	1114,1	8,0	308	1113,5	1134,9	21,4
89	932,8	952,9	20,1	309	1016,1	1029,1	13,0
90	913,8	919,6	5,8	310	1165,2	1170,2	4,9
91	999,8	1005,4	5,6	311	930,2	950,3	20,1
92	959,7	967,4	7,7	312	813,8	818,2	4,5
93	829,1	848,9	19,8	313	1021,6	1029,8	8,2
94	1076,6	1082,6	6,1	314	819,9	850,7	30,9
95	921,3	927,1	5,7	315	921,0	948,1	27,1
96	1072,2	1074,6	2,4	316	1169,1	1199,7	30,6
97	1127,0	1134,9	7,9	317	1065,6	1067,3	1,6
98	1071,3	1095,2	23,9	318	1003,0	1014,0	11,0
99	1119,3	1130,1	10,7	319	1146,5	1169,1	22,6
100	1021,6	1035,5	13,9	320	1066,0	1076,1	10,1
101	882,3	911,3	29,0	321	810,0	824,3	14,3
102	930,7	937,3	6,6	322	1097,5	1130,9	33,3
103	914,5	921,1	6,6	323	804,6	830,6	26,1
104	899,7	908,4	8,7	324	937,4	947,9	10,6
105	1065,5	1079,4	13,9	325	1074,8	1078,5	3,6
106	1061,2	1064,8	3,6	326	1128,5	1132,1	3,5
107	1057,6	1061,4	3,7	327	926,0	946,1	20,1
108	1111,4	1155,0	43,6	328	1028,7	1043,2	14,6
109	1103,5	1121,2	17,6	329	971,4	982,6	11,2
110	970,1	978,3	8,2	330	921,0	924,9	3,9
111	1028,0	1036,2	8,2	331	1005,9	1018,7	12,8
112	1009,8	1013,2	3,4	332	1198,7	1203,8	5,1
113	1048,2	1062,2	14,0	333	1044,9	1048,0	3,1
114	1049,5	1052,4	2,9	334	1132,6	1139,1	6,5
115	1023,3	1028,8	5,5	335	963,9	980,2	16,3
116	1012,6	1016,4	3,8	336	925,0	947,8	22,8
117	1039,9	1044,6	4,7	337	1098,7	1117,7	19,0
118	1032,0	1036,9	4,9	338	1109,5	1118,6	9,1
119	1043,0	1046,8	3,7	339	1116,8	1122,3	5,5
120	1007,0	1014,3	7,2	340	1064,4	1074,9	10,5
121	1038,3	1041,0	2,7	341	837,8	850,2	12,5
122	1030,5	1036,8	6,4	342	1002,2	1027,3	25,1
123	1021,2	1026,8	5,6	343	1135,1	1162,3	27,2
124	1234,7	1247,2	12,5	344	1002,6	1056,0	53,4
125	1091,8	1107,9	16,1	345	1161,6	1174,4	12,8
126	997,4	1007,0	9,7	346	1133,9	1157,3	23,5
127	998,1	1008,0	9,9	347	1086,3	1109,0	22,7
128	993,0	999,5	6,5	348	1056,8	1067,6	10,8
129	915,7	939,2	23,5	349	969,8	1005,3	35,6
130	839,9	848,6	8,7	350	807,9	841,4	33,5

Poço	NA mínimo	NA máximo	Variação máxima do NA	Poço	NA mínimo	NA máximo	Variação máxima do NA
131	1000,0	1013,2	13,2	351	808,8	813,1	4,3
132	922,8	926,0	3,1	352	989,4	999,8	10,3
133	881,3	892,0	10,7	353	790,9	793,9	3,0
134	859,2	868,8	9,6	354	799,4	806,2	6,9
135	1151,6	1153,1	1,5	355	1080,1	1084,0	3,9
136	919,9	935,9	16,0	356	1016,7	1068,4	51,7
137	890,5	898,0	7,5	357	1017,3	1047,8	30,5
138	1183,9	1191,5	7,5	358	998,3	1022,0	23,8
139	1126,6	1174,2	47,6	359	952,1	973,2	21,1
140	888,7	894,7	6,0	360	887,3	892,7	5,5
141	837,3	854,8	17,4	361	1158,8	1168,6	9,8
142	933,4	950,4	17,0	362	1046,5	1085,2	38,7
143	1107,3	1127,5	20,2	363	869,4	885,4	16,0
144	1109,9	1117,9	8,0	364	1003,4	1026,9	23,5
145	1104,8	1121,0	16,2	365	1003,7	1036,9	33,2
146	1058,3	1069,8	11,5	366	1153,0	1167,0	14,0
147	1149,1	1171,5	22,3	367	1047,0	1091,2	44,2
148	902,5	924,4	21,9	368	781,3	800,3	19,1
149	1054,8	1092,2	37,3	369	906,2	942,6	36,4
150	852,6	860,5	7,9	370	1013,4	1022,9	9,5
151	1167,6	1175,8	8,2	371	1007,7	1020,1	12,4
152	1090,4	1105,1	14,8	372	1164,5	1168,6	4,1
153	1093,1	1117,7	24,6	373	750,0	756,9	6,9
154	952,0	954,4	2,4	374	830,5	834,0	3,4
155	1103,8	1118,3	14,6	375	780,0	787,1	7,1
156	1087,2	1093,8	6,6	376	938,0	972,1	34,1
157	1088,7	1102,1	13,4	377	1062,4	1081,4	19,0
158	1094,4	1110,0	15,5	378	1045,3	1066,8	21,5
159	1183,5	1186,7	3,2	379	779,8	797,7	17,9
160	1143,3	1148,0	4,7	380	1091,5	1112,3	20,8
161	1067,4	1089,8	22,4	381	993,9	1040,4	46,5
162	982,3	986,6	4,3	382	814,4	834,1	19,7
163	1153,6	1160,1	6,5	383	904,8	932,8	28,0
164	1099,6	1104,5	4,9	384	1003,9	1008,5	4,6
165	1067,3	1078,2	10,8	385	1022,8	1040,2	17,4
166	1108,9	1139,0	30,1	386	821,0	839,4	18,4
167	970,1	972,5	2,5	387	1009,2	1026,5	17,3
168	1069,4	1074,9	5,5	388	979,3	994,4	15,1
169	1086,9	1101,6	14,7	389	1015,3	1038,0	22,7
170	1082,8	1102,9	20,1	390	1087,0	1096,5	9,4
171	1005,2	1015,8	10,7	391	871,6	888,4	16,8
172	809,5	815,9	6,5	392	820,3	829,3	9,0
173	874,4	892,4	17,9	393	821,5	831,9	10,4
174	886,5	892,6	6,1	394	858,7	897,8	39,1
175	871,8	888,4	16,6	395	923,1	962,8	39,8

Poço	NA mínimo	NA máximo	Variação máxima do NA	Poço	NA mínimo	NA máximo	Variação máxima do NA
176	1054,3	1061,8	7,5	396	805,9	840,6	34,8
177	941,8	962,9	21,0	397	887,8	922,1	34,3
178	992,7	1001,8	9,1	398	834,4	853,7	19,4
179	854,2	868,5	14,3	399	814,8	830,9	16,2
180	875,6	883,3	7,7	400	805,7	839,1	33,5
181	870,0	883,0	13,0	401	862,8	883,6	20,8
182	973,5	979,5	6,0	402	851,3	863,2	11,8
183	1004,3	1019,4	15,1	403	1321,4	1339,6	18,2
184	927,7	938,7	11,0				
185	949,8	956,2	6,4				
186	992,8	1033,5	40,6				
187	992,1	1000,6	8,5				
188	1054,5	1058,7	4,1				
189	1067,8	1115,6	47,8				
190	887,2	898,5	11,4				
191	925,4	942,7	17,3				
192	931,4	940,2	8,8				
193	1034,9	1039,7	4,8				
194	1005,4	1009,1	3,7				
195	1008,7	1048,6	39,9				
196	1118,2	1126,2	8,0				
197	1003,6	1007,4	3,8				
198	1029,2	1036,7	7,4				
199	1002,3	1008,8	6,5				
200	890,3	903,1	12,8				
201	907,7	920,4	12,7				
202	1073,6	1083,0	9,4				
203	852,8	859,4	6,6				
204	918,9	927,6	8,7				
205	916,6	924,8	8,3				
206	1148,6	1158,5	9,9				
207	1042,9	1079,0	36,1				
208	916,7	920,9	4,1				
209	1115,8	1138,6	22,8				
210	1149,6	1171,1	21,5				
211	918,6	958,3	39,8				
212	1136,3	1138,5	2,3				
213	1012,3	1014,0	1,7				
214	1171,2	1175,0	3,8				
215	1157,2	1169,5	12,2				
216	928,6	937,7	9,1				
217	1002,8	1011,6	8,8				
218	863,5	867,2	3,7				
219	836,0	853,4	17,5				
220	978,8	991,1	12,3				

APÊNDICE B

Tabela 25 – Poços que obtiveram resultados satisfatórios por meio do uso da RNA feita especificamente para outro poço

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
1	1, 120, 183, 186, 287, 299, 331, 342, 358, 381, 387
2	2, 10, 63, 142, 256, 288, 304, 311, 324
3	3, 4, 24, 26, 39, 42, 43, 51, 54, 55, 65, 129, 136, 148, 191, 201, 211, 221, 236, 240, 247, 254, 255, 315, 327, 369, 383, 397
4	4, 30
5	5, 12, 24, 54, 63, 89, 191, 211, 236, 247, 288, 304, 315, 327
6	6
7	7, 113
8	8, 329, 335
9	9, 20, 70, 87, 94, 98, 125, 152, 156, 157, 158, 169, 170, 189, 202, 234, 245, 286, 293, 301, 347, 380
10	2, 10, 28, 43, 54, 60, 63, 89, 102, 142, 177, 192, 211, 216, 236, 247, 256, 288, 304, 311, 324, 327, 336, 376, 395
11	1, 11, 120, 318, 331
12	5, 12, 55, 129, 240
13	13, 231
14	14, 105, 146, 149, 161, 202, 232, 250, 263, 269, 320, 340, 348, 362, 367
15	15
16	16, 110, 335
17	17
18	18, 35, 56, 109, 145, 155
19	18, 19, 108, 109, 145, 155, 166, 239, 262, 281, 322
20	9, 20, 23, 125, 152, 158, 169, 170, 234, 245, 301, 347, 380
21	21
22	22, 31, 69, 76, 80, 108, 139, 143, 145, 147, 166, 210, 215, 235, 239, 249, 253, 281, 285, 289, 308, 319, 343, 346, 361, 366
23	9, 14, 23, 98, 125, 149, 153, 155, 158, 161, 170, 189, 195, 207, 232, 235, 239, 251, 252, 263, 269, 281, 301, 322, 337, 344, 362, 367
24	3, 4, 5, 12, 24, 26, 30, 42, 51, 54, 55, 58, 63, 64, 89, 90, 95, 103, 129, 132, 136, 142, 177, 191, 204, 211, 221, 227, 228, 236, 240, 247, 254, 255, 256, 288, 304, 315, 327, 336, 349, 359, 369, 376
25	25, 41, 58, 110, 177, 227, 228, 260, 349, 359, 376
26	26, 30, 55, 103, 240, 254
27	27
28	28
29	29
30	30
31	31, 147, 210, 285, 289, 316, 319
32	8, 32, 66, 227, 228
33	33, 128
34	34
35	35, 88, 109, 153, 251, 322, 337
36	36, 125, 152, 380
37	37
38	38, 58, 335, 376
39	39, 57, 65, 104, 148, 200
40	40, 52

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
41	41, 75
42	3, 4, 24, 26, 39, 42, 43, 51, 54, 55, 57, 63, 65, 89, 95, 101, 129, 136, 142, 148, 175, 191, 204, 211, 221, 236, 240, 247, 254, 255, 288, 315, 327, 336, 369, 383, 391, 397
43	5, 12, 43, 54, 191, 211, 221
44	44, 138
45	7, 37, 45
46	46
47	47, 128
48	48
49	49, 287, 371
50	50, 59, 71, 80, 82, 253, 316
51	3, 24, 42, 43, 51, 54, 55, 65, 136, 142, 148, 191, 211, 236, 240, 247, 315, 383, 397
52	52
53	53
54	2, 3, 4, 8, 10, 24, 25, 27, 32, 41, 42, 43, 51, 54, 55, 58, 62, 63, 66, 72, 75, 89, 110, 129, 136, 142, 177, 191, 211, 221, 227, 228, 236, 240, 247, 255, 256, 259, 288, 300, 304, 315, 324, 327, 329, 335, 336, 359, 369, 376, 395
55	24, 26, 54, 55, 89, 103, 129, 211, 236, 240, 254, 315, 369
56	18, 56, 143, 262, 281, 338
57	3, 39, 42, 51, 57, 65, 101, 133, 148, 190, 200, 255, 383, 397
58	8, 25, 27, 41, 58, 62, 63, 89, 92, 142, 223, 227, 228, 256, 259, 260, 288, 292, 304, 335, 349, 359, 376
59	59, 71, 151
60	60
61	61
62	62
63	2, 10, 28, 58, 63, 89, 102, 142, 177, 216, 236, 256, 304, 376
64	64
65	3, 24, 39, 42, 43, 46, 51, 55, 57, 65, 104, 129, 136, 148, 191, 201, 221, 240, 247, 255, 315, 327, 336, 369, 383, 397
66	66
67	6, 47, 67, 126, 128, 178, 223, 225, 248, 352
68	68
69	69, 108, 125, 153, 166, 241, 322
70	14, 23, 70, 149, 189, 232, 250, 258
71	71, 151
72	72, 260, 329, 349
73	73
74	74
75	8, 25, 41, 66, 75, 110, 335
76	22, 31, 69, 71, 76, 80, 97, 139, 147, 206, 210, 241, 249, 253, 285, 289, 303, 319, 343, 346
77	77, 82, 316
78	2, 10, 63, 78, 256, 304, 324
79	79, 96, 258
80	22, 31, 71, 76, 80, 82, 139, 147, 210, 215, 249, 253, 285, 289, 316, 319, 343, 345, 346, 361, 366
81	81
82	59, 71, 82, 138, 316
83	83, 141, 219, 246, 272, 314, 398
84	84

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
85	85, 126, 127, 186, 199, 358
86	86
87	9, 23, 87, 153, 189, 301
88	18, 88, 108, 109, 145, 153, 155, 158, 166, 239, 251, 264, 281, 286, 301, 322, 337
89	2, 10, 58, 63, 89, 102, 142, 177, 192, 228, 236, 247, 256, 288, 304, 324, 336, 376, 395
90	90, 148, 397
91	85, 91, 126, 127, 186
92	25, 92, 110, 227, 228, 335
93	93, 179, 237, 276, 314, 398
94	9, 94, 98, 149, 161, 189, 367
95	4, 30, 55, 95
96	96
97	97, 241
98	9, 14, 94, 98, 105, 149, 156, 161, 170, 189, 202, 232, 234, 269, 320, 362, 367
99	69, 97, 99, 196, 209, 274, 308
100	100, 186, 195, 252, 280, 287, 299, 342, 344, 371, 381, 387, 389
101	57, 101, 133, 200
102	102
103	103
104	104
105	105, 161, 269
106	106
107	107
108	18, 22, 76, 80, 88, 108, 109, 139, 143, 145, 155, 166, 210, 235, 239, 249, 253, 281, 289, 308, 319, 361
109	18, 19, 88, 99, 108, 109, 143, 144, 145, 155, 166, 209, 235, 239, 242, 251, 262, 264, 274, 281, 286, 308, 322, 337
110	58, 72, 110, 228, 260, 329, 376
111	111
112	112
113	113, 114, 188
114	114
115	100, 115, 280, 309, 313, 389
116	116, 243
117	37, 117, 193, 238, 283, 328, 357
118	111, 118, 122, 198
119	119
120	1, 120, 171, 331, 384
121	121
122	48, 111, 118, 122, 198
123	100, 115, 116, 123, 195, 243, 261, 267, 280, 287, 299, 302, 309, 364, 365, 370, 371, 389
124	124
125	125, 152, 153, 156, 158, 169, 301, 347, 380
126	85, 91, 126, 127, 199
127	85, 91, 126, 127, 307
128	47, 67, 91, 128
129	3, 26, 42, 51, 55, 65, 103, 129, 148, 240, 254, 383, 397
130	130
131	1, 131, 358, 371
132	132

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
133	133
134	134
135	135
136	26, 55, 102, 129, 136, 221, 240, 369
137	137
138	138
139	22, 31, 69, 76, 80, 108, 139, 147, 206, 209, 210, 215, 241, 249, 274, 285, 289, 308, 319, 343, 345, 346, 361, 366
140	140
141	83, 141, 219, 246, 272, 398
142	2, 10, 24, 42, 43, 54, 58, 63, 65, 89, 142, 177, 211, 236, 240, 247, 256, 288, 304, 315, 324, 327, 376, 395
143	18, 19, 35, 109, 143, 145, 153, 155, 166, 235, 239, 251, 262, 281, 286, 308, 322, 337, 338
144	18, 19, 109, 143, 144, 145, 155, 235, 262, 281
145	9, 18, 23, 35, 88, 108, 109, 143, 145, 153, 155, 166, 189, 209, 235, 239, 251, 262, 281, 301, 308, 322, 337
146	146, 250
147	22, 31, 147, 206, 249
148	3, 24, 42, 51, 54, 55, 65, 129, 148, 191, 211, 240, 255, 315, 369, 383, 397
149	9, 14, 20, 36, 98, 105, 113, 125, 149, 152, 156, 157, 161, 169, 170, 189, 202, 207, 232, 234, 245, 263, 269, 293, 301, 347, 356, 362, 367, 378
150	150, 179, 237, 257
151	59, 71, 82, 151, 316
152	9, 23, 125, 152, 156, 157, 158, 169, 170, 251, 301, 347, 380
153	99, 108, 125, 152, 153, 155, 158, 166, 209, 235, 239, 274, 308, 322, 337
154	154
155	18, 19, 56, 88, 109, 143, 145, 155, 235, 239, 242, 262, 281
156	156
157	157
158	9, 23, 125, 152, 153, 158, 170, 251, 301
159	159
160	160
161	9, 14, 20, 23, 94, 98, 105, 125, 149, 152, 156, 158, 161, 169, 170, 189, 202, 245, 263, 269, 301, 320, 340, 347, 367, 380
162	162
163	22, 31, 80, 147, 163, 206, 210, 215, 249, 253, 285, 289, 319, 361
164	153, 164
165	165
166	18, 35, 69, 88, 99, 108, 109, 143, 145, 155, 166, 209, 235, 239, 241, 262, 274, 281, 308
167	167
168	14, 96, 98, 105, 149, 161, 168, 207, 232, 263, 269, 362, 367
169	125, 158, 169, 301
170	9, 14, 20, 23, 98, 105, 125, 149, 161, 169, 170, 189, 202, 232, 234, 263, 269, 301, 340, 355, 367, 380
171	120, 171, 183, 194, 197, 318, 331, 371
172	172, 350, 396, 400
173	173, 200
174	101, 133, 174, 200
175	101, 175, 391
176	176
177	2, 10, 28, 40, 52, 58, 60, 63, 78, 89, 142, 177, 256, 288, 304, 311, 324, 376

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
178	6, 128, 178
179	179
180	180
181	181, 391
182	182, 300
183	91, 116, 120, 127, 171, 183, 342, 358, 371
184	184, 221
185	185
186	183, 186, 195, 207, 238, 252, 280, 287, 299, 342, 344, 358, 381, 387
187	6, 67, 128, 178, 187
188	113, 146, 188, 207, 250, 362, 378
189	9, 14, 20, 23, 98, 105, 125, 149, 152, 153, 155, 161, 169, 170, 189, 202, 207, 232, 238, 263, 269, 301, 347, 357, 362, 367, 378
190	190
191	5, 12, 54, 136, 191, 211, 236, 247, 315, 327
192	24, 43, 54, 102, 136, 191, 192, 211, 216, 221, 236, 240, 247, 315, 327, 336
193	111, 193
194	194
195	45, 120, 123, 183, 195, 238, 252, 280, 287, 299, 309, 331, 342, 344, 364, 371, 387, 389
196	196
197	197
198	48, 111, 118, 122, 198
199	91, 127, 197, 199, 358
200	57, 65, 101, 140, 200
201	3, 42, 65, 201, 255
202	70, 94, 161, 202
203	203, 237
204	4, 26, 55, 204, 240
205	4, 30, 205
206	135, 206, 249
207	9, 14, 23, 45, 98, 100, 107, 109, 113, 116, 145, 149, 155, 186, 189, 195, 198, 207, 232, 235, 238, 243, 250, 252, 269, 278, 280, 287, 299, 309, 313, 322, 328, 342, 344, 348, 362, 364, 367, 370, 371, 378, 381, 385, 387, 389
208	208
209	22, 69, 76, 80, 97, 99, 108, 139, 209, 241, 249, 253, 274, 275, 289, 308, 319, 343, 346
210	22, 31, 80, 82, 135, 147, 206, 210, 249, 253, 285, 289, 316, 319
211	4, 5, 12, 24, 27, 30, 54, 55, 58, 62, 63, 89, 95, 129, 136, 191, 204, 211, 221, 236, 240, 247, 288, 304, 315, 327, 359, 369, 376, 395
212	212
213	49, 213, 243
214	214, 316
215	210, 215, 345, 361
216	10, 43, 63, 89, 102, 142, 216, 236, 247, 256, 288, 324, 336
217	120, 183, 186, 195, 217, 287, 299, 342, 358, 371, 381, 387
218	218
219	130, 141, 219, 246, 270, 398
220	220, 224, 329
221	5, 102, 191, 221
222	222
223	223
224	41, 66, 72, 75, 110, 224, 244, 329

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
225	6, 47, 225
226	26, 226
227	8, 25, 41, 61, 66, 75, 110, 227, 228, 259, 292, 335
228	8, 25, 27, 75, 227, 228, 259, 300, 329, 335
229	229
230	230
231	22, 76, 80, 231, 249, 289, 319, 343, 346
232	9, 14, 23, 98, 105, 146, 149, 153, 161, 170, 188, 189, 232, 250, 251, 263, 269, 301, 320, 348, 362, 367
233	233
234	9, 20, 23, 87, 98, 149, 170, 189, 234, 245, 293, 301, 347
235	18, 19, 35, 56, 69, 88, 99, 108, 109, 143, 145, 153, 155, 158, 166, 209, 235, 239, 251, 262, 274, 281, 286, 308, 322, 337, 338
236	5, 12, 43, 54, 58, 63, 89, 102, 136, 184, 191, 192, 211, 216, 221, 236, 247, 256, 288, 304, 315, 327, 359, 376
237	237
238	37, 45, 100, 115, 117, 120, 123, 171, 183, 186, 195, 238, 243, 252, 280, 287, 299, 302, 309, 313, 328, 342, 344, 358, 364, 370, 371, 381, 389
239	9, 18, 23, 35, 88, 98, 109, 143, 145, 153, 155, 166, 189, 235, 239, 251, 262, 274, 281, 301, 308, 322, 337, 338
240	3, 24, 39, 42, 43, 51, 54, 57, 65, 101, 102, 129, 133, 136, 142, 148, 184, 191, 192, 211, 221, 240, 247, 315, 327, 369, 383, 397
241	99, 108, 166, 209, 241, 253, 274
242	19, 99, 108, 166, 235, 242, 262, 274, 308
243	243, 370
244	244, 329
245	245
246	219, 246
247	24, 54, 55, 211, 247, 315
248	248
249	22, 31, 69, 76, 80, 97, 139, 147, 206, 210, 241, 249, 253, 275, 285, 289, 303, 319, 343, 346, 366
250	9, 14, 23, 98, 107, 113, 146, 149, 161, 188, 189, 207, 232, 250, 263, 269, 362, 367, 378
251	158, 251
252	45, 100, 113, 195, 198, 238, 252, 280, 287, 299, 328, 344, 364, 371, 378, 389
253	22, 31, 50, 59, 69, 71, 76, 80, 82, 139, 147, 210, 241, 253, 285, 289, 316, 319, 343, 345, 346
254	24, 26, 30, 51, 54, 55, 90, 103, 129, 211, 240, 254, 315, 369, 383
255	3, 24, 26, 42, 43, 51, 54, 55, 63, 65, 89, 90, 103, 129, 136, 142, 148, 191, 211, 221, 236, 240, 247, 254, 255, 315, 327, 336, 369, 383, 397
256	2, 10, 28, 58, 60, 63, 89, 142, 177, 256, 304, 376
257	150, 257, 398
258	70, 79, 94, 96, 161, 165, 202, 258, 269
259	75, 259, 329, 335
260	8, 25, 66, 110, 120, 197, 217, 259, 260, 329, 335, 349, 358
261	261
262	18, 19, 88, 99, 109, 143, 145, 155, 166, 196, 235, 239, 242, 262, 281, 308, 322, 338
263	14, 18, 23, 107, 109, 113, 146, 155, 188, 189, 232, 250, 262, 263, 348, 367
264	264, 286
265	Nenhum
266	58, 260, 266, 376

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
267	243, 267, 370
268	268
269	14, 20, 79, 98, 105, 149, 161, 168, 207, 232, 245, 258, 263, 269, 320, 340, 362, 367
270	270
271	271
272	141, 272, 398
273	273
274	97, 99, 209, 274
275	69, 275, 303
276	276
277	277
278	116, 243, 278
279	279
280	11, 14, 37, 45, 70, 100, 105, 111, 113, 115, 118, 122, 123, 149, 161, 165, 188, 193, 195, 198, 202, 207, 232, 238, 243, 252, 263, 269, 280, 287, 299, 302, 313, 328, 344, 356, 357, 362, 367, 371, 378, 381, 389
281	18, 22, 23, 76, 108, 109, 139, 143, 145, 155, 166, 189, 209, 235, 239, 249, 274, 281, 308, 343, 346
282	9, 20, 98, 149, 156, 170, 189, 234, 245, 282, 367
283	37, 283
284	284
285	31, 285, 366
286	125, 153, 158, 251, 286, 337
287	11, 48, 100, 111, 115, 123, 183, 186, 195, 198, 243, 267, 280, 287, 299, 302, 313, 342, 358, 364, 370, 371, 381, 387, 389
288	43, 54, 63, 89, 102, 142, 192, 211, 216, 236, 247, 256, 288, 327
289	31, 206, 210, 249, 285, 289, 319
290	290, 329
291	291
292	110, 259, 292, 335
293	20, 170, 234, 245, 293
294	Nenhum
295	295
296	260, 296
297	297
298	298
299	85, 123, 127, 171, 183, 195, 197, 199, 238, 252, 280, 287, 299, 342, 344, 358, 389
300	72, 260, 300, 329, 349
301	125, 153, 155, 158, 169, 251, 301, 337
302	100, 195, 238, 252, 280, 287, 302, 344, 389
303	69, 303
304	2, 10, 27, 43, 58, 60, 62, 63, 81, 89, 92, 102, 142, 177, 227, 228, 236, 247, 256, 288, 304, 311, 324, 327, 335, 336, 359, 376, 395
305	Nenhum
306	306
307	85, 91, 126, 127, 194, 197, 199, 307
308	18, 19, 99, 108, 145, 155, 166, 209, 235, 262, 274, 308
309	100, 123, 238, 252, 280, 302, 309, 389
310	310
311	142, 311, 324
312	312, 323, 350, 382, 396, 400

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
313	313
314	93, 134, 141, 179, 203, 219, 237, 257, 272, 314, 398
315	3, 4, 5, 10, 12, 24, 26, 39, 42, 43, 51, 54, 55, 58, 63, 65, 89, 95, 129, 136, 142, 148, 177, 191, 204, 211, 221, 227, 228, 236, 240, 247, 255, 256, 288, 304, 315, 324, 327, 335, 359, 369, 376, 383, 395, 397
316	44, 50, 59, 71, 82, 138, 151, 316
317	317
318	318, 342
319	22, 31, 76, 80, 147, 206, 210, 249, 253, 285, 289, 319
320	14, 105, 161, 258, 263, 269, 320, 340, 367
321	321, 382
322	9, 23, 35, 109, 125, 143, 145, 152, 153, 155, 157, 158, 166, 169, 170, 189, 235, 239, 251, 262, 281, 286, 301, 322, 337, 347
323	93, 314, 323, 350, 382, 386, 396, 400
324	52, 78, 324
325	161, 165, 202, 269, 325
326	326
327	2, 10, 12, 24, 43, 54, 58, 63, 65, 89, 101, 142, 191, 200, 211, 236, 240, 247, 256, 288, 304, 311, 315, 327, 336, 376, 397
328	45, 195, 207, 238, 252, 328, 344
329	8, 16, 32, 66, 110, 227, 260, 329, 349
330	330
331	120, 171, 194, 331
332	332
333	333
334	334
335	75, 110, 227, 259, 260, 335, 349
336	10, 24, 54, 63, 89, 136, 142, 177, 211, 236, 240, 247, 255, 256, 288, 304, 315, 324, 327, 336, 369, 376
337	99, 109, 125, 143, 145, 153, 155, 158, 166, 235, 239, 251, 264, 281, 286, 308, 322, 337
338	18, 56, 155, 262, 338
339	339
340	9, 14, 98, 149, 189, 263, 269, 340, 367
341	246, 341
342	58, 171, 186, 195, 260, 287, 299, 329, 342, 349, 358, 381
343	22, 76, 80, 139, 212, 249, 289, 303, 319, 343, 346
344	37, 45, 117, 195, 238, 252, 280, 283, 299, 344, 356, 378
345	71, 345, 361
346	22, 76, 80, 210, 249, 289, 319, 343, 346
347	125, 152, 156, 158, 169, 189, 301, 347
348	113, 188, 348
349	41, 72, 75, 100, 110, 123, 186, 195, 207, 238, 252, 260, 261, 267, 280, 287, 299, 300, 302, 309, 313, 329, 342, 344, 349, 358, 364, 370, 371, 378, 381, 387, 389
350	314, 323, 350, 382, 396, 400
351	351
352	352
353	353
354	354
355	355
356	14, 70, 79, 96, 105, 107, 146, 149, 161, 165, 168, 202, 207, 232, 238, 250, 258, 263, 269, 320, 340, 348, 356, 362, 367, 378

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
357	45, 149, 189, 207, 238, 252, 357, 362, 367, 378
358	91, 123, 126, 127, 183, 186, 197, 199, 267, 307, 318, 331, 342, 358, 370, 387
359	58, 359
360	360
361	22, 31, 80, 147, 206, 210, 249, 253, 285, 289, 319, 345, 361, 366
362	9, 14, 20, 45, 70, 94, 98, 105, 149, 161, 202, 207, 232, 234, 245, 263, 269, 355, 362, 367, 378
363	363, 394, 401
364	100, 195, 252, 280, 287, 309, 364, 389
365	45, 238, 328, 365
366	31, 80, 135, 147, 206, 210, 249, 285, 289, 319, 366
367	9, 14, 20, 70, 94, 96, 98, 105, 149, 161, 168, 170, 202, 207, 232, 234, 245, 258, 269, 362, 367
368	368, 379
369	3, 4, 5, 24, 26, 42, 51, 55, 65, 95, 129, 136, 191, 204, 240, 247, 254, 255, 315, 324, 327, 369, 383
370	116, 120, 183, 186, 243, 267, 299, 342, 358, 370, 371, 381, 387
371	116, 120, 183, 186, 195, 243, 280, 287, 299, 309, 342, 344, 358, 370, 371, 381, 389
372	372
373	373
374	374
375	375
376	2, 10, 58, 62, 177, 227, 228, 256, 304, 324, 335, 349, 359, 376, 38
377	149, 161, 189, 207, 232, 362, 367, 377
378	207, 238, 378
379	379
380	9, 23, 36, 125, 152, 153, 156, 157, 158, 169, 170, 189, 251, 301, 322, 337, 347, 380
381	37, 45, 72, 100, 117, 186, 193, 195, 198, 207, 224, 238, 244, 252, 259, 260, 280, 287, 299, 300, 328, 329, 335, 342, 344, 349, 356, 357, 358, 365, 381, 385, 389
382	93, 141, 179, 219, 272, 314, 350, 382, 396, 398
383	3, 4, 24, 26, 42, 51, 54, 55, 63, 65, 89, 90, 103, 129, 136, 148, 204, 208, 211, 221, 236, 240, 247, 254, 255, 256, 304, 315, 327, 336, 369, 383, 395, 397
384	384
385	45, 193, 195, 238, 252, 280, 328, 344, 385
386	386
387	58, 177, 370, 371, 376, 387
388	260, 349, 388
389	123, 195, 267, 287, 299, 302, 309, 313, 381, 389
390	Nenhum
391	175, 391
392	392
393	314, 323, 350, 393, 396, 400
394	65, 148, 383, 394, 397
395	8, 10, 41, 43, 58, 63, 66, 75, 89, 110, 142, 177, 191, 236, 247, 256, 304, 335, 376, 395
396	314, 350, 396
397	3, 24, 39, 42, 51, 55, 65, 136, 148, 221, 240, 255, 369, 383, 397
398	219, 398
399	314, 323, 350, 382, 386, 396, 399, 400
400	323, 350, 382, 396, 399, 400

Poço cuja RNA foi testada	Poços que obtiveram resultados satisfatórios
401	314, 394, 401
402	150, 203, 219, 237, 246, 398, 402
403	403