



Cesar Hernando Valencia Niño

Modelo híbrido **ESN-GA-SRG**: uma abordagem
de otimização e seleção de topologias em *Echo State Networks*
para previsão de séries temporais

Tese de Doutorado

Tese apresentada ao Programa de Pós-Graduação
em Engenharia Elétrica da PUC-Rio como requisito parcial para
obtenção do grau de Doutor em Engenharia Elétrica.

Orientadoras: Profa. Marley Maria Bernardes Rebuzzi Vellasco
Profa. Karla Tereza Figueiredo

Rio de Janeiro
Abril de 2016



Cesar Hernando Valencia Niño

Modelo híbrido **ESN-GA-SRG**: uma abordagem de otimização e seleção de topologias em *Echo State Networks* para previsão de séries temporais

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Profa. Marley Maria Bernardes Rebuzzi Vellasco

Orientadora

Departamento de Engenharia Elétrica – PUC-Rio

Profa. Karla Tereza Figueiredo

Co-orientadora

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Reinaldo Castro Souza

Departamento de Engenharia Elétrica – PUC-Rio

Prof. André Vargas Abs da Cruz

Universidade Estadual da Zona Oeste - UEZO

Prof. Luiz Pereira Calôba

Programa de Engenharia Elétrica - COPPE

Prof. Jorge Luís Machado do Amaral

Universidade Estadual do Rio de Janeiro - UERJ

Prof. Marco Aurélio Botelho da Silva

Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio

Prof. Marcio de Silveira Carvalho

Coordenador Setorial do Centro

Técnico Científico – PUC-Rio

Rio de Janeiro, 01 de abril de 2016

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e da orientadora.

Cesar Hernando Valencia Niño

Graduou-se em Engenharia Eletrônica pela Universidad Antonio Nariño em 2002. Obteve o título de Mestre em Engenharia Mecânica com ênfase em Projeto de Maquinas e Robótica pela UFRJ em 2009.

Ficha Catalográfica

Valencia Niño, Cesar Hernando

Modelo híbrido ESN-GA-SRG : uma abordagem de otimização e seleção de topologias em *Echo State Networks* para previsão de séries temporais / Cesar Hernando Valencia Niño ; orientadoras: Marley Maria Bernardes Rebuzzi Vellasco, Karla Tereza Figueiredo. – 2016.

207 f. : il. color. ; 30 cm

Tese (doutorado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2016.

Inclui bibliografia

1. Engenharia Elétrica – Teses. 2. Echo State Networks. 3. Algoritmos genéticos. 4. Separation Ratio Graph. 5. Previsão Multi-step de séries temporais. I. Vellasco, Marley Maria Bernardes Rebuzzi. II. Figueiredo, Karla Tereza. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

A minha filha Valeria e a minha família

Agradecimentos

Agradeço a todas as pessoas que me apoiaram nesta etapa da minha vida, em particular aos meus pais Hernando e Flor Marina, aos meus Irmãos Laura e Carlos, à minha esposa Silvia Nathalia e, de uma forma muito especial, à minha orientadora a Professora Marley Vellasco e à minha co-orientadora a Professora Karla Figueiredo, assim como ao CNPq e à PUC-Rio pelo apoio econômico e a infraestrutura necessária para o desenvolvimento desta pesquisa.

Resumo

Valencia Niño, Cesar Hernando; Vellasco, Marley Maria Bernardes Rebuzzi. **Modelo híbrido ESN-GA-SRG: uma abordagem de otimização e seleção de topologias em Echo State Networks para previsão de séries temporais.** Rio de Janeiro, 2016. 207p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

A utilização de modelos de inteligência computacional para tarefas de previsão *Multi-Step* de séries temporais tem apresentado resultados que permitem considerar estes modelos como alternativas viáveis para este tipo de problema. Baseados nos requerimentos computacionais e a melhora de desempenho, recentemente novas áreas de pesquisa têm sido apresentadas na comunidade científica. Este é o caso do *Reservoir Computing*, que apresenta novos campos de estudo para redes neurais do tipo recorrentes, as quais, no passado, não foram muito utilizados devido à complexidade de treinamento e ao alto custo computacional. Nesta nova área são apresentados modelos como *Liquid State Machine* e *Echo State Networks*, que proporcionam um novo entendimento no conceito de processamento dinâmico para redes recorrentes e propõem métodos de treinamento com baixo custo computacional. Neste trabalho determinou-se como foco de pesquisa a otimização de parâmetros globais para o projeto das *Echo State Networks*. Embora as *Echo State Networks* sejam objeto de estudo de pesquisadores reconhecidos, ainda apresentam comportamentos desconhecidos, em parte pela sua natureza dinâmica, mas também, pela falta de estudos que aprofundem o entendimento no comportamento dos estados gerados. Utilizando como fundamento o modelo *Separation Ratio Graph* para análise do desempenho, é proposto um novo modelo, denominado *ESN-GA-SRG*, que usa como base redes *ESN* com otimização de parâmetros globais utilizando *GA* e seleção de topologias para *Reservoir* por meio de análise de estados empregando *SRG*. O desempenho deste novo modelo é avaliado na previsão das 11 séries que compõem a versão reduzida do *NN3 Forecasting Competition* e em 36 séries da competição *M3*, selecionadas segundo características de periodicidade na amostragem, assimetria, sazonalidade e estacionaridade. O desempenho do modelo *ESN-GA-SRG* na previsão dessas séries temporais foi superior na maioria dos casos, com significância estatística, quando comparado com outros modelos da literatura.

Palavras-chave

Echo State Networks; Algoritmos Genéticos; *Separation Ratio Graph*; Previsão multi-step de séries temporais.

Abstract

Valencia Niño, Cesar Hernando; Vellasco, Marley Maria Bernardes Rebuzzi (Advisor). **ESN-GA-SRG Hybrid Model: an optimization and topology selection approach in Echo State Networks for time series forecasting.** Rio de Janeiro, 2016. 207p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

The use of computational intelligence models for Multi-Step time series prediction tasks has presented results that allow us to consider these models as viable alternatives for this type of problem. Based on computational requirements and performance improvement, new areas of research have recently been presented in the scientific community. This is the case of Reservoir Computing, which presents new fields of study for recurrent-type neural networks, which in the past were not widely used because of training complexity and high computational cost. In this new area are presented models such as Liquid State Machine and Echo State Networks, which provide a new understanding of the concept of dynamic processing for recurring networks and propose methods of training with low computational cost. In this work, we determined the optimization of global parameters for the Echo State Networks project. Although Echo State Networks are the object of study by recognized researchers, they still present unknown behavior, partly due to their dynamic nature, but also due to the lack of studies that deepen behavior understanding of the generated states. Based on the Separation Ratio Graph model for performance analysis, a new model, called ESN-GA-SRG, is proposed, which uses ESN networks with global parameter optimization using GA and selection of topologies for Reservoir through analysis of States employing SRG. The performance of this new model is evaluated to forecast the 11 series that made up the reduced version of the NN3 Forecasting Competition and for 36 series of the M3 competition, selected according to characteristics of periodicity in sampling, asymmetry, seasonality and stationary. The performance of the ESN-GA-SRG model in predicting these time series was superior in most cases, with statistical significance when compared with other models in the literature.

Keywords

Echo State Networks; Genetic Algorithms; Separation Ratio Graph; multi-step ahead time series-prediction.

Sumário

1. Introdução	17
1.1. Motivação	17
1.2. Objetivos	19
1.3. Contribuições	19
1.4. Organização do Trabalho	20
 2. Reservoir computing	 21
2.1. Modelos de Redes Neurais Recorrentes	21
2.1.1. <i>Temporal Recurrent Neural Network</i>	22
2.1.2. <i>Echo State Networks</i>	23
2.1.3. <i>Liquid State Machines</i>	24
2.1.4. <i>Decorrelation-Backpropagation Learning</i>	26
2.2. Treinamento das <i>Echo State Networks</i>	27
2.2.1. Métodos de avaliação da qualidade do Reservoir	30
2.2.2. Métricas de avaliação dos estados do Reservoir	31
2.2.2.1 Coeficiente de Separacao	31
2.2.2.2 Coeficiente de Lyapunov	33
2.2.2.3 <i>Separation Ratio Graph</i>	34
 3. Análise de sensibilidade	 37
3.1. Apresentação do Benchmark	39
3.2. Parâmetros selecionados	44
3.2.1. Tamanho do <i>Reservoir</i>	45
3.2.2. Percentual de Conexões e Tamanho do <i>Reservoir</i>	49
3.2.2.1 Percentual de conexões 20%	49
3.2.2.2 Percentual de conexões 35%	53
3.2.2.3 Percentual de conexões 50%	57
3.2.2.4 Percentual de conexões 75%	61
3.2.2.5 Percentual de conexões 100%	65
3.2.3. Fator de escala nos pesos das conexões de entrada	70
3.2.4. Fator de escala nos pesos das conexões de recorrentes	74
3.2.5. Raio Espectral	78
3.3. Análise e discussão dos resultados	82

4. Modelo híbrido <i>Echo State Network-Genetic Algorithm-Separation Ratio Graph</i>	83
4.1. Visão geral	84
4.2. Modulo Neuro-Evolução	85
4.2.1. Representação e codificação	86
4.2.2. Inicialização e seleção	86
4.2.3. Reprodução	87
4.2.4. Função Aptidão	88
4.3. <i>Separation Raio Graph</i>	88
4.3.1. Topologias de conexão	91
4.3.2. Agrupamento e vizinhança	92
 5. Resultados	 93
5.1. Estudo de casos	93
5.2. Resultados para a versão reduzida <i>NN3</i>	96
5.2.1. Resultados testes de significancia estatistica para as séries <i>NN3</i>	120
5.3. Resultados para as séries selecionadas da competição <i>M3</i>	123
5.3.1. Resultados para as séries do grupo anual da competicao <i>M3</i>	127
5.3.2. Resultados para as séries do grupo trimestral da competicao <i>M3</i>	141
5.3.3. Resultados para as séries do grupo mensal da competicao <i>M3</i>	154
5.3.4. Resultados testes de significancia estatistica para as séries <i>M3</i>	168
 6. Considerações finais	 175
6.1. Conclusões	175
6.2. Próximas etapas	176
 Referências bibliográficas	 177
 A. Versão reduzida <i>NN3</i>	 182
B. Séries selecionadas Base de Dados <i>M3</i>	188

Lista de figuras

Figura 1. <i>Reservoir Computing</i> . (Gang, 2015)	24
Figura 2. Estrutura LSM (Maass, 2002).	26
Figura 3. Exemplos topologias do médio líquido. (Ju, 2013).	26
Figura 4. Estrutura das Redes <i>BPDC</i> . (Steil, 2004).	27
Figura 5. Região objetivo e Zonas caóticas e de atratores.	36
Figura 6. Regressão linear de 3 redes com dinâmicas diferentes. (Gibbons 2010).	36
Figura 7. Estrutura dos testes realizados.	39
Figura 8. Séries <i>Mackey-Glass</i> utilizadas, a) $\tau = 10$, b) $\tau = 20$, c) $\tau = 70$.	42
Figura 9. Conjunto de validação vs previsão <i>ESN</i> , a) $\tau = 10$, b) $\tau = 20$, c) $\tau = 70$.	45
Figura 10. Séries <i>Mackey-Glass</i> com conexões e pesos variáveis. ($\#U$)	47
Figura 11. Série <i>Mackey-Glass</i> com conexões fixas e pesos variáveis. ($\#U$)	49
Figura 12. Série <i>Mackey-Glass</i> com Conexões e pesos variáveis. ($\%C = 20$)	52
Figura 13. Série <i>Mackey-Glass</i> com conexões fixas e pesos variáveis. ($\%C = 20$)	54
Figura 14. Série <i>Mackey-Glass</i> com conexões e pesos variáveis. ($\%C = 35$)	56
Figura 15. Série <i>Mackey-Glass</i> com conexões fixas e pesos variáveis. ($\%C = 35$)	58
Figura 16. Série <i>Mackey-Glass</i> com conexões e pesos variáveis. ($\%C = 50$)	60
Figura 17. Série <i>Mackey-Glass</i> com conexões fixas e pesos variáveis. ($\%C = 50$)	62
Figura 18. Série <i>Mackey-Glass</i> com conexões e pesos variáveis. ($\%C = 75$)	64
Figura 19. Série <i>Mackey-Glass</i> com conexões fixas e pesos variáveis. ($\%C = 75$)	66
Figura 20. Série <i>Mackey-Glass</i> com conexões e pesos variáveis. ($\%C = 100$)	68
Figura 21. Série <i>Mackey-Glass</i> com conexões fixas e pesos variáveis. ($\%C = 100$)	70
Figura 22. Série <i>Mackey-Glass</i> com conexões e pesos variáveis. (sWe)	72
Figura 23. Série <i>Mackey-Glass</i> com conexões fixas e pesos variáveis. (sWe)	74
Figura 24. Série <i>Mackey-Glass</i> com conexões e pesos variáveis. ($sWrr$)	76
Figura 25. Série <i>Mackey-Glass</i> com conexões fixas e pesos variáveis. ($sWrr$)	78
Figura 26. Série <i>Mackey-Glass</i> com conexões e pesos variáveis. ($\rho(W)$)	80
Figura 27. Série <i>Mackey-Glass</i> com conexões fixas e pesos variáveis. ($\rho(W)$)	82
Figura 28. Diagrama completo do modelo proposto.	85
Figura 29. Fluxograma do Algoritmo Genético.	86
Figura 30. Composição do Cromossomo.	87
Figura 31. Exemplo de cruzamento de um ponto.	88
Figura 32. Região objetivo e Zonas caóticas e de atratores.	90
Figura 33. Exemplos de 4 Regressões obtidas entre os padrões e estados.	92
Figura 34. Resultados Série <i>NN3-101</i> fora da amostra.	98
Figura 35. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-101</i>).	98
Figura 36. Resultados Série <i>NN3-102</i> fora da amostra.	100
Figura 37. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-102</i>).	100
Figura 38. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	101

Figura 39. Resultados Série <i>NN3-103</i> fora da amostra.	102
Figura 40. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-103</i>).	102
Figura 41. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	103
Figura 42. Resultados Série <i>NN3-104</i> fora da amostra.	104
Figura 43. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-104</i>).	104
Figura 44. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	105
Figura 45. Resultados Série <i>NN3-105</i> fora da amostra.	106
Figura 46. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-105</i>).	106
Figura 47. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	107
Figura 48. Resultados Série <i>NN3-106</i> fora da amostra.	108
Figura 49. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-106</i>).	108
Figura 50. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	109
Figura 51. Resultados Série <i>NN3-107</i> fora da amostra.	110
Figura 52. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-107</i>).	110
Figura 53. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	111
Figura 54. Resultados Série <i>NN3-108</i> fora da amostra.	112
Figura 55. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-108</i>).	112
Figura 56. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	113
Figura 57. Resultados Série <i>NN3-109</i> fora da amostra.	114
Figura 58. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-109</i>).	114
Figura 59. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	115
Figura 60. Resultados Série <i>NN3-110</i> fora da amostra.	116
Figura 61. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-110</i>).	116
Figura 62. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	117
Figura 63. Resultados Série <i>NN3-111</i> fora da amostra.	118
Figura 64. Evolução do <i>sMAPE</i> a longo do horizonte de previsão (Série <i>NN3-111</i>).	118
Figura 65. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	119
Figura 66. Diagrama de barras comparativas dos resultados obtidos.	120
Figura 67. Resultados Série <i>M3-0033</i> fora da amostra.	128
Figura 68. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	129
Figura 69. Resultados Série <i>M3-0050</i> fora da amostra.	129
Figura 70. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	130
Figura 71. Resultados Série <i>M3-0070</i> fora da amostra.	130
Figura 72. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	131
Figura 73. Resultados Série <i>M3-0105</i> fora da amostra.	132
Figura 74. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	132
Figura 75. Resultados Série <i>M3-0116</i> fora da amostra.	133
Figura 76. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	133
Figura 77. Resultados Série <i>M3-0119</i> fora da amostra.	134
Figura 78. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	134

Figura 79. Resultados Série <i>M3-0162</i> fora da amostra.	135
Figura 80. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	135
Figura 81. Resultados Série <i>M3-0171</i> fora da amostra.	136
Figura 82. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	136
Figura 83. Resultados Série <i>M3-0200</i> fora da amostra.	137
Figura 84. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	137
Figura 85. Resultados Série <i>M3-0351</i> fora da amostra.	138
Figura 86. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	138
Figura 87. Resultados Série <i>M3-0564</i> fora da amostra.	139
Figura 88. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	139
Figura 89. Resultados Série <i>M3-0599</i> fora da amostra.	140
Figura 90. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	140
Figura 91. Resultados Série <i>M3-0642</i> fora da amostra.	141
Figura 92. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	141
Figura 93. Evolução do sMAPE para cada modelo.	142
Figura 94. Resultados Série <i>M3-0671</i> fora da amostra.	143
Figura 95. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	143
Figura 96. Resultados Série <i>M3-0677</i> fora da amostra.	144
Figura 97. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	144
Figura 98. Resultados Série <i>M3-0692</i> fora da amostra.	145
Figura 99. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	145
Figura 100. Resultados Série <i>M3-0713</i> fora da amostra.	146
Figura 101. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	146
Figura 102. Resultados Série <i>M3-0730</i> fora da amostra.	147
Figura 103. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	147
Figura 104. Resultados Série <i>M3-0758</i> fora da amostra.	148
Figura 105. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	148
Figura 106. Resultados Série <i>M3-0773</i> fora da amostra.	149
Figura 107. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	149
Figura 108. Resultados Série <i>M3-0829</i> fora da amostra.	150
Figura 109. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	150
Figura 110. Resultados Série <i>M3-0847</i> fora da amostra.	151
Figura 111. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	151
Figura 112. Resultados Série <i>M3-0854</i> fora da amostra.	152
Figura 113. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	152
Figura 114. Resultados Série <i>M3-0865</i> fora da amostra.	153
Figura 115. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	153
Figura 116. Resultados Série <i>M3-0896</i> fora da amostra.	154
Figura 117. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	154
Figura 118. Evolução do sMAPE para cada modelo.	155

Figura 119. Resultados Série <i>M3-1430</i> fora da amostra.	156
Figura 120. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	156
Figura 121. Resultados Série <i>M3-1460</i> fora da amostra.	157
Figura 122. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	157
Figura 123. Resultados Série <i>M3-1725</i> fora da amostra.	158
Figura 124. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	158
Figura 125. Resultados Série <i>M3-1900</i> fora da amostra.	159
Figura 126. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	159
Figura 127. Resultados Série <i>M3-1906</i> fora da amostra.	160
Figura 128. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	160
Figura 129. Resultados Série <i>M3-1912</i> fora da amostra.	161
Figura 130. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	161
Figura 131. Resultados Série <i>M3-1996</i> fora da amostra.	162
Figura 132. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	162
Figura 133. Resultados Série <i>M3-2001</i> fora da amostra.	163
Figura 134. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	163
Figura 135. Resultados Série <i>M3-2013</i> fora da amostra.	164
Figura 136. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	164
Figura 137. Resultados Série <i>M3-2016</i> fora da amostra.	165
Figura 138. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	165
Figura 139. Resultados Série <i>M3-2089</i> fora da amostra.	166
Figura 140. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	166
Figura 141. Resultados Série <i>M3-2579</i> fora da amostra.	167
Figura 142. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.	167
Figura 143. Evolução do sMAPE para cada modelo.	168
Figura 144. Diagrama de barras comparativas para as séries anuais.	168
Figura 145. Diagrama de barras comparativas para as séries trimestrais.	169
Figura 146. Diagrama de barras comparativas para as séries mensais.	169
Figura 147. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-101</i> .	183
Figura 148. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-102</i> .	183
Figura 149. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-103</i> .	184
Figura 150. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-104</i> .	184
Figura 151. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-105</i> .	185
Figura 152. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-106</i> .	185
Figura 153. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-107</i> .	186
Figura 154. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-108</i> .	186
Figura 155. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-109</i> .	187
Figura 156. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-110</i> .	187
Figura 157. Conjuntos de Treinamento, Validação e Teste da Série <i>NN3-111</i> .	188
Figura 158. Conjuntos de Treinamento, Validação e Teste da Série <i>M3-0033</i> .	189

Figura 159. Conjuntos de Treinamento, Validação e Teste da Série M3-0050.	190
Figura 160. Conjuntos de Treinamento, Validação e Teste da Série M3-0070.	190
Figura 161. Conjuntos de Treinamento, Validação e Teste da Série M3-0105.	190
Figura 162. Conjuntos de Treinamento, Validação e Teste da Série M3-0116.	191
Figura 163. Conjuntos de Treinamento, Validação e Teste da Série M3-0119.	191
Figura 164. Conjuntos de Treinamento, Validação e Teste da Série M3-0162.	192
Figura 165. Conjuntos de Treinamento, Validação e Teste da Série M3-0171.	192
Figura 166. Conjuntos de Treinamento, Validação e Teste da Série M3-0200.	193
Figura 167. Conjuntos de Treinamento, Validação e Teste da Série M3-0351.	193
Figura 168. Conjuntos de Treinamento, Validação e Teste da Série M3-0564.	194
Figura 169. Conjuntos de Treinamento, Validação e Teste da Série M3-0599.	194
Figura 170. Conjuntos de Treinamento, Validação e Teste da Série M3-0642.	195
Figura 171. Conjuntos de Treinamento, Validação e Teste da Série M3-0671.	195
Figura 172. Conjuntos de Treinamento, Validação e Teste da Série M3-0671.	196
Figura 173. Conjuntos de Treinamento, Validação e Teste da Série M3-0692.	196
Figura 174. Conjuntos de Treinamento, Validação e Teste da Série M3-0713.	197
Figura 175. Conjuntos de Treinamento, Validação e Teste da Série M3-0730.	197
Figura 176. Conjuntos de Treinamento, Validação e Teste da Série M3-0758.	198
Figura 177. Conjuntos de Treinamento, Validação e Teste da Série M3-0773.	198
Figura 178. Conjuntos de Treinamento, Validação e Teste da Série M3-0829.	199
Figura 179. Conjuntos de Treinamento, Validação e Teste da Série M3-0847.	199
Figura 180. Conjuntos de Treinamento, Validação e Teste da Série M3-0854.	200
Figura 181. Conjuntos de Treinamento, Validação e Teste da Série M3-0865.	200
Figura 182. Conjuntos de Treinamento, Validação e Teste da Série M3-0896.	201
Figura 183. Conjuntos de Treinamento, Validação e Teste da Série M3-1430.	201
Figura 184. Conjuntos de Treinamento, Validação e Teste da Série M3-1460.	202
Figura 185. Conjuntos de Treinamento, Validação e Teste da Série M3-1725.	202
Figura 186. Conjuntos de Treinamento, Validação e Teste da Série M3-1900.	203
Figura 187. Conjuntos de Treinamento, Validação e Teste da Série M3-1906.	203
Figura 188. Conjuntos de Treinamento, Validação e Teste da Série M3-1912.	204
Figura 189. Conjuntos de Treinamento, Validação e Teste da Série M3-1996.	204
Figura 190. Conjuntos de Treinamento, Validação e Teste da Série M3-2001.	205
Figura 191. Conjuntos de Treinamento, Validação e Teste da Série M3-2001.	205
Figura 192. Conjuntos de Treinamento, Validação e Teste da Série M3-2016.	206
Figura 193. Conjuntos de Treinamento, Validação e Teste da Série M3-2089.	206
Figura 194. Conjuntos de Treinamento, Validação e Teste da Série M3-2579.	207

Lista de tabelas

Tabela 1. Quantitativo de redes testadas por série.	39
Tabela 2. Parâmetros de entrada para gerar a série <i>Mackey-Glass</i> (Matlab versão 2013a).....	41
Tabela 3. Valores do coeficiente de assimetria para as 3 séries utilizadas.	43
Tabela 4. Limites dos parâmetros do Cromossomo.	43
Tabela 5. Parâmetros otimizados por GA para as <i>ESNs</i> base das 3 séries utilizadas.	44
Tabela 6. Resultados com conexões e pesos variáveis. ($\#U$).....	46
Tabela 7. Resultados com conexões fixas e pesos variáveis. ($\#U$)	48
Tabela 8. Resultados com conexões e pesos variáveis ($\%C = 20$).....	51
Tabela 9. Resultados com conexões fixas e pesos variáveis. ($\%C = 20$)	53
Tabela 10. Resultados com conexões e pesos variáveis. ($\%C = 35$).....	55
Tabela 11. Resultados com conexões fixas e pesos variáveis. ($\%C = 35$)	57
Tabela 12. Resultados com conexões e pesos variáveis. ($\%C = 50$).....	59
Tabela 13. Resultados com conexões fixas e pesos variáveis. ($\%C = 50$)	61
Tabela 14. Resultados com conexões e pesos variáveis. ($\%C = 75$).....	63
Tabela 15. Resultados com conexões fixas e pesos variáveis. ($\%C = 75$)	65
Tabela 16. Resultados com conexões e pesos variáveis. ($\%C = 100$).....	67
Tabela 17. Resultados conexões fixas e pesos variáveis. ($\%C = 100$)	69
Tabela 18. Resultados com conexões e pesos variáveis. (sWe).....	71
Tabela 19. Resultados com conexões fixas e pesos variáveis. (sWe)	73
Tabela 20. Resultados com conexões e pesos variáveis. ($sWrr$)	75
Tabela 21. Resultados com conexões fixas e pesos variáveis. ($sWrr$)	77
Tabela 22. Resultados com conexões e pesos variáveis. ($\rho(W)$)	79
Tabela 23. Resultados com conexões fixas e pesos variáveis. ($\rho(W)$)	81
Tabela 24. Matriz de distancias entre padrões.	93
Tabela 25. Informações gerais das 11 séries <i>NN3</i> utilizadas.	95
Tabela 26. Informações gerais das 12 séries <i>M3</i> com amostragem anual.	95
Tabela 27. Informações gerais das 12 séries <i>M3</i> com amostragem trimestral.	95
Tabela 28. Informações gerais das 12 séries <i>M3</i> com amostragem mensal.	96
Tabela 29. Limites dos parâmetros do Cromossomo.	96
Tabela 30. Resultados comparativos para as séries <i>NN3</i> com os modelos utilizados.	97
Tabela 31. Resultados comparativos para cada série <i>NN3</i>	97
Tabela 32. Parâmetros finais de configuração para a série <i>NN3-101</i>	98
Tabela 33. Parâmetros finais de configuração para a série <i>NN3-102</i>	99
Tabela 34. Parâmetros finais de configuração para a série <i>NN3-103</i>	101
Tabela 35. Parâmetros finais de configuração para a série <i>NN3-104</i>	103
Tabela 36. Parâmetros finais de configuração para a série <i>NN3-105</i>	105
Tabela 37. Parâmetros finais de configuração para a série <i>NN3-106</i>	107

Tabela 38. Parâmetros finais de configuração para a série <i>NN3-107</i> .	109
Tabela 39. Parâmetros finais de configuração para a série <i>NN3-108</i> .	111
Tabela 40. Parâmetros finais de configuração para a série <i>NN3-109</i> .	113
Tabela 41. Parâmetros finais de configuração para a série <i>NN3-110</i> .	115
Tabela 42. Parâmetros finais de configuração para a série <i>NN3-111</i> .	117
Tabela 43. Resultados comparativos para as séries <i>NN3</i> .	120
Tabela 44. Resultados para cada uma das séries.	120
Tabela 45. Resultados dos primeiros 10 lugares da competição <i>NN3</i> .	121
Tabela 46. Resultados do teste <i>t-student</i> para as séries <i>NN3</i> .	122
Tabela 47. Resultados do teste <i>sync</i> para as séries <i>NN3</i> .	
Tabela 48. Resultados do teste <i>Wilcoxon</i> para as séries <i>NN3</i> .	122
Tabela 49. Resultados para o teste Friedman.	123
Tabela 50. Resultados para o teste Post-Hoc.	123
Tabela 51. Séries selecionadas para o grupo de amostragem anual.	124
Tabela 52. Séries selecionadas para o grupo de amostragem trimestral.	124
Tabela 53. Séries selecionadas para o grupo de amostragem mensal.	125
Tabela 54. Resultados comparativos para as séries <i>M3</i> .	125
Tabela 55. Resultados comparativos para cada série <i>M3</i> anual.	125
Tabela 56. Resultados comparativos para cada série <i>M3</i> trimestral.	126
Tabela 57. Resultados comparativos para cada série <i>M3</i> mensal.	126
Tabela 58. Comparação do <i>sMAPE</i> para as séries anuais.	127
Tabela 59. Comparativa do <i>sMAPE</i> para as séries trimestrais.	127
Tabela 60. Comparativa do <i>sMAPE</i> para as séries mensais.	127
Tabela 61. Parâmetros finais de configuração para a série <i>M3-0033</i> .	128
Tabela 62. Parâmetros finais de configuração para a série <i>M3-0050</i> .	129
Tabela 63. Parâmetros finais de configuração para a série <i>M3-0070</i> .	130
Tabela 64. Parâmetros finais de configuração para a série <i>M3-0105</i> .	131
Tabela 65. Parâmetros finais de configuração para a série <i>M3-0116</i> .	132
Tabela 66. Parâmetros finais de configuração para a série <i>M3-0119</i> .	133
Tabela 67. Parâmetros finais de configuração para a série <i>M3-0162</i> .	134
Tabela 68. Parâmetros finais de configuração para a série <i>M3-0171</i> .	135
Tabela 69. Parâmetros finais de configuração para a série <i>M3-0200</i> .	136
Tabela 70. Parâmetros finais de configuração para a série <i>M3-0351</i> .	137
Tabela 71. Parâmetros finais de configuração para a série <i>M3-0564</i> .	138
Tabela 72. Parâmetros finais de configuração para a série <i>M3-0599</i> .	139
Tabela 73. Parâmetros finais de configuração para a série <i>M3-0642</i> .	140
Tabela 74. Parâmetros finais de configuração para a série <i>M3-0671</i> .	142
Tabela 75. Parâmetros finais de configuração para a série <i>M3-0677</i> .	143
Tabela 76. Parâmetros finais de configuração para a série <i>M3-0692</i> .	144
Tabela 77. Parâmetros finais de configuração para a série <i>M3-0713</i> .	145

Tabela 78. Parâmetros finais de configuração para a série M3-0730.	146
Tabela 79. Parâmetros finais de configuração para a série M3-0758.	147
Tabela 80. Parâmetros finais de configuração para a série M3-0773.	148
Tabela 81. Parâmetros finais de configuração para a série M3-0829.	149
Tabela 82. Parâmetros finais de configuração para a série M3-0847.	150
Tabela 83. Parâmetros finais de configuração para a série M3-0854.	151
Tabela 84. Parâmetros finais de configuração para a série M3-0865.	152
Tabela 85. Parâmetros finais de configuração para a série M3-0896.	153
Tabela 86. Parâmetros finais de configuração para a série M3-1430.	155
Tabela 87. Parâmetros finais de configuração para a série M3-1460.	156
Tabela 88. Parâmetros finais de configuração para a série M3-1725.	157
Tabela 89. Parâmetros finais de configuração para a série M3-1900.	158
Tabela 90. Parâmetros finais de configuração para a série M3-1906.	159
Tabela 91. Parâmetros finais de configuração para a série M3-1912.	160
Tabela 92. Parâmetros finais de configuração para a série M3-1996.	161
Tabela 93. Parâmetros finais de configuração para a série M3-2001.	162
Tabela 94. Parâmetros finais de configuração para a série M3-2013.	163
Tabela 95. Parâmetros finais de configuração para a série M3-2016.	164
Tabela 96. Parâmetros finais de configuração para a série M3-2089.	165
Tabela 97. Parâmetros finais de configuração para a série M3-2579.	166
Tabela 98. Resultados do teste <i>t-student</i> para as séries M3 com amostragem anual.	170
Tabela 99. Resultados do teste <i>sync</i> para as séries M3 com amostragem anual.	170
Tabela 100. Resultados do teste <i>Wilcoxon</i> para as séries M3 com amostragem anual.	170
Tabela 101. Resultados para o teste Friedman.	171
Tabela 102. Resultados para o teste Post-Hoc.	171
Tabela 103. Resultados do teste <i>t-student</i> para as séries M3 com amostragem trimestral.	172
Tabela 104. Resultados do teste <i>Sync</i> para as séries M3 com amostragem trimestral.	172
Tabela 105. Resultados do teste <i>Wilcoxon</i> para as séries M3 com amostragem trimestral. ...	172
Tabela 106. Resultados para o teste Friedman.	172
Tabela 107. Resultados para o teste Post-Hoc.	173
Tabela 108. Resultados do teste <i>t-student</i> para as séries M3 com amostragem mensal.	173
Tabela 109. Resultados do teste <i>Sync</i> para as séries M3 com amostragem mensal.	174
Tabela 110. Resultados do teste <i>Wilcoxon</i> para as séries M3 com amostragem mensal.	174
Tabela 111. Resultados para o teste Friedman.	174
Tabela 112. Resultados para o teste Post-Hoc.	175

1

Introdução

1.1

Motivação

A utilização dos modelos de Redes Neurais Artificiais Recorrentes (*RNARs*) tem se expandido para diversas áreas e aplicações, como por exemplo: reconhecimento de padrões (Graves, 2009), Previsão de sistemas caóticos em comunicações sem fio (Jaeger, 2004), Representação dinâmica das primitivas da robótica (Paine, 2004), Identificação e controle de processos não lineares em sistemas dinâmicos (Tutschku, 1995), Modelagem não-linear auto regressiva para aplicações de previsão (Jeong, 2008), Controle no deslocamento de bípedes (Santos, 2012), Agrupamento e Previsão de Séries temporais (Cherif, 2011), Ferramenta de apoio à decisão para compras industriais (Zimmerman, 2012), Ferramenta de apoio à decisão para diagnóstico de Laringopatias (Szkola, 2011) e Reconhecimento de dígitos falados (Graves, 2013), além de outros tantos. Um dos casos mais relevantes é a previsão de séries temporais, cuja utilização foi intensificada após os resultados obtidos no “*Artificial Neural Network & Computational Intelligence Forecasting Competition – NN3*”, onde estas redes conseguiram o primeiro lugar na versão completa (111 séries) e segundo lugar na versão reduzida (11 séries). Assim, este tipo de aplicação constituiu-se em um dos *Benchmark* mais utilizados para avaliar a capacidade computacional dessas redes neurais (Jaeger, 2007).

Apresentadas por (Hopfield, 1982), as redes que levam seu nome foram o ponto de partida para a posterior apresentação dos primeiros modelos de redes recorrentes, como o caso das *Jordan Networks* introduzidas por (Jordan, 1986), as *Bidirectional associative memory* (*BAM*), que foram expostas por (Kosko, 1988), e as *Elman Networks* que foram apresentadas por (Elman, 1990).

Posteriormente, (Dominey, 1995) apresentou as *Temporal Recurrent Neural Networks* como mais um modelo de redes recorrentes, demonstrando o potencial deste tipo de rede na resolução de diversos problemas. Porém, o alto custo computacional necessário para realizar o treinamento fez com que ficassem fora do foco da maioria dos pesquisadores. Recentemente, o termo “*Reservoir Computing*” (*RC*) foi apresentado à comunidade científica como uma nova categoria de redes recorrentes que visa a explorar as propriedades dessas redes com um menor custo computacional. Dentre os modelos desta nova categoria destacam-se as redes *Echo State Networks* (*ESNs*) propostas por (Jaeger, 2001), as redes *Liquid State Machines* (*LSM*) propostas por (Maass e Natschläger, 2002) e as redes *Decorrelation-Backpropagation Learning* introduzidas por (Steil, 2004), sendo que uma das maiores contribuições foi a redução

do custo computacional, com a utilização de novos modelos de treinamento. É importante ressaltar que as redes *Extreme Learning*, apesar de não serem consideradas parte da família *Reservoir Computing*, têm um princípio de operação similar às *ESNs*.

As *Echo State Networks (ESNs)* foram as redes que conseguiram o melhor desempenho em previsão *Multi-Step* no “*Artificial Neural Network & Computational Intelligence Forecasting Competition – NN3*”, transformando-as em um modelo potencial para diversas outras aplicações, conforme pode-se observar nas publicações abaixo:

1. Controle (Salmen e Ploger, 2005), (Han, 2014).
2. Previsão de Séries Temporais (Jianhui Xi et al. 2005), (de Aquino et al. 2015).
3. Reconhecimento de padrões (Verstraeten et al. 2005), (Wen, 2015).
4. Identificação de Sistemas (Bian e Mou, 2011), (Soh, 2015).
5. Modelagem (Bai Zhang et al. 2012), (Magerl et al. 2015).
6. Classificação (Mitul et al, 2013), (Shi et al. 2015).

Os diversos trabalhos realizados com essas redes demonstraram que, devido à natureza das arquiteturas das *ESNs*, estas têm grande potencial para resolver problemas em que os dados possuem atrelamento temporal, com tempo computacional reduzido. Entretanto, algumas características ainda geram questionamentos, como por exemplo, a configuração dos parâmetros gerais da rede (Jaeger, 2005), (Prokhorov, 2005), o tipo de *Reservoir* a ser utilizado (Qiao, 2016), o comportamento dos estados internos do *Reservoir* e a contribuição das conexões recorrentes no mapeamento dos padrões apresentados (Ishii, 2004), (Yao, 2014), (Yuenyong, 2015).

Como será detalhado neste trabalho, todos estes aspectos foram investigados, estudados e analisados, visando a obter uma caracterização mais adequada do comportamento dos parâmetros neste tipo de redes. Alguns artigos na literatura recente como (Quiao, 2016), (Hongyun, 2015), (Jaeger, 2013), (Jaeger, 2005) vêm abordando aspectos relativos aos parâmetros deste tipo de rede neural e como alguns podem ser otimizados utilizando computação evolucionária. Entretanto, ainda não foi apresentada uma pesquisa que tenha sido mais efetiva ou completa acerca desta questão.

O comportamento do *Reservoir* também é foco deste trabalho, onde a topologia utilizada e os pesos internos são avaliados para determinar o desempenho dos estados internos quando as entradas são apresentadas. Assim, após a avaliação dos diversos trabalhos na literatura e mediante a constatação da ausência de uma teoria abrangendo todos estes conceitos, determinou-se a principal motivação para o desenvolvimento desta pesquisa.

1.2

Objetivos

Apesar de se observar as lacunas ainda abertas com relação à modelagem deste novo tipo de redes neurais, não se pretende com esse trabalho desenvolver uma análise que envolva a parametrização das *ESNs* para qualquer tipo de aplicação. Sendo assim, o principal objetivo do presente trabalho consiste em desenvolver e avaliar um modelo global de otimização conjunta, que contemple a identificação da melhor configuração de parâmetros e topologia do Reservoir para as *ESNs*, aplicado em previsão *Multi-Step*, onde dependendo do tamanho da janela faz com que valores previstos pela *ESN* sejam utilizados como entradas para previsões de passos a frente.

Os objetivos secundários são:

- a) Realizar uma análise de sensibilidade envolvendo a configuração dos parâmetros das *ESNs* para tarefas de previsão *Multi-Step* em séries temporais;
- b) Efetuar um estudo do comportamento dos estados do *Reservoir* que permita avaliar quais topologias atingem os melhores resultados para tarefas de previsão *Multi-Step* em séries temporais;

1.3

Contribuições

As principais contribuições desse trabalho são:

- a) Desenvolvimento de uma metodologia que permita a configuração automática das *ESNs*, considerando os diversos parâmetros e a topologia, determinando a contribuição de cada variável observada na previsão *Multi-Step* de séries temporais com características diferenciadas.
- b) Desenvolvimento de um algoritmo de otimização conjunta para avaliar o desempenho das topologias de *Reservoir*, fundamentado na métrica *Separation Ratio Graph* (Gibbons, 2010).
- c) Desenvolvimento de uma análise de sensibilidade que permita caracterizar o grau de influência dos parâmetros gerais na configuração das *ESNs* para previsão de séries temporais.

1.4

Organização do Trabalho

Este trabalho está dividido em mais cinco capítulos, descritos a seguir:

No capítulo 2 faz-se um resumo dos fundamentos teóricos necessários para uma melhor compreensão do trabalho. Basicamente, os conceitos fundamentais de *Reservoir Computing* e os modelos que o compõem são apresentados, assim como métodos e métricas de avaliação dos estados do *Reservoir*.

No capítulo 3 é apresentada a análise de sensibilidade realizada para determinar a contribuição de cada um dos parâmetros de configuração das *ESNs*.

No capítulo 4 apresenta-se, detalhadamente, o modelo proposto de otimização de parâmetros gerais e seleção de topologias para previsão de séries temporais, denominado ESN-GA-SRG.

No capítulo 5 são apresentados os resultados obtidos, os quais permitem validar o desempenho do modelo proposto para previsão *Multi-Step* de séries temporais.

No capítulo 6 são apresentadas as conclusões e as perspectivas de trabalhos futuros a serem desenvolvidos.

2.

Reservoir Computing

Esse capítulo tem por objetivo apresentar a fundamentação do tema “*Reservoir Computing*”. Pretende-se estabelecer uma base teórica e conceitual para o desenvolvimento dos capítulos posteriores.

Reservoir Computing (RC) é um tipo de rede neural recorrente que, diferentemente das redes *Multi Layer Perceptron (MLP)*, apresentam na camada escondida um grande conjunto de neurônios parcialmente interconectados entre si e com conexões destes neurônios para as camadas de entrada e de saída.

Os princípios fundamentais das redes do tipo *RC* podem ser resumidos em:

- a) Utiliza uma grande quantidade de neurônios (conhecido como *Reservoir*), acionados por meio de sinais de entrada, sendo que cada unidade do *Reservoir* cria a sua própria transformação não linear da entrada considerando a transformação gerada pelo padrão anterior. Esse processo acontece em um ambiente com uma dimensão bem mais alta do que a utilizada originalmente.
- b) Sinais de saída são obtidos a partir de um *Reservoir* excitado utilizando, geralmente, uma simples combinação linear.

A seção a seguir fará uma apresentação sucinta dos diferentes modelos de RNR e na seção 2.2 apresentará em mais detalhes as ESN, uma vez que foram as escolhidas para o desenvolvimento deste trabalho pela vantagem oferecida no que tange ao custo computacional.

2.1.

Modelos de Redes Neurais Recorrentes

Nas seções a seguir serão apresentados e descritos os modelos de redes neurais recorrentes que formam parte da categoria *Reservoir Computing*; a ordem de apresentação destes modelos foi organizada cronologicamente, isto é, o modelo mais antigo é apresentado inicialmente e posteriormente os modelos subsequentes.

2.1.1.

Temporal Recurrent Neural Network

Este tipo de rede, apresentado por Peter F. Dominey em 1995, foi o modelo pioneiro de *RC*, embora o termo *Reservoir Computing* não fosse utilizado na época. Entretanto, a metodologia utilizada para o projeto e treinamento das *Temporal Recurrent Neural Networks* (TRNNs) compartilha os mesmos princípios de *RC*. A partir do momento que outros modelos foram desenvolvidos e a categoria *Reservoir Computing* foi utilizada para redes recorrentes com os princípios descritos anteriormente, o modelo *Temporal Recurrent Neural Network* foi encaixado nesta categoria. Apresentado como um estudo de caso, utilizou como benchmark a aplicação “*Complex Sensory-Motor Sequence Learning*”. Este modelo é composto por duas sub-redes: a primeira “*pre-frontal córtex (PFC)*” é responsável por gerar padrões únicos de atividade que representam estados contínuos da execução de uma sequência; já a segunda sub-rede, “*Stratium*”, associa os padrões de estados codificados com uma resposta para cada ponto de execução da sequência.

A partir da perspectiva da Neurociência, este modelo baseia-se no conhecimento cortical e sub-cortical do sistema óculomotor de um primata, os quais são os encarregados de transformar observações em padrões de entrada, que posteriormente são representados por excitações (estados) e finalmente são associados com uma resposta.

O treinamento é realizado da mesma forma que no mundo animal, sendo que o comportamento desejado é moldado gradualmente: os padrões são gerados pelas informações capturadas pela retina, o *PFC* (, que basicamente é o *Reservoir*) cria os estados internos e a sub-rede “*Stratium*” associa estes estados com os *targets*.

Na estratégia utilizada para o aprendizado, três parâmetros são considerados importantes segundo (Dominey, 1995):

- a) **Taxa de aprendizado:** No processo inicial é utilizada uma taxa com um valor alto, que permite ter uma relação mais forte dos estados gerados, mesmo quando as sequências apresentadas contenham erros. Após o processo inicial a taxa de aprendizado vai decrescendo.
- b) **Taxa de esquecimento:** Inicialmente tem um valor pequeno e vai aumentando gradualmente conforme o treinamento é realizado, permitindo manter controladamente a informação de estados anteriores.
- c) **Formato de treinamento:** No período de treinamento cada sequência é apresentada em conjunto com a resposta certa (*target*); já no período de não instrução (pós-treinamento), a sequência é apresentada e o modelo deve selecionar uma de quatro respostas possíveis selecionadas pelo projetista da rede. Se a resposta selecionada for a correta, vai passar para o treinamento de uma

nova sequência; caso contrário, o processo de treinamento será repetido novamente para a sequência que não conseguiu a resposta certa.

Em geral o modelo *TRNN* apresentou resultados satisfatórios para o benchmark selecionado, mas o tempo de treinamento (“instrução”) pode-se estender por longos períodos. Tal dificuldade fez com que inicialmente este modelo não fosse muito utilizado, mas o que se deseja destacar é que o princípio de *Reservoir Computing* faz parte deste modelo.

2.1.2.

Echo State Networks

Este modelo, apresentado pelo Herbert Jaeger em 2001, tornou-se amplamente utilizado pela simplicidade no treinamento e os excelentes resultados obtidos em *benchmarks* como “*Japanese Vowel Competition*” (Jaeger, 2007) onde obteve uma taxa de erro de 0% no conjunto de teste e o “*International Forecasting Competition NN3*” (Ilies, 2007) onde alcançou a segunda colocação, dentre 45 modelos, na previsão *Multi-Step* de séries temporais para a versão reduzida.

Este tipo de rede foi inspirado em experimentos neuropsicológicos (Jaeger, 2005), onde uma das características é a presença de pelo menos uma conexão recorrente, ou seja, o valor obtido na camada de saída no tempo n é utilizado na camada de entrada e/ou no reservatório no tempo $n + 1$. Na figura 1 é apresentado o esquema de conexão nas redes *ESNs*.

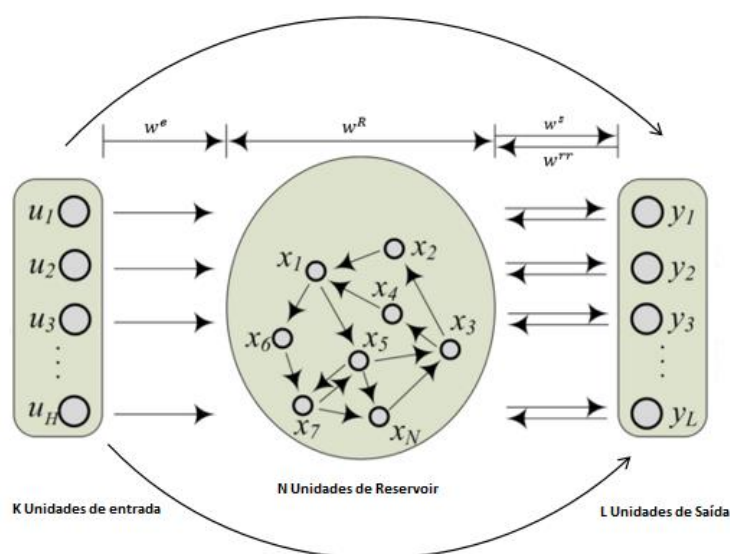


Figura 1. *Reservoir Computing*. (Gang, 2015)

O elemento mais importante nas *ESNs* é o *Reservoir*, o qual contém um grande número de unidades internas interconectadas aleatoriamente e/ou autoconectadas. Os pesos das sinapses dos neurônios do *Reservoir*, após terem sido inicializados, permanecem inalterados. Somente pesos das conexões da camada de saída são modificados durante o processo de treinamento.

A atualização de pesos pode ser realizada *off-line*, utilizando uma regressão linear, ou *on-line* utilizando o método de mínimos quadrados recursivos (Cernasky, 2008), oferecendo baixo custo computacional quando treinadas *off-line*.

2.1.3.

Liquid State Machines

Também conhecidas como máquinas de estado líquido, foram apresentadas por Maass e Natschläger em 2002 como resultados de pesquisas em redes recorrentes. Como princípio de operação, é feita uma analogia ao comportamento gerado por uma pedra (entrada) quando lançada em uma lagoa (meio líquido). Nessa analogia ocorrem ondulações no meio líquido, as quais são observadas como uma reação (estado líquido). As ondulações são propagadas pela superfície da água e têm a possibilidade de interagir com outras já existentes; portanto, a água permitiria manter e integrar as informações dos eventos acontecidos. Quando essas ondulações são capturadas por uma câmera (unidade de leitura), podem fornecer informações dos eventos, e assim associar estes eventos às ondulações observadas (Maass e Natschläger, 2002).

Utilizando uma rede *Liquid State Machine (LSM)* é possível transformar uma entrada temporal de baixa dimensão em uma dimensão maior gerada pelos estados líquidos. Além disso, adiciona-se informações sobre os estímulos acontecidos no passado ao sinal de entrada atual.

Este tipo de redes é composto por duas partes principais: o meio líquido e a unidade de leitura. O meio líquido, também conhecido como “Micro-circuito neural”, pode ser analogamente visto como o *Reservoir* das *ESNs*. Ele possui uma dinâmica complexa que gera estados líquidos a partir dos dados apresentados nas entradas, também é importante destacar que é aqui onde se apresenta a recorrência dos sinais. A unidade de leitura é basicamente a função que decodifica os estados líquidos em um tempo t para posteriormente gerar a resposta da rede.

A utilização de neurônios com funções de ativação pulsante, quer dizer do tipo *Integrate-and-Fire* (Grando, 2010) que também podem ser usados nas *ESNs*, faz com que seja gerado um sistema dinâmico complexo que realiza um processamento em “tempo real” sobre entradas contínuas. Nessa operação, uma entrada $U(\cdot)$ é apresentada ao meio líquido; os estados líquidos $X(n)$ são gerados para cada instante de tempo t e, finalmente,

estes estados são mapeados por uma função de leitura f para, assim, produzir a função de saída $Y(n)$. Na figura 2 é apresentada a estrutura de uma rede LSM.

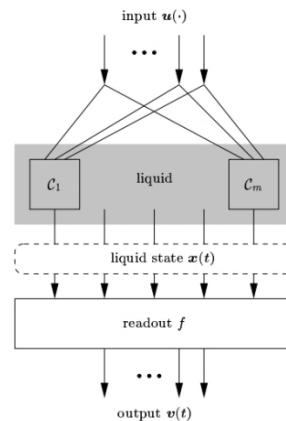


Figura 2. Estrutura LSM (Maass, 2002).

O meio líquido geralmente é representado como um arranjo tridimensional, que contém um ou mais grupos de neurônios (piscina/as) onde as conexões podem ser arbitrárias o que apresenta uma diferença com as *ESNs*. As piscinas estão compostas por neurônios os quais são distribuídos uniformemente, dependendo da topologia, e as dimensões que são definidas *a priori*. Os principais tipos de neurônios que o meio líquido utiliza são: *Leaky Integrate-and-fire (LIF)* (Wang, 2014), *Hodgkin-Huxley Model (HH)* (Csercsik, 2009) e *Spike-response Model (SRM)* (Feldbusch, 2009) onde as funções de ativação têm uma resposta diferente para cada modelo mencionado e as conexões sinápticas podem ser estáticas ou dinâmicas (Grando, 2010). No caso das conexões sinápticas estáticas, a amplitude das respostas pós-sinápticas é constante; no caso das conexões dinâmicas, as respostas dos estados têm um acréscimo, considerando o estado atual e a taxa de esquecimento que pode ser considerado análogo ao leaking rate nas *ESNs*. Na figura 3 são apresentados 2 exemplos de topologias do meio Líquido.

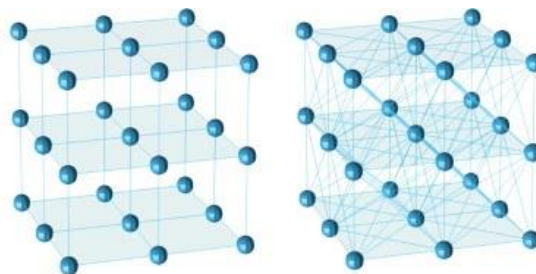


Figura 3. Exemplos topologias do médio líquido. (Ju, 2013).

A unidade de leitura pode ser realizada a partir de uma rede *Multilayer Perceptron (MLP)*, a qual vai receber os estados líquidos como entrada e gera uma ou mais saídas necessárias para a tarefa. Em alguns casos, se faz necessário utilizar um filtro antes da

unidade de leitura, pois, altas frequências devem ser eliminadas de modo que os estados líquidos sejam adequados para a análise.

O treinamento é realizado utilizando técnicas de aprendizado que permitam mapear a saída do meio líquido para a saída desejada. A complexidade da unidade de leitura vai depender das informações fornecidas pelos estados líquidos. Assim, estas podem não ser linearmente separáveis, o que induz à utilização de modelos mais complexos, como por exemplo, os modelos propostos por Maass (2002) que utiliza *Pools of Perceptrons* ou segundo (Goldenholtz, 2002) que utiliza redes *MLP*.

Concretamente este modelo tem maior aplicação na área de inferência, para aplicações de previsão de series temporais a utilização de neurônios pulsados representa uma nova área de desenvolvimento.

2.1.4.

Decorrelation-Backpropagation Learning

Apresentada por Steil em 2004, a regra de treinamento *Decorreltion-Backpropagation (BPDC)* é aplicada a redes similares às *ESNs* ou *LSMs*, com a diferença que as unidades do *Reservoir* estão 100% conectadas entre si, isto é, todas as unidades estão conectadas a todas as outras unidades dentro do *Reservoir*. No entanto, da mesma forma que nas *ESNs*, apenas os pesos da camada de saída são ajustados. A partir do princípio da decorrelação que deve existir entre os estados gerados e as entradas, torna-se possível que a camada de saída combine estes estados para obter a saída desejada.

Na Figura 4, apresenta-se a estrutura proposta em (Steil, 2004), onde a rede é 100% conectada e somente os pesos da camada de saída podem ser treinados.

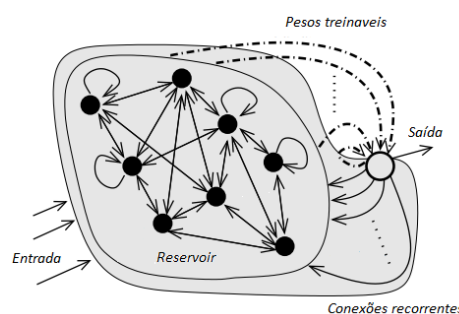


Figura 4. Estrutura das Redes BPDC. (Steil, 2004).

Pesquisas posteriores apresentaram resultados concretos na estabilidade da regra de aprendizado *on-line* (Steil, 2006), assim como a adaptação *on-line* do *Reservoir* mediante a propriedade de plasticidade intrínseca que é a adaptação da excitabilidade dos neurônios de forma conjunta com os estímulos recebidos (Steil, 2007).

Os modelos que foram apresentados compõem a recente área denominada *Reservoir Computing*, onde todos oferecem grandes vantagens quando observadas as potencialidades que oferecem as configurações de redes recorrentes. Porém, observando o quesito de custo computacional, as *ESNs* se apresentam como as melhores (menor custo computacional), sendo, atualmente, o modelo de rede neural recorrente mais utilizado. Assim, as redes *ESNs* são apresentadas em mais detalhes na seção a seguir, discutindo o processo de atualização de estados e o algoritmo de treinamento.

2.2.

Treinamento das *Echo State Networks*

Como apresentado na seção anterior, a grande vantagem das *ESNs* sobre as outras redes recorrentes da área de *Reservoir Computing* é o baixo custo computacional quando o treinamento é realizado *offline*. Nesta seção são apresentados com maior detalhe os processos de atualização de estados e treinamento, assim como as características principais de configuração.

O estado de ativação dos neurônios no *Reservoir* é atualizado quando cada padrão é apresentado. Na equação 2.1 são apresentados todos os elementos que influenciam o estado de ativação dos neurônios do *Reservoir*.

$$\mathbf{X}(n) = (1 - \alpha)\mathbf{X}(n - 1) + \alpha \varphi(w^e \mathbf{U}(n) + w^R \mathbf{X}(n - 1) + w^{rr} \mathbf{D}(n - 1)) \quad (2.1)$$

Onde $x(n)$ é o estado interno do *Reservoir* no tempo n ; α é o *Leaking Rate*, cujo o objetivo é controlar a velocidade com que os estados são atualizados quando neurônios do tipo *Integrate-and-Fire* são utilizados; $\varphi(\cdot)$ é a função de ativação dos neurônios do *Reservoir*; $\mathbf{D}(n - 1)$ é o potencial gerado nas unidades de saída no tempo $n - 1$; $\mathbf{U}(n)$ é o vetor de entrada; w^e , w^R e w^{rr} são os pesos para as conexões de entrada, *Reservoir* e conexões recorrentes, que são inicializados aleatoriamente com uma distribuição uniforme.

Segundo (Jaeger, 2002), considerando algumas condições de configuração da rede, o estado de ativação $\mathbf{X}(n)$ de uma rede neural recorrente é uma função (infinita) de históricos de entrada $\mathbf{U}(n), \mathbf{U}(n - 1), \dots$ apresentados à rede. A condição fundamental para que estes estados contenham informação útil é que deve existir a *Echo Function* $E = (e_1, \dots, e_N)$, onde $e_i: U^{\mathbb{N}} \rightarrow \mathbb{R}$, ou seja, o histórico está composto por todas as entradas anteriores que pertencem ao conjunto dos números reais. O efeito do estado anterior deve desaparecer gradualmente à medida que o tempo transcorre e não deve persistir.

Uma forma de manter esta propriedade é, segundo (Jaeger, 2001), garantir que o raio espectral $\rho(W)$, que é o maior valor absoluto entre os autovalores da matriz de pesos do *Reservoir*, seja menor do que 1. No entanto, em (Ferreira, 2011) foi apresentado que esta condição nem sempre garante a propriedade de *Echo State*, o que tem levado ao equívoco de acreditar que tal parâmetro não deve passar do valor de 1 (Lukoševičius, 2009). O trabalho mencionado anteriormente (Ferreira, 2011) estabelece que “A conexão correta entre o raio espectral e a propriedade de *Echo State* é que a propriedade de *Echo State* é violada se $\rho(W) < 1$ apenas em RC que usam a função de ativação *tanh* e que não possuem entrada. Contrariamente a esse equívoco, a propriedade de *Echo State* pode ser obtida mesmo que $\rho(W) > 1$ em sistemas com entrada diferente de zero, e ela também pode ser perdida mesmo se $\rho(W) < 1$, embora seja muito difícil construir sistemas onde isso ocorra” (Ferreira, 2011).

As saídas das ESNs podem ser obtidas utilizando a equação 2.2.

$$D(n) = \varphi^s \left(w^s \cdot \begin{bmatrix} X(n) \\ D(n-1) \end{bmatrix} \right) \quad (2.2)$$

Onde, φ^s pode ser uma função linear ou uma sigmoide, e w^s é a matriz que contém os pesos das conexões de saída, que são determinados após o treinamento. Existem modelos que podem ter conexões da entrada diretamente para a saída, mas o modelo utilizado neste trabalho utiliza a configuração mais usual, ou seja, as ditas conexões entre a entrada e a saída não são consideradas.

Utilizando K unidades de entrada, N unidades no *Reservoir* e L unidades na camada de saída, as dimensões das matrizes de pesos são: $w^e \in \mathbb{R}^{N \times K}$, $w^R \in \mathbb{R}^{N \times N}$, $w^{rr} \in \mathbb{R}^{N \times L}$ e $w^s \in \mathbb{R}^{L \times N}$.

O treinamento *offline* é realizado em dois estágios. Primeiro são apresentados todos os padrões uma única vez e coletados os estados do *Reservoir* no intervalo $[A_p, \dots, A_q]$, onde A_p e A_q são os pontos de início e fim da faixa de amostragem (i.e. A_p é o primeiro e A_q é o último padrão do conjunto de treinamento), respectivamente. Estes estados são armazenados em uma matriz J de tamanho $(A_q - A_p) + 1 \times N$, como apresentado na equação 2.3.

$$J = \begin{bmatrix} x_{p,1} & \cdots & x_{p,N} \\ \vdots & \ddots & \vdots \\ x_{q,1} & \cdots & x_{q,N} \end{bmatrix} \quad (2.3)$$

No processo de treinamento, somente os pesos da camada de saída w^s são adaptados. Assim, após terem sido coletados os estados do *Reservoir*, todas as saídas (*targets*) são coletadas em uma matriz D . Uma vez obtidas as matrizes J e D , os pesos da camada de saída são obtidos com o método da matriz pseudo-inversa, segundo a equação 2.4.

$$w^s = (J^T J)^{-1} J^T D \quad (2.4)$$

O erro para o conjunto de teste (que também pode ser utilizado no conjunto de validação, caso necessário) é obtido utilizando a equação 2.5, onde a saída da rede é comparada com a saída desejada.

$$MSE = \frac{1}{(A_f - A_e) + 1} \sum_{n=e}^f (y(n) - d(n))^2 \quad (2.5)$$

Parâmetros adicionais foram criados e são utilizados nas redes *ESNs* com a finalidade de determinar configurações mais adequadas, isto é, ajustar componentes da rede de tal forma que possam ser obtidos os melhores resultados possíveis. Tais parâmetros são apresentados a seguir:

- a) **Conectividade:** Quando gerado o *Reservoir*, dois parâmetros devem ser especificados: o primeiro é o número de unidades (neurônios), e o segundo a porcentagem de conexão, isto é, a quantidade de conexões entre os neurônios no *Reservoir*. Geralmente a representação utilizada emprega uma matriz de $N \times N$ que contém, além das conexões, os pesos gerados aleatoriamente. Como foi recomendado pelo autor do modelo, uma conexão de 20% deve ser o valor mais baixo a ser utilizado (Jaeger, 2007), e assim garantir um *Reservoir* disperso que permita a decomposição da dinâmica em pequenos subsistemas acoplados, podendo obter até 100% de conexões.
- b) **Raio Espectral:** valor utilizado como fator de escala dos pesos das conexões dentro do *Reservoir*. Este valor procura manter a propriedade *Echo State* pela qual, segundo Jaeger, é possível reconstruir o sinal de entrada através dos estados coletados ao longo da apresentação dos padrões.
- c) **Fator de escala das conexões de entrada:** este parâmetro é utilizado com o intuito de privilegiar algumas conexões de entrada. Isso ocorre durante o processo de formação da matriz de estados quando os padrões são apresentados, e determina o grau de não linearidade da dinâmica do *Reservoir*. Se os pesos forem pequenos o comportamento pode ser próximo do linear; já se forem grandes podem saturar a função de ativação das unidades do *Reservoir*, alternando os resultados entre 0 e 1.
- d) **Fator de escala das conexões recorrentes:** utilizando este parâmetro é possível gerar uma retroalimentação adicional a partir das conexões entre a saída e as conexões do *Reservoir*. No modelo clássico, não existe este tipo de conexão entre

a saída e os neurônios do *Reservoir*. A existência desse tipo de realimentação pode conduzir a dinâmicas instáveis. No modelo clássico é recomendado utilizar esta conexão quando se tem problemas que precisam conservar a informação dos estados anteriores por mais tempo, além de dispor de informações que se encontram na saída (*target*) e que não estão disponíveis nos padrões de entrada (Breve, 2005).

2.2.1.

Métodos de avaliação da qualidade do *Reservoir*

Baseados na necessidade de avaliar a qualidade do *Reservoir*, diferentes pesquisadores têm apresentado modelos que permitem determinar métricas de desempenho para as redes dentro da categoria de *Reservoir Computing*. A seguir serão apresentados os métodos mais utilizados.

Capacidade de Memória: Uma medida da capacidade de memória de curta duração permite avaliar quão bem os padrões de entrada $u(n)$ podem ser reconstruídos pelo *Reservoir* como $y(n+k)$ após k atrasos (Jaeger, 2002).

Para uma rede *ESN* de uma única entrada, a capacidade de memória de curto prazo pode ser quantificada utilizando a equação 2.6, onde o termo $r^2(u(n-i), y_i(n))$ é o coeficiente de correlação quadrática entre o sinal de entrada atrasado por i e a saída (*target*) utilizada.

$$C = \sum_{i=1,2,\dots} r^2(u(n-i), y_i(n)) \quad (2.6)$$

O limite na capacidade de memória para uma rede *ESN* de tamanho N (número de unidades internas no *Reservoir*) está delimitado por N , isto é, a natureza auto-regressiva das conexões internas e as propriedades dos estados de conter parte da história do passado fazem com que o limite seja determinado pelo tamanho de unidades ativas dentro do *Reservoir*.

Eigenvalue Spread (EV): segundo (Jaeger, 2005), em alguns casos é observado que a matriz de correlação dos pesos do *Reservoir*, ou seja, a relação entre o autovalor máximo e o autovalor mínimo dos pesos do *Reservoir* pode atingir valores de 1×10^{12} , ou até maiores. Se a esta condição for adicionada a possibilidade de que os pesos treinados

na camada de saída sejam da ordem de 1×10^8 , o comportamento da rede seria, neste caso, caótico, portanto, a condição de que o valor de *EV* seja tão grande deve ser evitada, uma vez que:

- a) Valores muito altos de *EV* tornam inviáveis computacionalmente a utilização de algoritmos de treinamento *on-line*.
- b) Pesos de saída muito grandes implicam em uma diminuição na capacidade de generalização da rede, assim como requerem representações muito precisas nos estados do *Reservoir*.

Uma das técnicas utilizadas para reduzir o crescimento do *EV* é a adição de ruído aos padrões de entrada, mas isto faz com que a precisão do modelo diminua, em algumas outras pesquisas como (Barabasi, 1999), (Kaiser, 2004) e (Dunca, 1998) são estudados outros tipos de topologias como “Small World Networks” e “Real World Networks”.

2.2.2.

Métricas de avaliação dos estados do *Reservoir*

Diversas técnicas vêm sendo utilizadas para avaliar a qualidade dos estados gerados após a apresentação dos padrões para as redes que compõem a área de *Reservoir Computing*. Nesta seção são apresentadas 3 métricas consideradas as mais importantes segundo (Gibbons, 2010) sendo que a última é o resultado da fusão das duas.

2.2.2.1.

Coeficiente de Separação

Inicialmente, esta métrica foi utilizada para redes que realizam tarefas de classificação. O coeficiente de separação é a distância média entre os estados gerados pelo *Reservoir* após a apresentação dos vetores de entrada de cada uma das classes objetivo. Essa definição se baseia no fato que os vetores de entrada e os estados correspondentes estão divididos em classes discretas.

Uma vez que todos os estados do *Reservoir* de uma mesma classe têm a mesma saída, a distância entre estes estados deve ser pequena; correspondentemente, estados do *Reservoir* de diferentes classes devem ter uma distância grande. Se este princípio se cumpre, a rede estará em capacidade de realizar a classificação corretamente.

Tendo um conjunto de estados X , todos da mesma classe, $x_1, x_2, x_3 \dots x_k \in O_j$, a média do centro de massa da classe O_j , $u(O_j(t))$, no tempo t , é definida pela equação 2.7.

$$u(O_j(t)) = \frac{\sum_{i=1}^K x_i(t)}{K} \quad (2.7)$$

Onde, $x_i(t)$ representa o estado no tempo t , K é o número de estados totais e $u(O_j(t))$ é o centro de massa da classe O_j .

A separação do *Reservoir* é especificada em termos de dois fatores: a distância inter-classe $C_d(t)$, definida na equação 2.8, e a distância intra-classe $C_v(t)$, definida na equação 2.9.

$$C_d(t) = \sum_{m=1}^N \sum_{n=1}^N \frac{||u(O_m(t)) - u(O_n(t))||}{N^2} \quad (2.8)$$

$$C_v(t) = \frac{1}{N} \sum_{m=1}^N \frac{\sum_{i=1}^K ||u(O_m(t)) - x_i(t)||}{K} \quad (2.9)$$

Onde, $u(O_m(t))$ e $u(O_n(t))$ são os centros de massa das classes O_m e O_n .

O coeficiente de separação é, por fim, obtido utilizando a equação 2.10.

$$SEP_x(t) = \frac{C_d(t)}{C_v(t) + 1} \quad (2.10)$$

Como apresentado em (Goodman, 2006), quanto maior o coeficiente de separação, maior a precisão na classificação, uma vez que a distância inter-classe $C_d(t)$ maior que a distância intra-classe. Portanto, é possível determinar a qualidade do *Reservoir* antes do treinamento.

É possível estender o conceito de separação para sistemas que utilizam *Reservoirs* sem a necessidade de predefinir classes, isto é, pode ser utilizado em sistemas não supervisionados.

Assumindo dois vetores de entrada $u_i(t)$ e $u_j(t)$ que são significativamente diferentes, a distância entre eles pode ser denominada D , segundo a equação 2.11.

$$||u_i(t) - u_j(t)|| = D \quad (2.11)$$

Para que um *Reservoir* realize a classificação corretamente, ele deve produzir os estados $x_i(t)$ e $x_j(t)$, resultantes das entradas $u_i(t)$ e $u_j(t)$, com uma distância igual ou maior que valor D , segundo a equação 2.12.

$$||x_i(t) - x_j(t)|| \geq D \quad (2.12)$$

A razão entre a distância dos vetores de entrada e a distância dos estados gerados deve, portanto, cumprir a equação 2.13, ou seja, a razão deve ser igual ou maior que 1.

$$\frac{||x_i(t) - x_j(t)||}{||u_i(t) - u_j(t)||} \geq 1 \quad (2.13)$$

2.2.2.2.

Coeficiente de *Lyapunov*

Segundo pesquisas realizadas por (Scharauwen, 2008) e (Verstraeten, 2007), a natureza caótica do *Reservoir* pode ser mensurada empiricamente estimando o coeficiente de *Lyapunov*. Se o valor do coeficiente for positivo, o sistema tem um comportamento caótico e pequenas mudanças na entrada podem apresentar grandes mudanças nos estados do *Reservoir*. Por outro lado, um coeficiente de *Lyapunov* negativo resulta em uma rede estável, onde pequenas diferenças na entrada convergem em estados do *Reservoir* que tendem a um atrator.

Em (Rosenstein, 2012) é apresentado um método para calcular o coeficiente de *Lyapunov* para pequenos conjuntos de dados, estimando a taxa média de separação de distância entre dois vizinhos mais próximos (Luque, 2000). A distância entre estados no tempo t é definida pela equação 2.14.

$$d_j(t) = ||x_i(t) - x_j(t)|| \quad (2.14)$$

Onde, $x_i(t)$ é o vizinho mais próximo de $x_j(t)$. Supondo que a distância de estados no tempo 0 é $d_j(0)$, o coeficiente de *Lyapunov*, λ , pode ser calculado para k número de elementos a ser classificados utilizando a equação 2.15.

$$\lambda(t) = k \sum_{n=1}^K \ln \left(\frac{d_j(t)}{d_j(0)} \right) \quad (2.15)$$

Este método pode ser adaptado modificando a distância dos estados iniciais pela distância dos vetores de entrada, onde $d_j(0)$ é substituído pela equação 2.16.

$$d_j(t) = ||u_i(t) - u_j(t)|| \quad (2.16)$$

Onde, $u_i(t)$ é o vizinho mais próximo de $u_j(t)$ no vetor de entrada. O coeficiente de *Lyapunov* da equação 2.18 é modificado para a equação 2.17.

$$\lambda(t) \sim k \sum_{n=1}^N \ln \left(\frac{\|x_i(t) - x_j(t)\|}{\|u_i(t) - u_j(t)\|} \right) \quad (2.17)$$

Onde, $u_i(t)$ é o vizinho mais próximo de $u_j(t)$ e $x_i(t)$ e $x_j(t)$ são o estados do *Reservoir* resultantes dos vetores de entrada $u_i(t)$ e $u_j(t)$. Segundo (Legenstein, 2007) e (Nils, 2004), tem se observado que o desempenho ótimo do *Reservoir* acontece na “Beira do Caos” onde o coeficiente de Lyapunov esta perto de 0. Em outras palavras, o resultado da razão entre a diferença dos estados e a diferença dos vetores de entrada em um *Reservoir* com resposta ideal deve cumprir a equação 2.18.

$$\frac{\|x_i(t) - x_j(t)\|}{\|u_i(t) - u_j(t)\|} \cong 1 \quad (2.18)$$

2.2.2.3.

Separation Ratio Graph

Os conceitos de separação do *Reservoir* e o coeficiente de *Lyapunov* são unificados em um método nomeado *Separation Raio (SR)*. Por definição, o *SR* é a razão entre a distância dos vetores de entrada e a distância entre os estados gerados por essas entradas, como apresentado na equação 2.18, onde um *Reservoir* ótimo cumpre com $SR \cong 1$.

É possível apresentar graficamente a distância de separação dos vetores de entrada versus a distância de separação dos correspondentes estados do *Reservoir*. Para realizar esta apresentação gráfica utiliza-se um conjunto de vetores de entrada $\{u_1(t), u_2(t), \dots, u_p(t)\}$, e os estados gerados por estas entradas $\{x_1(t), x_2(t), \dots, x_p(t)\}$. Para cada par de estados, i e $j \in [1, p]$ onde p é o número total de padrões, a distância entre os vetores de entrada $\|u_i(t) - u_j(t)\|$, é apresentada graficamente contra a distância entre os estados correspondentes $\|x_i(t) - x_j(t)\|$.

Após realizar a apresentação gráfica é possível aplicar uma regressão linear para determinar a linha pontilhada da figura 6. A região onde a razão da separação está perto de 1 é definida como “Beira do caos” (Natschlager, 2005), sendo que a região objetivo está no meio da zona caótica e da zona de atratores, como apresentado na figura 5.

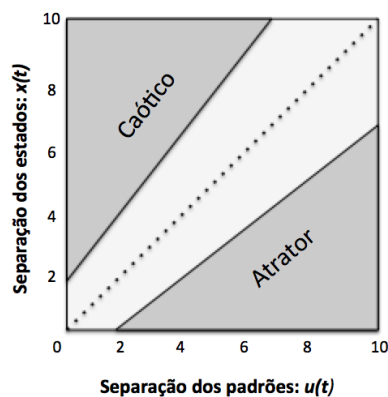


Figura 5. Região objetivo e Zonas caóticas e de atratores.

(adaptado de Gibbons 2010).

Na figura 6 são apresentadas as respostas de 3 redes com comportamentos diferentes do *Reservoir*. Estas redes foram utilizadas para classificação e o parâmetro k , que representa a quantidade de classes, foi modificado em cada uma das experiências realizadas em (Gibbons, 2010).

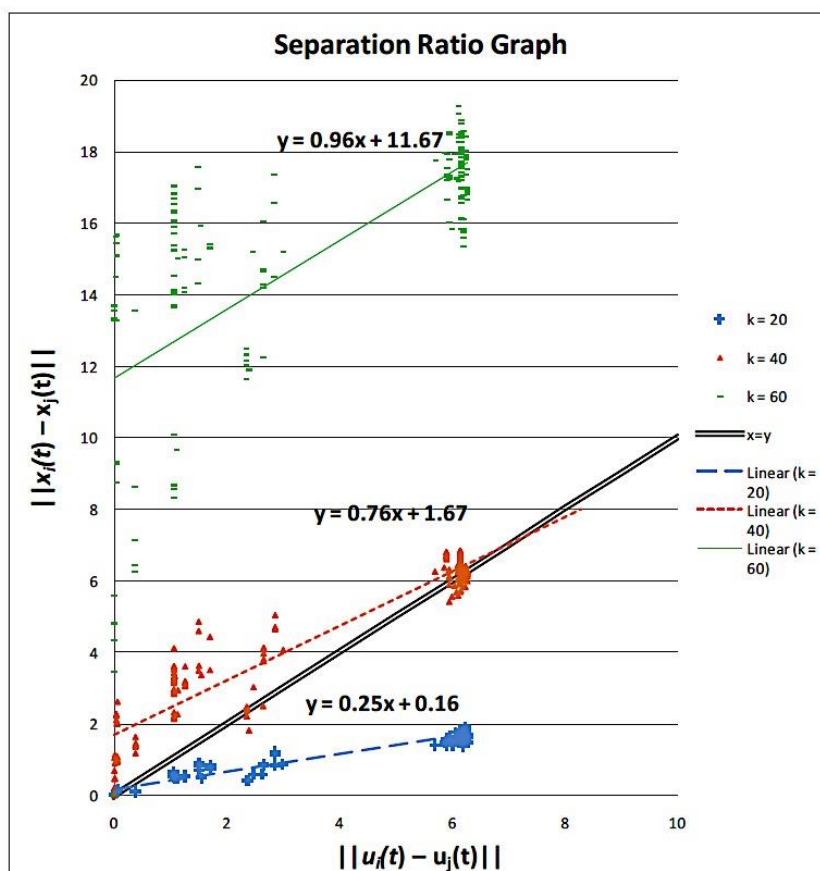


Figura 6. Regressão linear de 3 redes com dinâmicas diferentes. (Gibbons 2010).

É possível observar na figura 6 que existem comportamentos diferentes para cada tarefa de classificação: no caso da rede com $k = 20$ é possível observar que a regressão da rede tem um comportamento de um atrator (grandes diferenças na entrada produzem pequenas diferenças na saída); no caso de $k = 40$ o resultado obtido na regressão está muito perto do comportamento ideal; e no caso de $k = 60$ é possível observar se tem uma resposta caótica (pequenas mudanças na entrada geram grandes mudanças na saída). No caso dos eixos apresentados na figura 6 se tem no eixo X a diferença dos pares de entradas $u_i(t)$, $u_j(t)$, os quais geram os pares de estados $x_i(t)$, $x_j(t)$ e a diferença destes é apresentada no eixo Y, tudo isto representa a distância dos estados gerados pelo *Reservoir*.

No próximo capítulo será apresentada a análise de sensibilidade dos parâmetros das ESNs, a qual procura determinar a influência dos parâmetros gerais de projeto de ESNs no seu desempenho para a tarefa de previsão *Multi-Step* de séries temporais.

3.

Análise de Sensibilidade

Conforme descrito no capítulo anterior, as redes *ESNs* possuem uma quantidade considerável de parâmetros (10) a serem especificados de forma a se obter um bom desempenho do modelo. Com o objetivo de se conhecer a influência de cada um desses parâmetros no desempenho das redes *ESNs* e reduzir o espaço de busca no desenvolvimento de um modelo híbrido para a otimização automática dos parâmetros mais importantes, este capítulo apresenta uma análise de sensibilidade de cada um desses parâmetros na previsão de séries temporais.

Assim, no intuito de determinar como estes parâmetros interferem no desempenho das *ESNs*, este capítulo apresenta o resultado de análises envolvendo os parâmetros identificados no trabalho “*Reservoir Riddles: Suggestions for Echo State Network Research*” (Jaeger, 2005). Dentre os diversos parâmetros da *ESNs*, Jaeger indica 5 como os mais sensíveis para uma boa configuração das *ESN*, a saber:

- raio espectral - $\rho(w)$,
- fator de escala nos pesos das conexões de entrada - sW^e ,
- fator de escala nos pesos das conexões recorrentes - sW^{rr} ,
- número de unidades no *Reservoir* - $\#U$
- percentual de conexão no *Reservoir* - $\%C$.

A figura 7 apresenta a forma como os testes foram realizados para cada um dos parâmetros selecionados, destacando o tamanho da janela (*Jan*), pois este parâmetro é utilizado para determinar a quantidade de atributos utilizados. Inicialmente é realizada a otimização dos parâmetros gerais utilizando um algoritmo genético, resultando em uma rede *ESN* otimizada que será nomeada **Modelo Base**. Estes testes consideraram dois cenários distintos: no primeiro são geradas 100 redes com conexões diferentes e pesos diferentes; já no segundo cenário, são geradas 100 redes com as mesmas conexões mas com pesos diferentes, com o intuito de determinar a influência na modificação dos pesos durante o processo de geração do *Reservoir*.

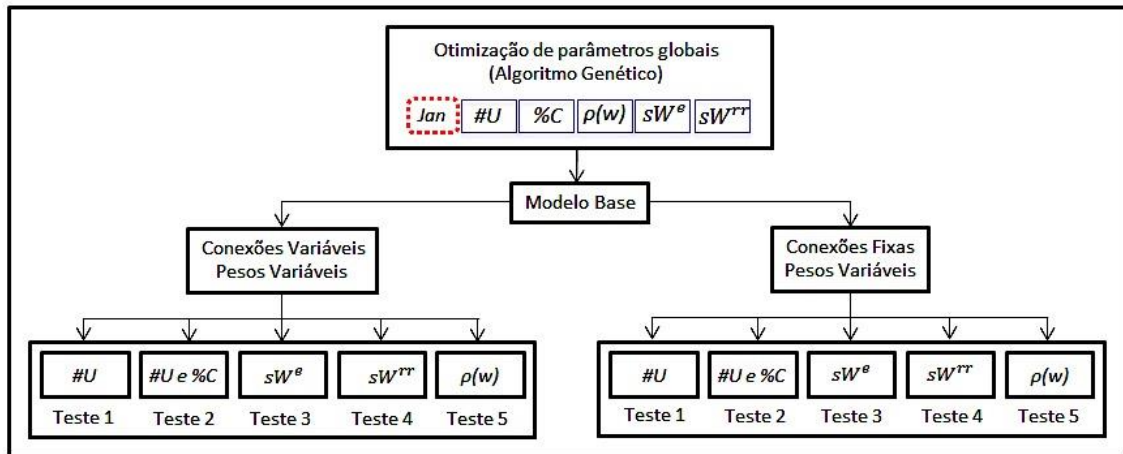


Figura 7. Estrutura dos testes realizados.

A tabela 1, a seguir, apresenta o número de configurações testadas para cada um dos parâmetros avaliados, tanto para as redes com conexões e pesos variáveis quanto para as redes com conexões fixas e pesos variáveis. Os diferentes valores utilizados na análise de sensibilidade de cada parâmetro são apresentados nas seções 3.2.1 a 3.2.5.

Tabela 1. Quantitativo de redes testadas por série.

Parâmetro Avaliado	Configurações (Quantidade de valores diferentes)	Número de Redes testadas para Conexões Variáveis e Pesos Variáveis	Número de Redes testadas para Conexões Fixas e Pesos Variáveis	Total de redes avaliadas
(#U)	10	100	100	2000
(#U) e (%C)	50 (10 #U x 5 %C)	100	100	10000
sW^e	11	100	100	2200
sW^{rr}	11	100	100	2200
$\rho(w)$	10	100	100	2000
Total de testes	18400			

Na equação 3.1 é apresentada a métrica utilizada como função aptidão do algoritmo genético. O SMAPE (erro médio simétrico absoluto percentual) indica o percentual de erro que o modelo comete nas previsões. O uso do SMAPE visa a minimizar alguns erros que o tradicional MAPE (erro médio absoluto percentual) comete quando o valor da previsão é maior ou menor que o observado, ou quando os valores das previsões são próximos de zero (MAKRIDAKIS e HIBON, 2000). Neste caso particular, a função aptidão (SMAPE) é calculada no conjunto de validação.

$$SMAPE_{Validação} (\%) = \frac{1}{N} \sum_{t=1}^N \frac{|y_t - y'_t|}{(y_t + y'_t)/2} * 100 \quad (3.1)$$

Os resultados dos testes realizados são apresentados através de três métricas: Erro Médio Quadrático (sigla *RMSE* em inglês) da melhor rede, dentre as 100 avaliadas; média do

$RMSE$ e desvio padrão das 100 redes. As métricas são apresentadas nas equações 3.2, 3.3 e 3.4, aplicadas aos conjuntos de validação e teste.

$$RMSE = \sqrt{\frac{\sum_{t=1}^N (y_t - y'_t)^2}{N}} \quad (3.2)$$

$$\overline{RMSE} = \frac{1}{N} \sum_{i=1}^N RMSE_i \quad (3.3)$$

$$DP = \sqrt{\frac{\sum_{t=1}^N (RMSE_t - \overline{RMSE})^2}{N - 1}} \quad (3.4)$$

Onde, y_t é o valor real a ser previsto, y'_t é o valor previsto pelo modelo e N é o número de amostras avaliadas.

3.1.

Apresentação do Benchmark

Para a análise de sensibilidade, foi selecionada a série temporal *Mackey-Glass*, proposta pelos professores da McGill University do Canada, Leon Glass e Michael Mackey (Mackey, 1977), em pesquisas conjuntas que procuravam determinar se a dinâmica caótica poderia existir em sistemas fisiológicos. Mackey trabalhava na modelagem do controle de produção de células do sangue e Glass estudava o controle respiratório, sendo que os dois modelos compartilhavam a necessidade de utilizar equações diferenciais devido à presença de comportamentos próprios da existência de dependência de atrasos de tempo. Uma das vantagens na utilização desta série é a possibilidade de ajustá-la para comportamento mais ou menos caótico, de acordo com o parâmetro τ (ver equação 3.5).

Conforme apresentado em (Ferreira, 2011), valores acima de 16,8 no parâmetro τ permite a existência de um comportamento caótico na série.

A série temporal *Mackey-Glass* é obtida através da equação diferencial 3.5, a qual pode ser resolvida numericamente empregando o método de quarta ordem de Runge-Kutta (Liu, 1999).

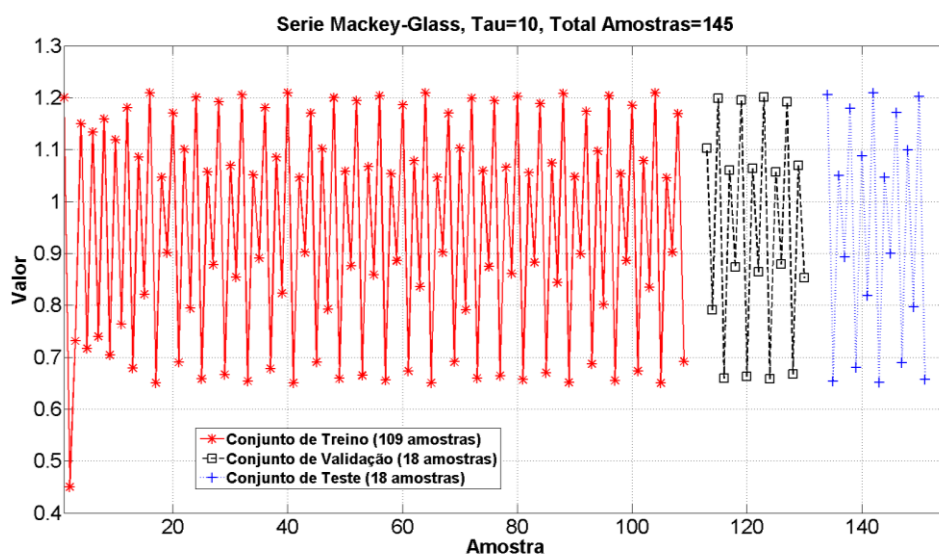
$$\frac{dx(t)}{dt} = \frac{ax(t - \tau)}{1 + x(t - \tau)^{10}} - bx(t) \quad (3.5)$$

onde, a e b são constantes, e τ é a constante de atraso. Outros parâmetros utilizados na solução são: a condição inicial $x(0)$; o tamanho do passo de tempo Δt ; o número de amostras N ; e o tamanho do intervalo ξ . Na tabela 2 são apresentados os parâmetros utilizados para gerar as três configurações da série *Mackey-Glass* utilizadas como benchmark (três valores distintos de τ), as quais foram obtidas do código existente na biblioteca do Matlab.

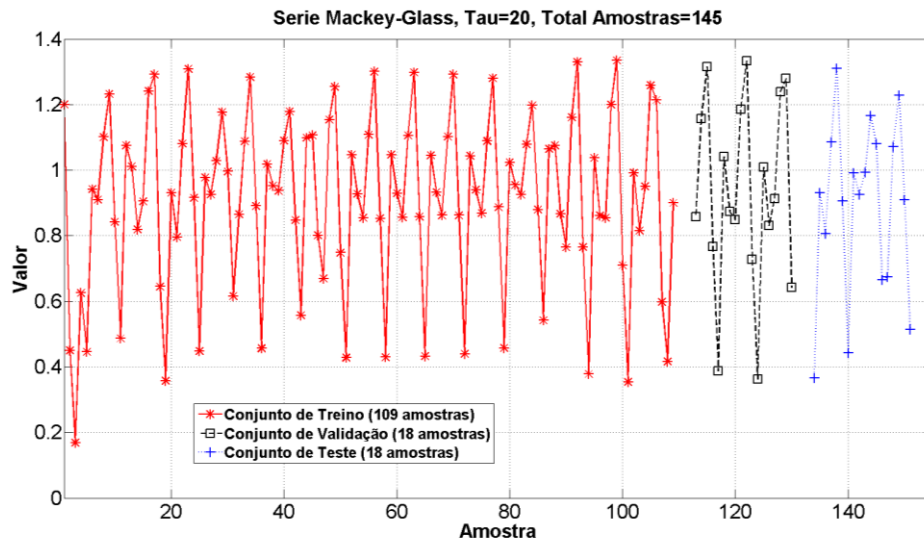
Tabela 2. Parâmetros de entrada para gerar a série *Mackey-Glass* (biblioteca do Matlab versão 2013a).

Parâmetro	Valor
a	0.2
b	0.1
τ	10/20/70
x_0	1.2
Δt	10
N	145
ξ	1

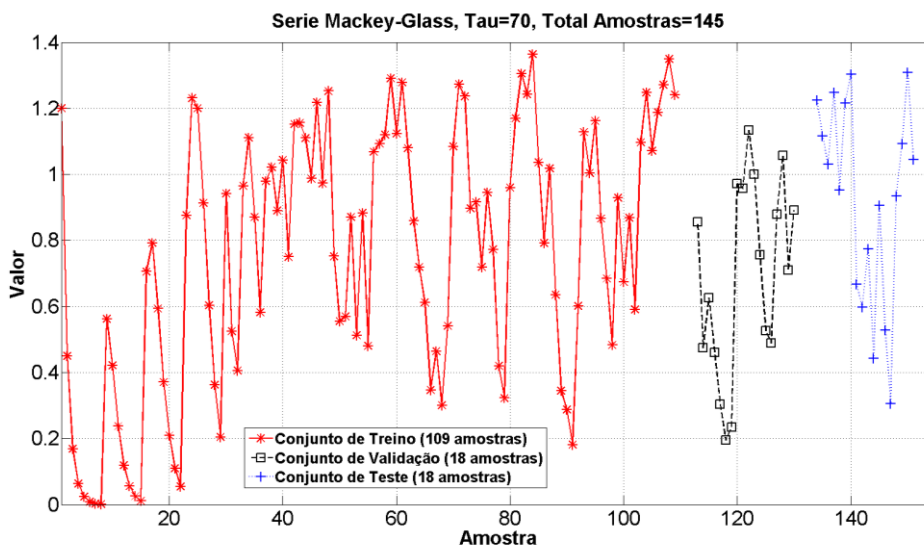
Na figura 8 são apresentadas as 3 séries completas, as quais contêm 145 amostras. O número de amostras é similar ao número de amostras da versão reduzida das séries *NN3* (que será utilizada nos testes do modelo proposto), ou seja, as 11 séries da versão reduzida apresentam amostras de 133 a 144 valores em cada série. Para conservar uma similaridade maior com a competição *NN3*, o tamanho dos conjuntos de validação e teste nesta análise de sensibilidade é composto pelos últimos 36 valores, sendo que os 18 penúltimos compõem o conjunto de validação e os 18 últimos compõem o conjunto de teste. A previsão é realizada de modo *Multi-Step*, onde a previsão futura utiliza previsões realizadas nos passos anteriores como dado de entrada da rede.



a)



b)



c)

Figura 8. Séries *Mackey-Glass* utilizadas, a) $\tau = 10$, b) $\tau = 20$, c) $\tau = 70$.

De forma a melhor analisar as características das 3 séries, utilizou-se o Coeficiente de Assimetria (equação 3.6), o qual calcula a variabilidade assimétrica dessas séries, onde valores próximos de 0 indicam uma distribuição de frequências simétrica, o que pode ser interpretado como uma série mais fácil de ser prevista. O coeficiente de assimetria pode ser obtido utilizando a equação 3.6.

$$\gamma = \frac{E[(X - \bar{X})^3]}{\sigma^3} \quad (3.6)$$

Onde $E[(X - \bar{X})^3]$ é o terceiro momento central de uma variável X aleatória e \bar{X} é o valor esperado de X , também conhecido como μ_3 e σ é o desvio padrão. A tabela 3 contém os

coeficiente de assimetria para cada uma das 3 séries utilizadas. Verifica-se, como esperado, que a série com $\tau = 10$ apresenta um coeficiente próximo de 0 (série pouco caótica); já as séries com $\tau = 20$ e $\tau = 70$ apresentam comportamentos mais complexos, com um coeficiente de assimetria médio e próximo de -1.

Tabela 3. Valores do coeficiente de assimetria para as 3 séries utilizadas.

Série	γ
Mackey-Glass ($\tau = 10$)	-0.1976
Mackey-Glass ($\tau = 20$)	-0.5505
Mackey-Glass ($\tau = 70$)	-0.9159

Conforme mencionado anteriormente, a análise de sensibilidade foi realizada a partir de um Modelo Base da ESN. A tabela 4 apresenta os limites utilizados na configuração do cromossomo do algoritmo genético utilizado para configurar o Modelo Base da ESN, assim como a resolução utilizada para cada parâmetro.

Tabela 4. Limites dos parâmetros do Cromossomo.

Parâmetro	Símbolo	Valor Mínimo	Valor Máximo	Resolução
Janela	Jan	1	36	1
Número de Unidades	$\#U$	100	2000	1
Percentual de Conexão	$\%C$	20	100	1
Raio Espectral	$\rho(W)$	0.1	1	0.1
Fator de Escala nos pesos das Conexões de entrada	sW^e	0	1	0.1
Fator de Escala nos pesos das Conexões recorrentes	sW^{rr}	0	1	0.1

Na tabela 5 são apresentados os parâmetros de configuração do GA e os tempos de convergência para cada uma das séries. Todas as simulações foram realizadas em Matlab utilizando um computador Core i5 2,3GHz com 4GB de RAM.

Tabela 5. Parâmetros do GA para as 3 séries utilizadas.

Parâmetro	$\tau = 10$	$\tau = 20$	$\tau = 70$
Tamanho da população	50	50	50
Critério de parada	100	100	100
Operador de Cruzamento	1 Ponto	1 Ponto	1 Ponto
Operador de Mutação	Uniforme	Uniforme	Uniforme
Taxa de Cruzamento	0,7	0,7	0,7
Taxa de Mutação	0,2	0,2	0,2
Tempo de Convergência	14,21m	14,65m	17,43m

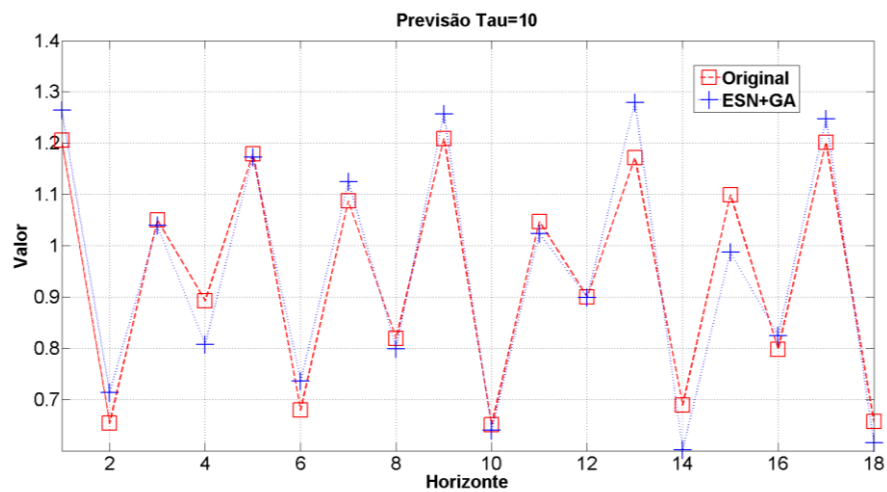
A tabela 6 apresenta os parâmetros otimizados pelo GA para cada uma das ESNs consideradas Modelos Base. Conforme apresentado na tabela 6, as funções de ativação dos neurônios do *Reservoir* e do neurônio de saída foram fixadas em tangente hiperbólica e linear, respectivamente. A análise de sensibilidade envolveu a modificação de apenas

um parâmetro por vez, sendo que os outros parâmetros restantes foram fixados nos valores especificados pelo algoritmo genético e apresentados na tabela 6.

Tabela 6. Parâmetros otimizados pelo GA para as *ESNs* base das 3 séries utilizadas.

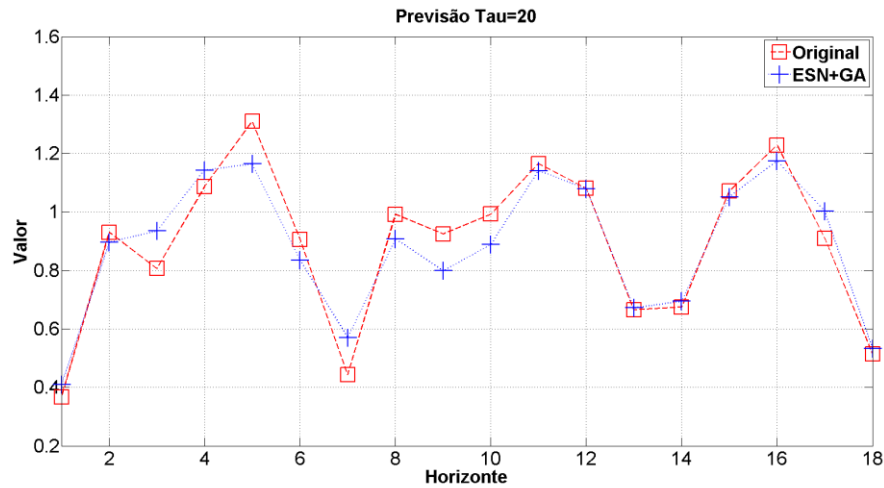
Parâmetro	$\tau = 10$	$\tau = 20$	$\tau = 70$
Número de unidades no <i>Reservoir</i>	538	750	1308
Janela	11	17	18
Porcentagem de conexão	24%	39%	61%
Raio Espectral	0.2	0.9	0.8
Fator de escala das conexões de entrada	0.4	0.6	0.8
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"	"Tanh"	"Tanh"
Função de ativação do neurônio de Saída	"Linear"	"Linear"	"Linear"
Fator de escala das conexões recorrentes	0.5	0.3	0.1

A previsão obtida pelo Modelo Base de *ESN* para os conjuntos de teste é apresentada na figura 9, assim como o valor de *RMSE*.



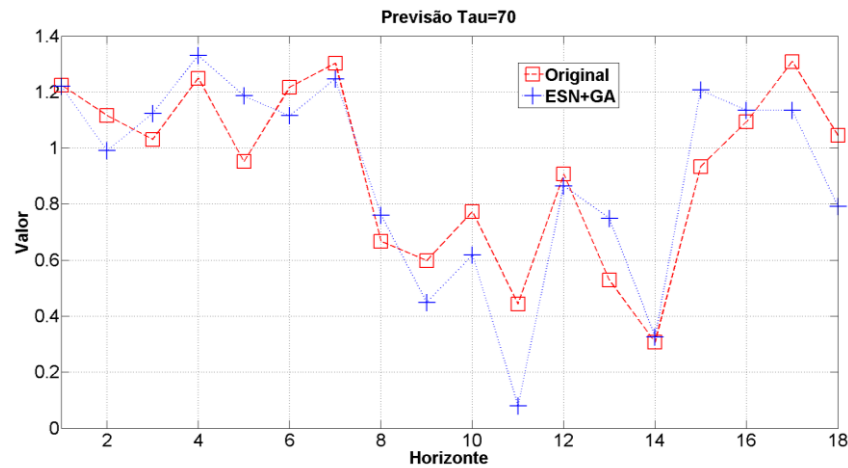
$RMSE_Treino=0,0216; RMSE_Validação=0,0512; RMSE_Teste=0,0571$

a)



$RMSE_Treino=0,0345; RMSE_Validação=0,0671; RMSE_Teste=0,0789$

b)



$RMSE_Treino=0,0872; RMSE_Validação=0,1159; RMSE_Teste=0,1681$

c)

Figura 9. Conjunto de validação vs previsão ESN, a) $\tau = 10$, b) $\tau = 20$, c) $\tau = 70$.

3.2.

Parâmetros Seleccionados

No texto “*Reservoir Riddles: Suggestions for Echo State Network Research*”, são feitos questionamentos tais como: “é possível caracterizar um determinado *Reservoir* para uma determinada tarefa?”, “Como pode ser adaptado o *Reservoir* para uma determinada tarefa?”, “Qual é o papel da organização topológica do *Reservoir*?”, entre outros. Como resposta a alguns destes questionamentos, (Jaeger, 2005) apresenta que a otimização de parâmetros globais é uma solução aceitável, mas ainda devem ser realizados análises complementares no que tange às ativações do *Reservoir*. Assim, o que se pretende com este estudo de sensibilidade dos parâmetros é obter uma compreensão teórica mais aprofundada que permita realizar ajustes no projeto das ESNs segundo a tarefa proposta.

As seções a seguir avaliam a influência dos 5 principais parâmetros de uma *ESN*: o tamanho do *Reservoir*; o Percentual de Conexões; o Fator de escala nos pesos das conexões de entrada; o Fator de Escala nas conexões recorrentes; e o Raio Espectral.

3.2.1.

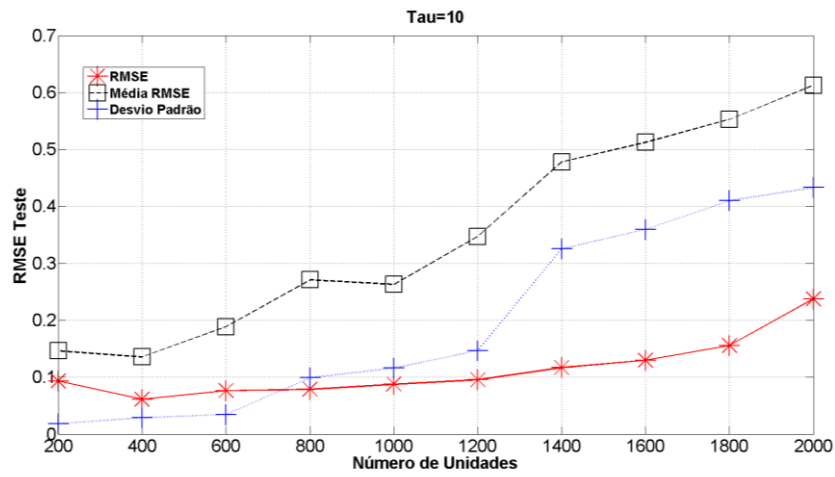
Tamanho do *Reservoir*

Nesta experiência foi alterado o número de unidades no *Reservoir*. Os valores utilizados foram selecionados segundo os seguintes critérios: o menor número de unidades utilizado nos testes é 200, onde este número mantém a propriedade de ser grande o suficiente para levar a uma dimensão maior os padrões apresentados, e o maior número de unidades foi determinado mediante uma análise prévia, onde observou-se que *Reservoirs* maiores que 2000 unidades não apresentavam uma melhora considerável no desempenho obtido.

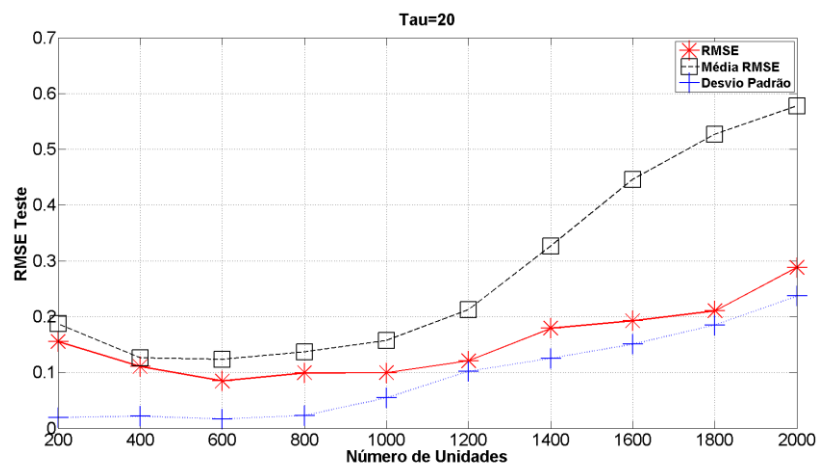
A seguir são apresentados, nas figuras 10 e 11, e nas tabelas 7 e 8, os resultados dos testes realizados para conexões e pesos variáveis, para conexões fixas e pesos variáveis, respectivamente. Nas ditas tabelas o valor do RMSE corresponde à melhor rede obtida; no caso da média são considerados todos os valores dos 100 experimentos.

Tabela 7. Resultados com conexões e pesos variáveis. (#U)

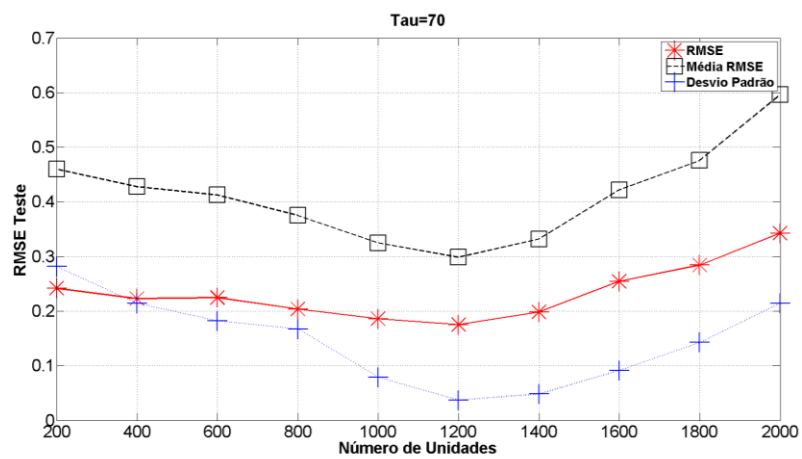
Tamanho do <i>Reservoir</i>	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,0934	0,0612	0,07641	0,0788	0,0877	0,0954	0,1168	0,1296	0,1556	0,2376
Média Teste	0,1465	0,1358	0,1884	0,2711	0,2628	0,3472	0,4781	0,5124	0,5526	0,6124
Desvio Padrão Teste	0,0183	0,0289	0,0346	0,0992	0,116	0,1466	0,3253	0,3594	0,4098	0,4329
$\tau = 20$										
RMSE Teste	0,1553	0,1105	0,08463	0,0992	0,0995	0,1207	0,1793	0,1926	0,2105	0,2882
Média Teste	0,1875	0,1264	0,1235	0,137	0,1576	0,2127	0,3268	0,4461	0,5269	0,5777
Desvio Padrão Teste	0,0195	0,0218	0,0167	0,023	0,0546	0,1021	0,1255	0,1509	0,1848	0,2373
$\tau = 70$										
RMSE Teste	0,2417	0,22289	0,2247	0,2038	0,1857	0,1754	0,1985	0,2547	0,2843	0,3425
Média Teste	0,4597	0,4279	0,4127	0,3752	0,3247	0,2987	0,3321	0,4217	0,4759	0,5961
Desvio Padrão Teste	0,2819	0,2143	0,1824	0,1672	0,0791	0,0371	0,0483	0,0914	0,1428	0,2143



a)



b)



c)

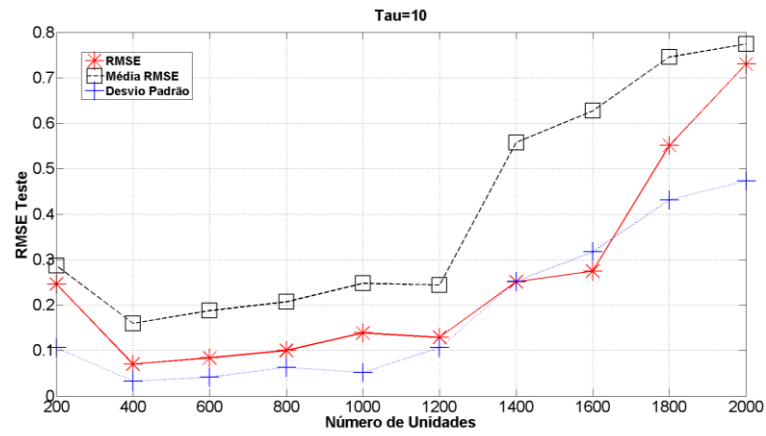
Figura 10. Séries Mackey-Glass com conexões e pesos variáveis. ($\#U$)

a) $\tau = 10$, b) $\tau = 20$, c) $\tau = 70$.

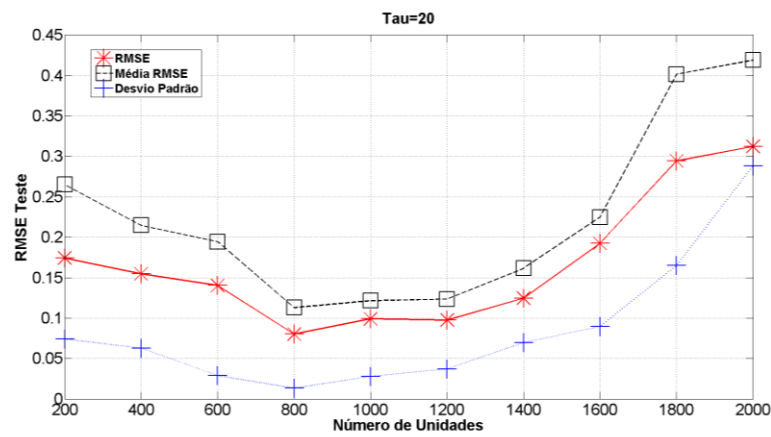
Os testes realizados a partir da modificação do tamanho de *Reservoir* permitiram identificar os melhores comportamentos das redes quando o dito parâmetro é modificado. Os melhores resultados foram obtidos nos valores próximos ao valor das configurações base (538, 750 e 1308 unidades para $\tau = 10$, $\tau = 20$ e $\tau = 70$, respectivamente), como apresentado na figura 10. No caso de configurações com tamanho de unidades maiores (1600 para frente) os resultados apresentaram previsões com erros maiores, o que pode ser entendido como *Reservoir* superdimensionados, onde não é possível conseguir uma configuração de pesos na camada de saída adequada durante o treinamento. , quer dizer os *Echoes* ou estados gerados pelo *Reservoir* em conjunto com os pesos da camada de saída em conjunto não conseguem mapear os padrões apresentados.

Tabela 8. Resultados com conexões fixas e pesos variáveis. (#U)

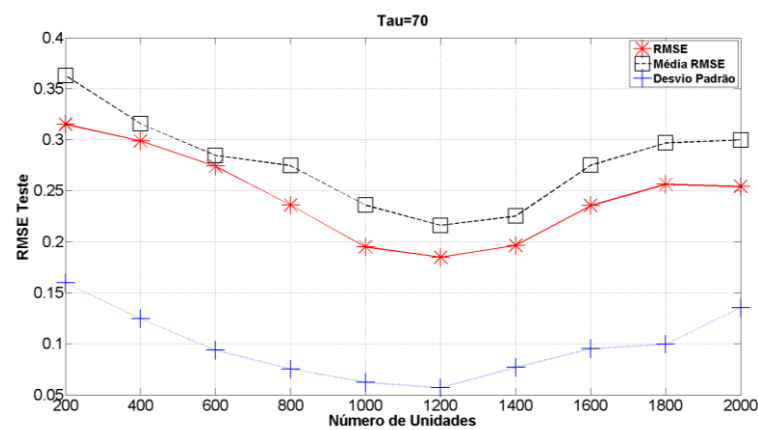
Tamanho do Reservoir	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,2462	0,0701	0,0837	0,1001	0,1386	0,1287	0,2511	0,2744	0,5509	0,7299
Média Teste	0,2864	0,1596	0,1875	0,2069	0,2478	0,2439	0,5576	0,627	0,7455	0,7742
Desvio Padrão Teste	0,1061	0,0323	0,0412	0,0626	0,0519	0,106	0,2524	0,3169	0,4314	0,4728
$\tau = 20$										
RMSE Teste	0,1743	0,1547	0,1403	0,0802	0,0991	0,0976	0,1247	0,1924	0,2943	0,3124
Média Teste	0,2651	0,2147	0,1945	0,1128	0,1214	0,1234	0,1617	0,2249	0,4018	0,4193
Desvio Padrão Teste	0,0741	0,0629	0,0289	0,0134	0,028	0,0371	0,0697	0,0897	0,165	0,2884
$\tau = 70$										
RMSE Teste	0,3151	0,2989	0,2741	0,2359	0,1949	0,1847	0,1965	0,2354	0,2563	0,2541
Média Teste	0,3627	0,3156	0,2846	0,2746	0,2359	0,2159	0,2251	0,2749	0,2967	0,2997
Desvio Padrão Teste	0,1596	0,1247	0,0937	0,0752	0,0621	0,0568	0,0769	0,0951	0,09947	0,1354



a)



b)



c)

Figura 11. Série Mackey-Glass com conexões fixas e pesos variáveis. ($\#U$)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

No caso da configuração de conexões fixas e pesos variáveis, segundo a figura 11, os resultados apresentaram um comportamento similar ao observado na primeira parte do teste. No entanto, neste caso o desempenho obtido para as redes com número de unidades 200 foi pior.

Nos dois casos analisados, percebe-se que o número de unidades necessárias está diretamente relacionado à complexidade da tarefa de previsão, isto é, conforme o coeficiente de assimetria aumenta, maior é o número de unidades necessárias.

3.2.2.

Percentual de Conexões e Tamanho do *Reservoir*

Na experiência a seguir são avaliadas configurações nas quais tanto o percentual de conexões quanto o tamanho do *Reservoir* são modificados. Segundo Jaeger (2001), o menor percentual de conexões é da ordem de 20%, que é considerado o maior percentual de esparsidade na matriz de conexões. Sendo assim, o parâmetro percentual de conexões foi avaliado para configurações com valores percentuais de 20%, 35%, 50%, 75% e 100%, onde o máximo valor considera que todos os neurônios do *Reservoir* estão conectados entre si. Neste caso, foram avaliados conjuntamente o percentual de conexões e o tamanho do *Reservoir* com o intuito de determinar a relação entre estes dois parâmetros. Deste modo, os valores utilizados para o percentual de conexão (20%, 35%, 50%, 75% e 100%) foram avaliados com as mesmas configurações de número de unidades do *Reservoir* da experiência anterior, isto é, alguns valores entre 200 e 2000 unidades.

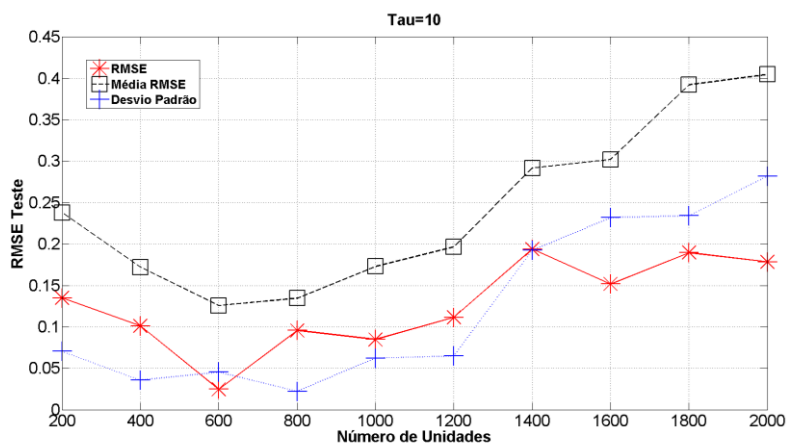
São avaliados os desempenhos das redes em conjunto com a variação do número de unidades. Nas figuras 12 a 21 e nas tabelas 9 a 18 são apresentados os resultados para as configurações descritas no parágrafo anterior, novamente para testes com conexões e pesos variáveis e com conexões fixas e pesos variáveis.

3.2.2.1 Percentual de conexões 20%

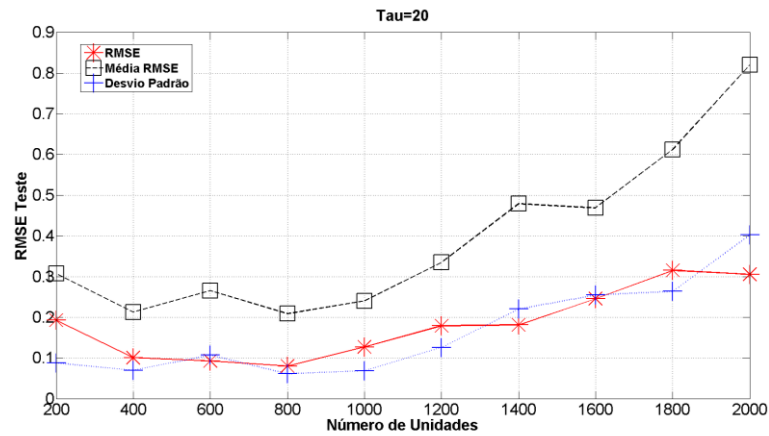
A seguir são apresentados os resultados obtidos para a duas possíveis configurações de conexões e percentual de conexões de 20%.

Tabela 9. Resultados com conexões e pesos variáveis (%C = 20)

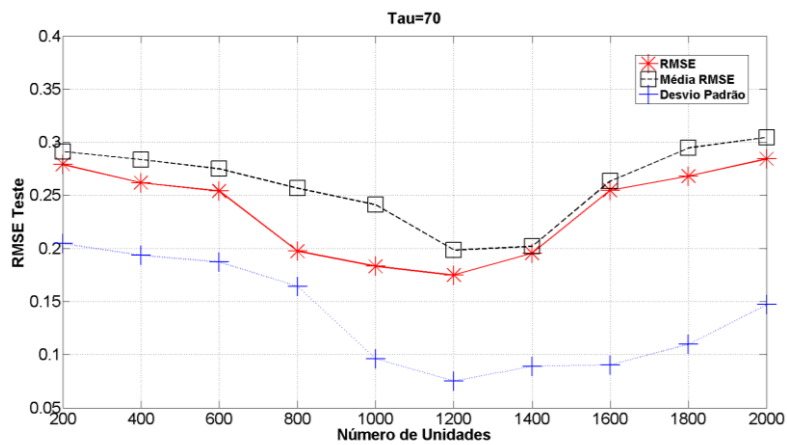
20% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,1351	0,1013	0,0247	0,0957	0,0849	0,1115	0,194	0,1518	0,1897	0,1786
Média Teste	0,2381	0,1721	0,1257	0,1347	0,1731	0,1966	0,2917	0,3022	0,3923	0,4049
Desvio Padrão Teste	0,0706	0,0357	0,0455	0,0222	0,0623	0,0648	0,1924	0,232	0,2342	0,2822
$\tau = 20$										
RMSE Teste	0,1927	0,1002	0,0923	0,0798	0,1268	0,1788	0,1811	0,2456	0,3151	0,3054
Média Teste	0,3073	0,2126	0,2653	0,2089	0,24	0,3348	0,4789	0,4688	0,6124	0,8203
Desvio Padrão Teste	0,0876	0,0693	0,1066	0,0612	0,0685	0,1255	0,2202	0,2545	0,2635	0,4023
$\tau = 70$										
RMSE Teste	0,2791	0,2619	0,2541	0,1975	0,1832	0,1749	0,1954	0,2547	0,2681	0,2843
Média Teste	0,2913	0,2836	0,2749	0,2567	0,2411	0,1985	0,2019	0,2634	0,2946	0,3045
Desvio Padrão Teste	0,2046	0,1934	0,1873	0,1643	0,0961	0,0753	0,0891	0,0904	0,1097	0,1472



a)



b)



c)

Figura 12. Série Mackey-Glass com Conexões e pesos variáveis. (%C = 20)

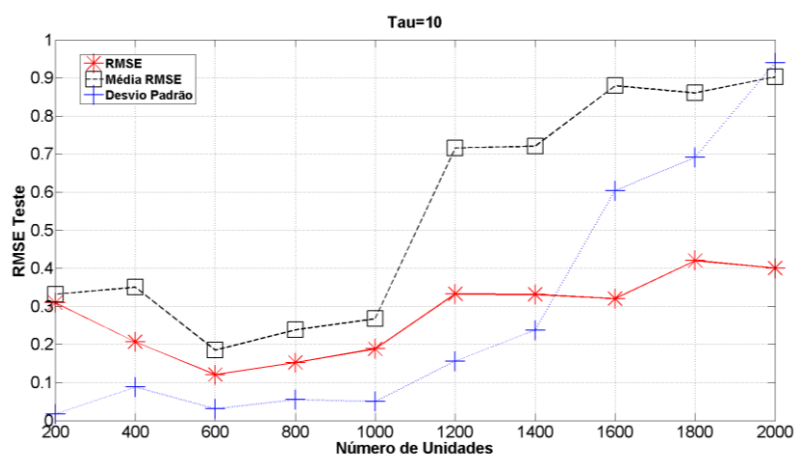
a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

No teste de variação do percentual de conexões foram observados comportamentos similares ao teste de avaliação de número de unidades entre os dois tipos de configurações de conexões e pesos utilizados.

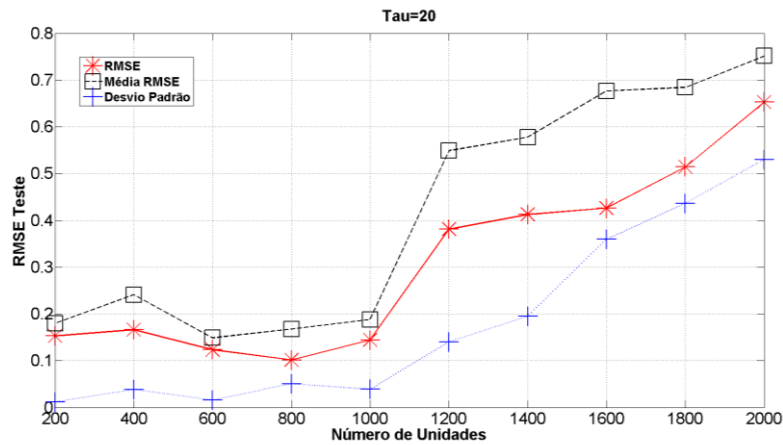
Na configuração base o percentual de conexões utilizado foi de 24%, 39% e 61% para $\tau = 10$, $\tau = 20$ e $\tau = 70$, respectivamente, e realizando a comparação direta com o *Reservoir* considerado pouco conectado (20%), os resultados obtidos para configurações de 1600 unidades em diante foram considerados insatisfatórios, chegando a obter respostas de 0,31 no *RMSE*, o que supõe previsões que estão longe de atingir os resultados reais. Pode-se constatar que realmente este tipo de configuração (quantidade de unidades muito grandes e poucas conexões) não aproveita a alta dimensionalidade oferecida pelo *Reservoir*.

Tabela 10. Resultados com conexões fixas e pesos variáveis. (%C = 20)

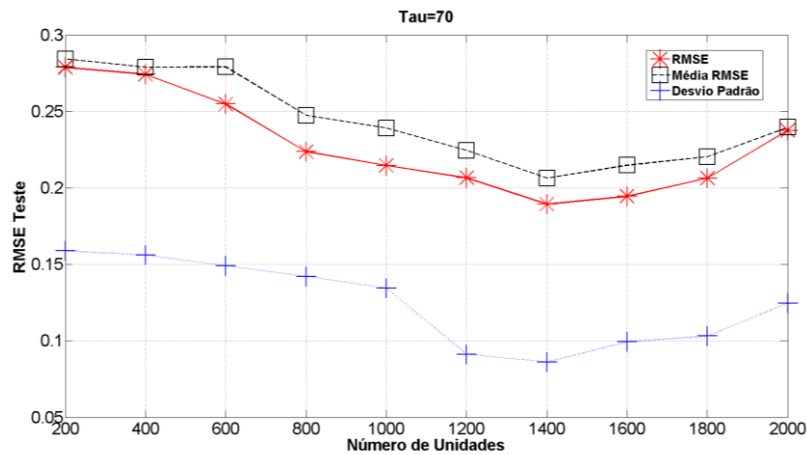
20% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0.3094	0.2067	0.1205	0.1525	0.1884	0.3325	0.3313	0.3206	0.4201	0.4003
Média Teste	0.3316	0.3503	0.1848	0.2380	0.2675	0.7159	0.7206	0.8793	0.8606	0.9032
Desvio Padrão Teste	0.0178	0.0881	0.0311	0.0551	0.0498	0.1564	0.2374	0.6041	0.6913	0.9397
$\tau = 20$										
RMSE Teste	0.1528	0.1661	0.1233	0.1016	0.1441	0.3810	0.4127	0.4261	0.5141	0.6525
Média Teste	0.1796	0.2411	0.1489	0.1677	0.1880	0.5489	0.5779	0.6764	0.6844	0.7515
Desvio Padrão Teste	0.0119	0.0383	0.0158	0.0504	0.0387	0.1401	0.1953	0.3597	0.4362	0.5300
$\tau = 70$										
RMSE Teste	0,2788	0,2743	0,2549	0,2237	0,2146	0,2066	0,1894	0,1945	0,2064	0,2375
Média Teste	0,2843	0,2789	0,2791	0,2475	0,2391	0,2245	0,2064	0,2149	0,2204	0,2397
Desvio Padrão Teste	0,1589	0,1562	0,1492	0,1421	0,1345	0,0912	0,0862	0,0994	0,1032	0,1247



a)



b)



c)

Figura 13. Série Mackey-Glass com conexões fixas e pesos variáveis. (%C = 20)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

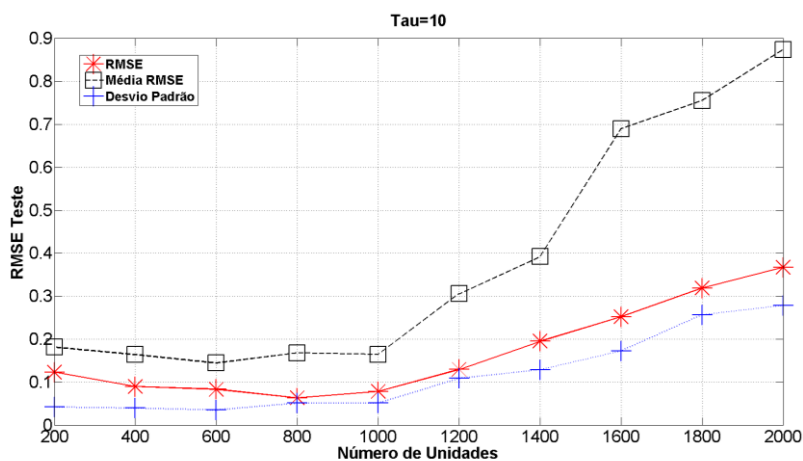
Para o caso de conexões fixas e pesos variáveis foi possível observar que nas séries $\tau = 10$ e $\tau = 20$, os erros, em média, atingiram valores altos, ou seja, para *Reservoirs* com 1200 unidades e maiores, a resposta na previsão ficou muito longe do valor esperado. Uma explicação para este tipo de comportamento pode ser a pouca quantidade de conexões para uma rede com um número de neurônios significativo, onde a condição de alta dimensionalidade não foi satisfeita.

3.2.2.2 Percentual de conexões 35%

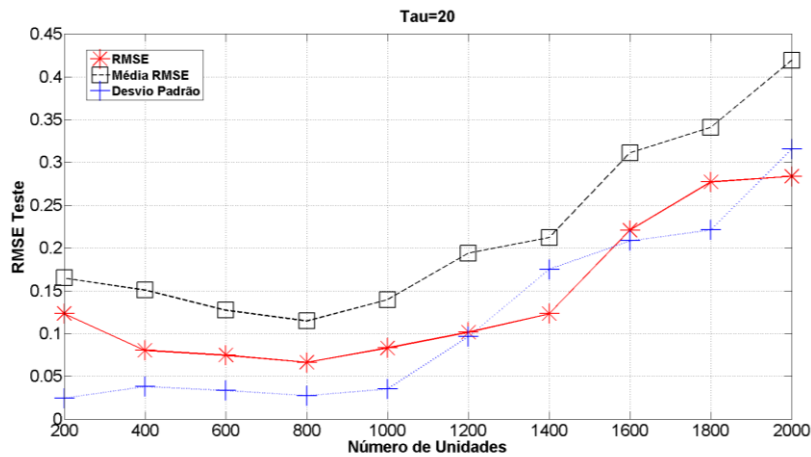
A seguir são apresentados os resultados obtidos para as duas possíveis configurações de conexões e percentual de conexões de 35%.

Tabela 11. Resultados com conexões e pesos variáveis. (%C = 35)

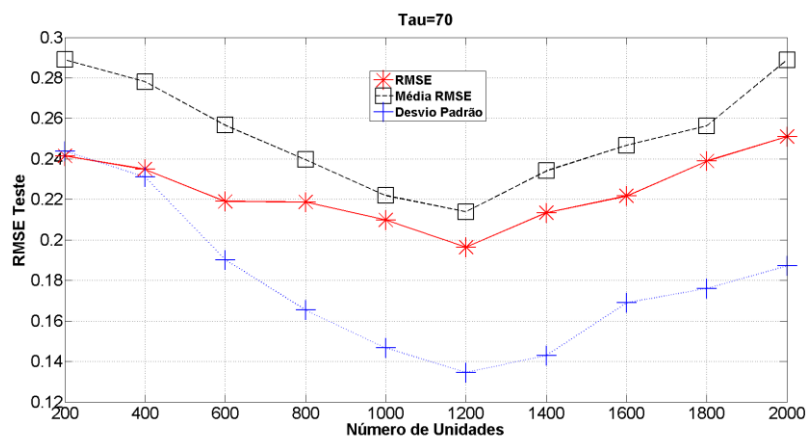
35% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,1233	0,0897	0,0831	0,0633	0,0783	0,1293	0,1952	0,2518	0,319	0,3675
Média Teste	0,1813	0,1641	0,1441	0,168	0,1647	0,3059	0,3919	0,6901	0,7555	0,8741
Desvio Padrão Teste	0,0417	0,0397	0,0349	0,0508	0,0511	0,1084	0,1286	0,1726	0,2565	0,2784
$\tau = 20$										
RMSE Teste	0,1233	0,0802	0,0747	0,0663	0,0831	0,1015	0,1231	0,2209	0,2772	0,2837
Média Teste	0,1649	0,1507	0,1273	0,1147	0,1395	0,1939	0,2122	0,3114	0,3411	0,4195
Desvio Padrão Teste	0,0244	0,0379	0,0333	0,0272	0,0351	0,0963	0,175	0,2084	0,2212	0,3159
$\tau = 70$										
RMSE Teste	0,2415	0,2349	0,219	0,2187	0,2098	0,1964	0,2134	0,2217	0,2389	0,2511
Média Teste	0,289	0,2781	0,2567	0,2397	0,2219	0,2139	0,2341	0,2467	0,2563	0,2889
Desvio Padrão Teste	0,2439	0,231	0,1902	0,1654	0,1467	0,1345	0,1429	0,1689	0,176	0,1872



a)



b)



c)

Figura 14. Série Mackey-Glass com conexões e pesos variáveis. (%C = 35)

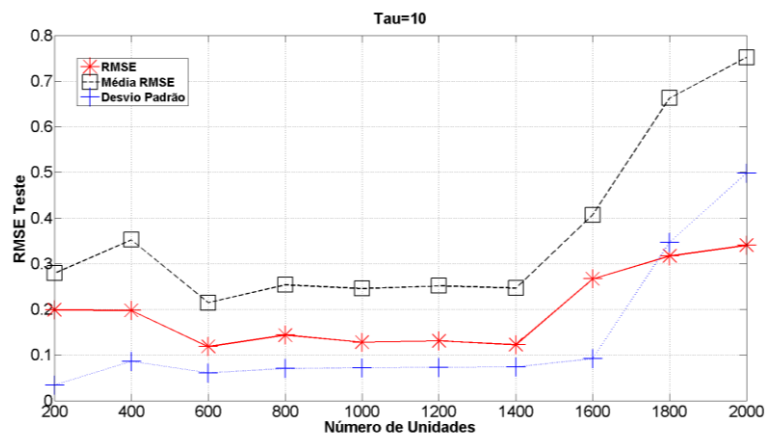
a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

No caso de percentual de conexões igual a 35%, é possível observar que para a série $\tau = 10$ o comportamento do erro médio teve um comportamento ruim, pois este percentual de conexões está mais longe do percentual do modelo base (24%), o que pode ser comprovado observando a figura 14.a no resultado da média *RMSE*.

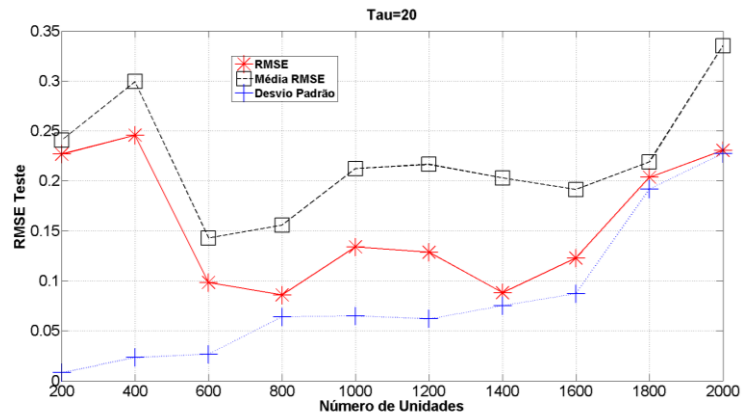
Nos resultados para $\tau = 20$ e $\tau = 70$, as respostas obtidas apresentaram o comportamento esperado, sem mudanças particulares no desempenho e obtendo os melhores resultados quando o número de unidades esteve perto do valor do modelo base (750 para $\tau = 20$ e 1300 para $\tau = 70$).

Tabela 12. Resultados com conexões fixas e pesos variáveis. (%C = 35)

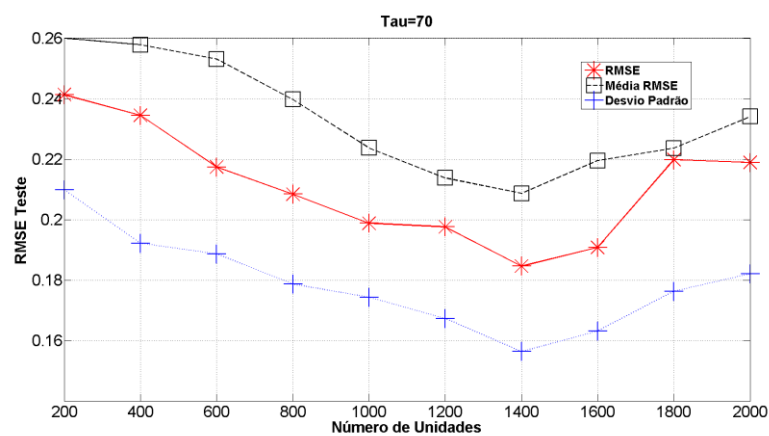
35% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,1988	0,1977	0,1185	0,1437	0,1282	0,1309	0,1225	0,2667	0,317	0,3403
Média Teste	0,2791	0,3523	0,2147	0,2538	0,2455	0,2514	0,2465	0,407	0,6627	0,7521
Desvio Padrão Teste	0,0341	0,0862	0,061	0,0705	0,0721	0,0732	0,0741	0,0917	0,3469	0,4983
$\tau = 20$										
RMSE Teste	0,2267	0,2458	0,0982	0,0859	0,1339	0,1286	0,0881	0,1228	0,2035	0,2305
Média Teste	0,2406	0,2993	0,1429	0,1557	0,2122	0,2166	0,203	0,1913	0,219	0,3351
Desvio Padrão Teste	0,0081	0,0233	0,0266	0,0641	0,0649	0,0619	0,0752	0,0871	0,1919	0,2275
$\tau = 70$										
RMSE Teste	0,2413	0,2345	0,2174	0,2084	0,1988	0,1976	0,1846	0,1908	0,2198	0,2189
Média Teste	0,26	0,2578	0,2531	0,2397	0,2237	0,2138	0,2087	0,2195	0,2236	0,2341
Desvio Padrão Teste	0,2098	0,1922	0,1887	0,1787	0,1743	0,1673	0,1564	0,1632	0,1763	0,1821



a)



b)



c)

Figura 15. Série Mackey-Glass com conexões fixas e pesos variáveis. (%C = 35)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

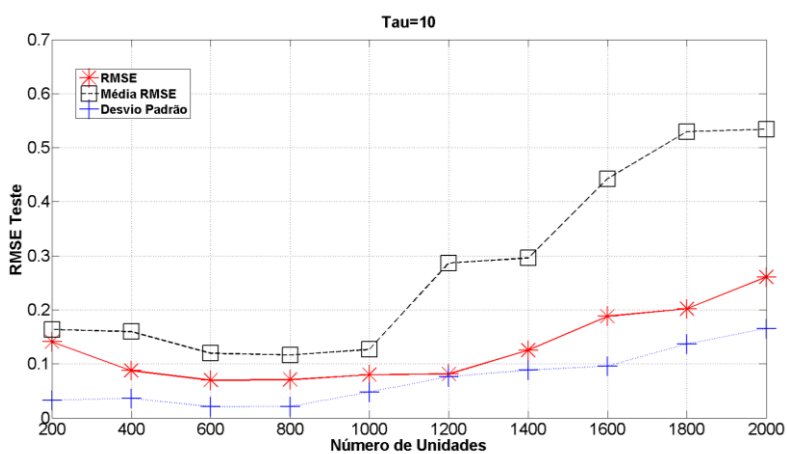
No caso dos resultados para conexões fixas e pesos variáveis, se observou um comportamento particular para as séries com $\tau = 70$, onde o erro *RMSE* obtido teve um valor alto (0.2458) com 400 unidades no *Reservoir*. Este resultado está em concordância com as configurações do modelo base (750 unidades), e permite concluir que realmente o número de unidades utilizado não consegue realizar um mapeamento completo dos padrões apresentados.

3.2.2.3 Percentual de conexões 50%

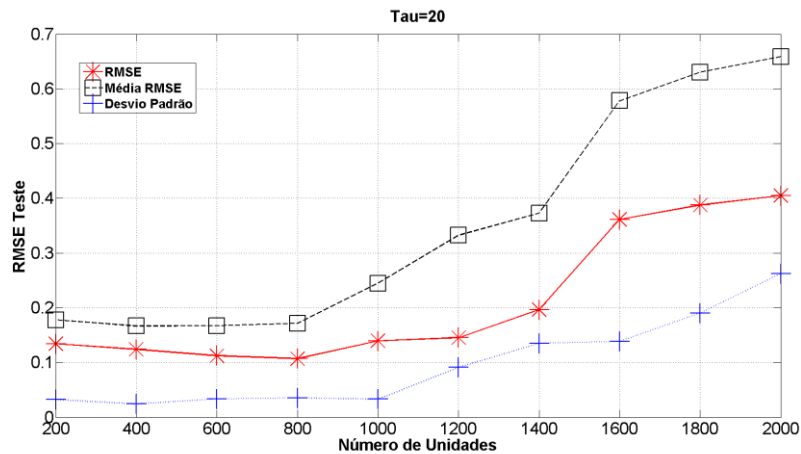
A seguir são apresentados os resultados obtidos para a duas possíveis configurações de conexões e percentual de conexões de 50%.

Tabela 13. Resultados com conexões e pesos variáveis. (%C = 50)

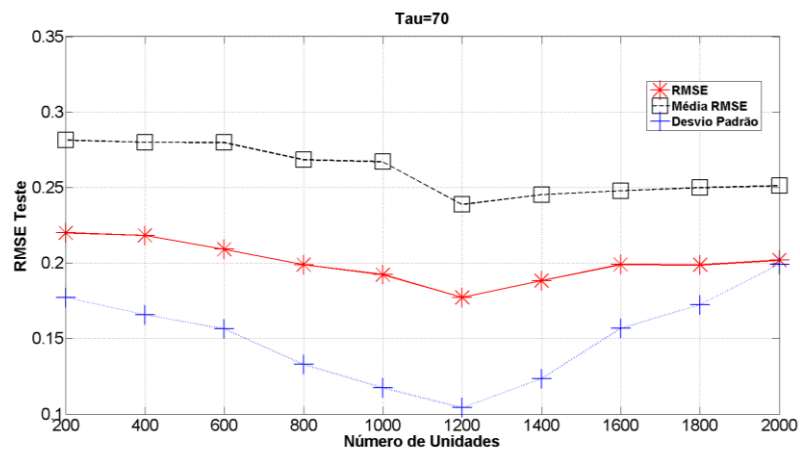
50% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,1408	0,0876	0,0698	0,0708	0,0798	0,0812	0,1254	0,1881	0,2021	0,2607
Média Teste	0,1638	0,1597	0,1196	0,1165	0,1269	0,2864	0,2959	0,4422	0,5295	0,5343
Desvio Padrão Teste	0,0329	0,0364	0,0206	0,0213	0,048	0,0765	0,0883	0,096	0,1371	0,1652
$\tau = 20$										
RMSE Teste	0,1343	0,1239	0,1123	0,1071	0,1395	0,1453	0,1965	0,3613	0,3876	0,4048
Média Teste	0,1781	0,1667	0,1672	0,1715	0,2445	0,3324	0,3728	0,5782	0,6301	0,659
Desvio Padrão Teste	0,0324	0,0243	0,0337	0,0353	0,0332	0,0911	0,1352	0,1386	0,1902	0,2624
$\tau = 70$										
RMSE Teste	0,22	0,2183	0,2091	0,1989	0,1923	0,1772	0,1883	0,199	0,1987	0,2019
Média Teste	0,2814	0,28	0,2798	0,2684	0,2671	0,2387	0,2452	0,2478	0,2499	0,2512
Desvio Padrão Teste	0,1772	0,1658	0,1564	0,1328	0,1172	0,1042	0,1235	0,1567	0,1725	0,1992



a)



b)



c)

Figura 16. Série Mackey-Glass com conexões e pesos variáveis. (%C = 50)

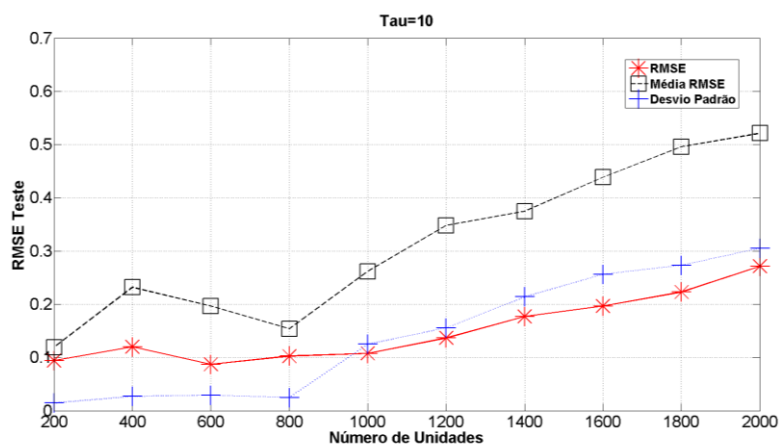
a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

Os resultados obtidos no caso do percentual de conexões 50% mostraram que o comportamento dos erros *RMSE* esteve muito perto do valor da média *RMSE* para a série $\tau = 20$ com um número de unidades entre 200 e 1000, o que pode ser entendido que nesta faixa de quantidade de unidades as previsões realizadas tiveram resultados similares, permitindo entender que o número de unidades não gerou diferenças significativas.

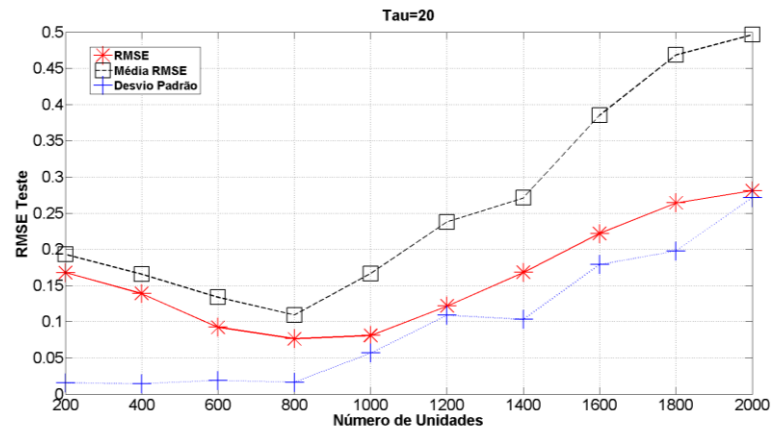
No caso de $\tau = 70$, os resultados tiveram um comportamento estável, isto é, os melhores resultados foram obtidos perto do número de unidades do modelo base (1308) e o comportamento da média *RMSE* e desvio padrão acompanharam o erro *RMSE*, com desempenhos similares. Percebe-se, portanto, que poucas unidades (200) ou muitas unidades (2000) geram os piores resultados.

Tabela 14. Resultados com conexões fixas e pesos variáveis. (%C = 50)

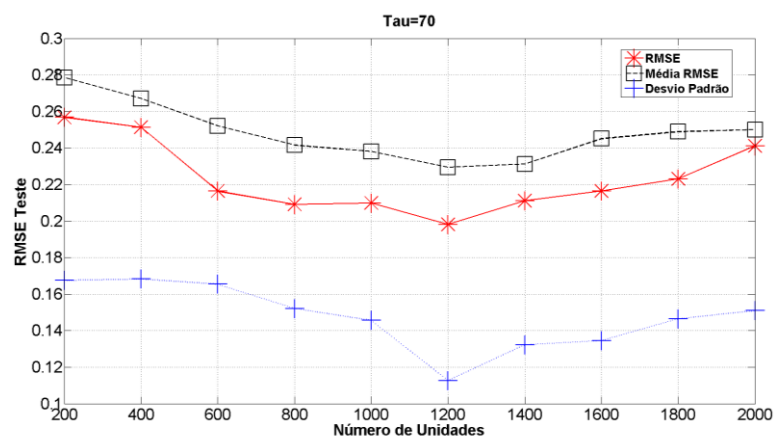
50% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,0939	0,1202	0,0872	0,1025	0,1076	0,1366	0,1768	0,1967	0,2229	0,2715
Média Teste	0,1190	0,2318	0,1964	0,1539	0,2618	0,3480	0,3749	0,4386	0,4958	0,5212
Desvio Padrão Teste	0,0143	0,0268	0,0292	0,0249	0,1248	0,1554	0,2138	0,2561	0,2734	0,3054
$\tau = 20$										
RMSE Teste	0,1677	0,1389	0,0924	0,0765	0,0810	0,1217	0,1682	0,2217	0,2640	0,2810
Média Teste	0,1930	0,1655	0,1337	0,1093	0,1665	0,2377	0,2710	0,3851	0,4684	0,4963
Desvio Padrão Teste	0,0155	0,0144	0,0190	0,0163	0,0567	0,1090	0,1032	0,1790	0,1977	0,2713
$\tau = 70$										
RMSE Teste	0,2568	0,2513	0,2164	0,2091	0,2098	0,1981	0,2111	0,2165	0,2231	0,2412
Média Teste	0,2786	0,2671	0,2521	0,2416	0,2381	0,2294	0,2312	0,2451	0,2489	0,2500
Desvio Padrão Teste	0,1675	0,1682	0,1654	0,1521	0,1456	0,1126	0,1322	0,1345	0,1465	0,1509



a)



b)



c)

Figura 17. Série Mackey-Glass com conexões fixas e pesos variáveis. (%C = 50)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

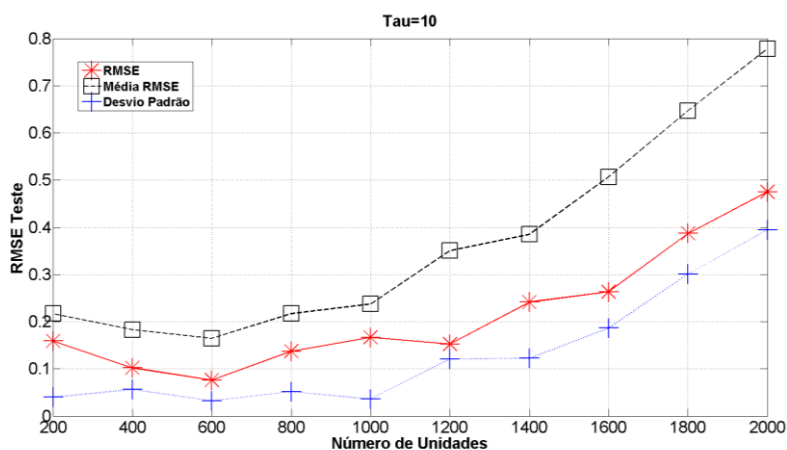
Para os testes realizados com conexões fixas e pesos variáveis, é possível constatar que os resultados obtidos apoiam a ideia de que pode existir uma configuração de rede (Conexões fixas) que gere melhores resultados do que outras configurações. Isto pode ser verificado pelos resultados do *RMSE*, onde é possível observar que o máximo valor do erro nos testes para conexões fixas é menor do que para conexões variáveis.

3.2.2.4 Percentual de conexões 75%

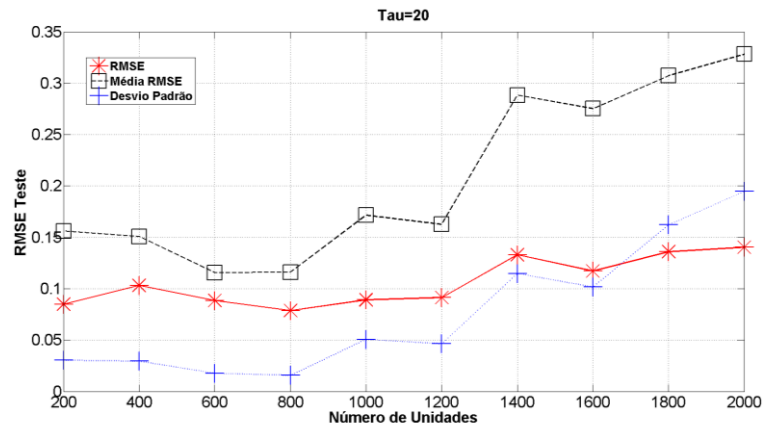
A seguir são apresentados os resultados obtidos para a duas possíveis configurações de conexões e percentual de conexões de 75%.

Tabela 15. Resultados com conexões e pesos variáveis. (%C = 75)

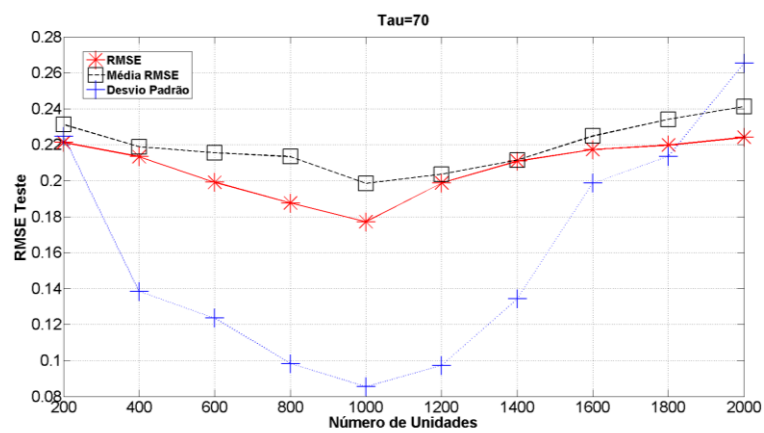
75% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,159	0,1019	0,0762	0,137	0,1666	0,1522	0,2413	0,2634	0,3869	0,4751
Média Teste	0,2168	0,1829	0,1645	0,217	0,2372	0,3504	0,3851	0,5068	0,6469	0,7784
Desvio Padrão Teste	0,0398	0,0563	0,0325	0,0516	0,0361	0,1209	0,1226	0,1866	0,3007	0,3949
$\tau = 20$										
RMSE Teste	0,0849	0,1032	0,0884	0,0786	0,0892	0,0913	0,1331	0,1172	0,136	0,1403
Média Teste	0,1561	0,1508	0,1157	0,1161	0,1716	0,1626	0,2884	0,2753	0,3073	0,3283
Desvio Padrão Teste	0,0303	0,0294	0,0178	0,0157	0,0506	0,0464	0,1148	0,1016	0,1621	0,1948
$\tau = 70$										
RMSE Teste	0,2213	0,2134	0,1991	0,1876	0,1772	0,1987	0,2109	0,2173	0,2198	0,2241
Média Teste	0,2312	0,2189	0,2156	0,2134	0,1985	0,2036	0,2115	0,2248	0,2341	0,2412
Desvio Padrão Teste	0,2247	0,1384	0,1236	0,0982	0,0854	0,0972	0,1342	0,1987	0,2135	0,2653



a)



b)



c)

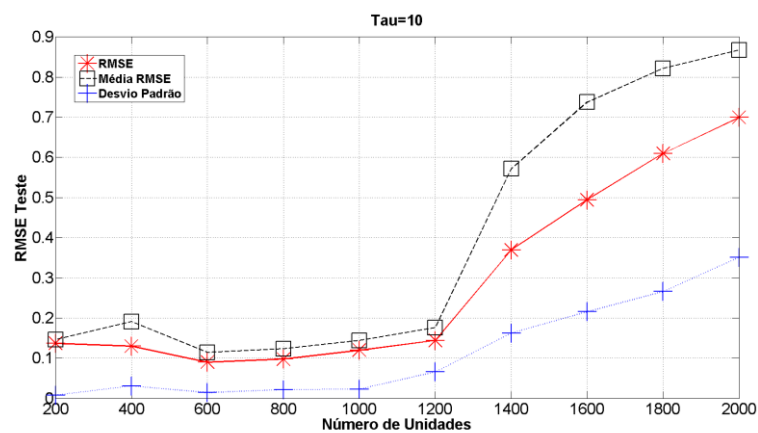
Figura 18. Série Mackey-Glass com conexões e pesos variáveis. (%C = 75)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

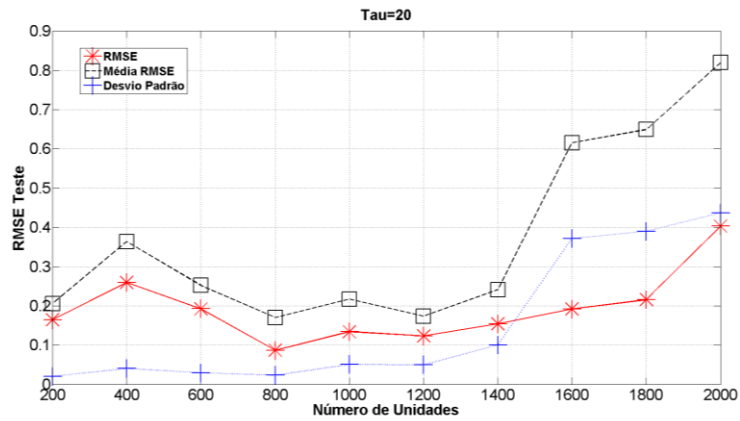
Nos testes para conexões e pesos variáveis para $\tau = 70$ a média do *RMSE* esteve muito perto do melhor *RMSE* obtido para cada configuração. Isto pode ser explicado pelo fato do percentual utilizado ser alto e estar próximo ao percentual do modelo base (61%). Mesmo assim, o melhor resultado obtido (0.1772) não conseguiu superar o resultado do modelo base 0.1681.

Tabela 16. Resultados com conexões fixas e pesos variáveis. (%C = 75)

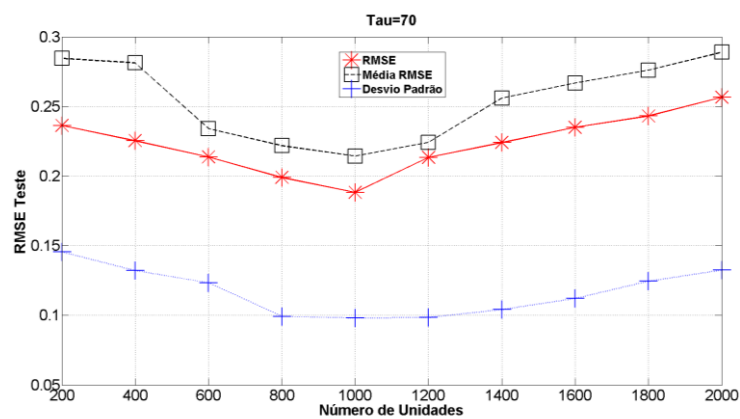
75% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,1368	0,1295	0,0898	0,0976	0,1196	0,1437	0,3692	0,4943	0,6097	0,6999
Média Teste	0,1463	0,1908	0,1142	0,1231	0,1437	0,1759	0,5715	0,7377	0,8217	0,8677
Desvio Padrão Teste	0,0074	0,0306	0,0143	0,0216	0,0227	0,0654	0,1622	0,2156	0,2657	0,3514
$\tau = 20$										
RMSE Teste	0,1647	0,2595	0,1925	0,0868	0,1342	0,1231	0,1542	0,1921	0,2157	0,4032
Média Teste	0,2060	0,3640	0,2521	0,1703	0,2170	0,1737	0,2411	0,6157	0,6494	0,8201
Desvio Padrão Teste	0,0201	0,0402	0,0295	0,0232	0,0509	0,0495	0,1003	0,3718	0,3899	0,4369
$\tau = 70$										
RMSE Teste	0,2363	0,2254	0,2138	0,199	0,1882	0,2134	0,2241	0,2351	0,2431	0,2567
Média Teste	0,2845	0,2813	0,2341	0,2218	0,2143	0,2243	0,2561	0,2668	0,2761	0,2889
Desvio Padrão Teste	0,1456	0,1321	0,1234	0,0992	0,0981	0,0984	0,1041	0,1121	0,1245	0,1325



a)



b)



c)

Figura 19. Série Mackey-Glass com conexões fixas e pesos variáveis. (%C = 75)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

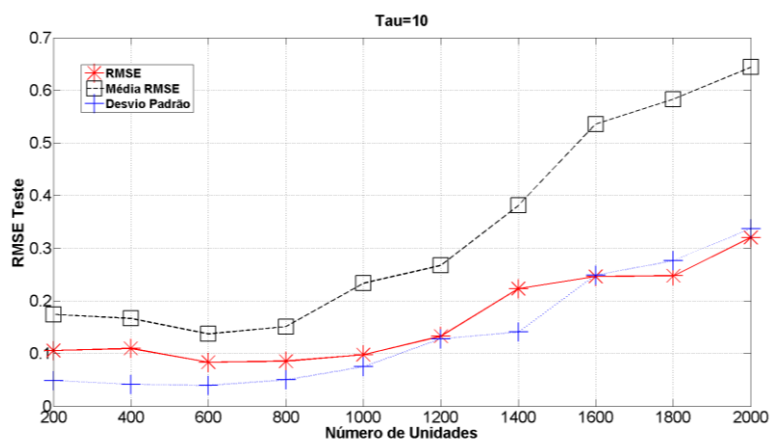
Nos testes para conexões fixas e pesos variáveis para $\tau = 10$ foi observado um comportamento particular para *Reservoirs* maiores de 1200 unidades, onde o *RMSE* teve um crescimento abrupto. Este desempenho pode ser explicado pelo fato do percentual de conexões ser maior ao utilizado no modelo base, além da grande quantidade de unidades no *Reservoir*. Estes dois parâmetros (%C e #U), em proporções maiores do que o valor ótimo, fazem com que o comportamento da rede realize uma separação dos padrões apresentados de uma forma não desejada.

3.2.2.5 Percentual de conexões 100%

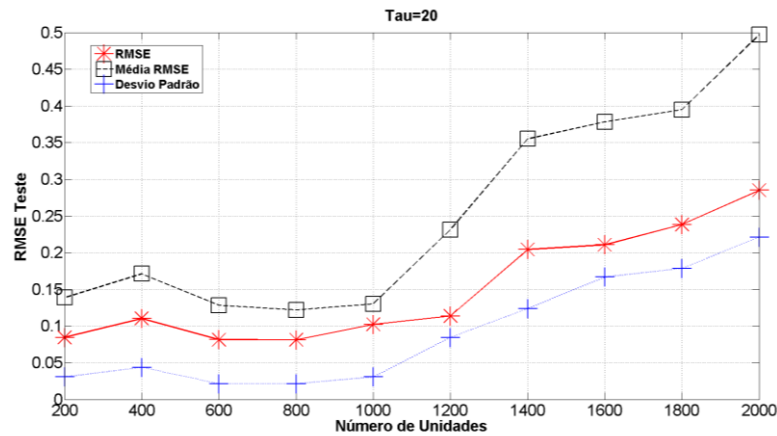
A seguir são apresentados os resultados obtidos para as duas possíveis configurações de conexões e percentual de conexões de 100%.

Tabela 17. Resultados com conexões e pesos variáveis. (%C = 100)

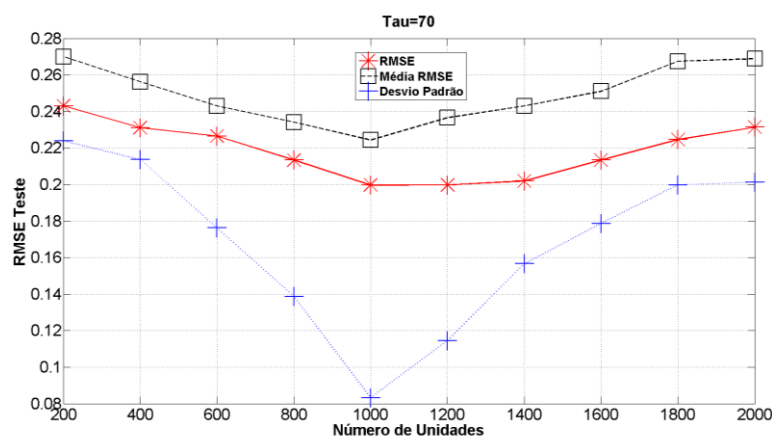
100% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,1058	0,1096	0,0837	0,0857	0,0979	0,1329	0,2232	0,2463	0,2476	0,3202
Média Teste	0,1742	0,1668	0,1375	0,1512	0,2336	0,2674	0,3812	0,5357	0,5829	0,6441
Desvio Padrão Teste	0,0489	0,0416	0,0393	0,0502	0,075	0,1279	0,1408	0,249	0,2765	0,3372
$\tau = 20$										
RMSE Teste	0,0846	0,1103	0,0815	0,0812	0,1021	0,1137	0,2042	0,2107	0,2385	0,285
Média Teste	0,1388	0,1713	0,1284	0,1219	0,1301	0,2314	0,3552	0,3784	0,3949	0,497
Desvio Padrão Teste	0,0307	0,0438	0,0215	0,0214	0,0308	0,0843	0,124	0,1669	0,1787	0,2212
$\tau = 70$										
RMSE Teste	0,2431	0,2311	0,2265	0,2134	0,1997	0,1998	0,2019	0,2134	0,2246	0,2314
Média Teste	0,27	0,2563	0,2431	0,2341	0,2243	0,2365	0,2431	0,2511	0,2675	0,2689
Desvio Padrão Teste	0,2239	0,2137	0,1762	0,1387	0,0832	0,1145	0,1567	0,1787	0,1998	0,2013



a)



b)



c)

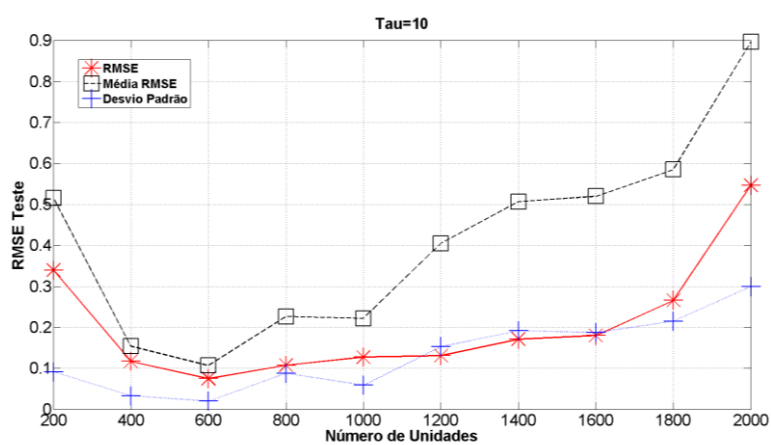
Figura 20. Série Mackey-Glass com conexões e pesos variáveis. (%C = 100)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

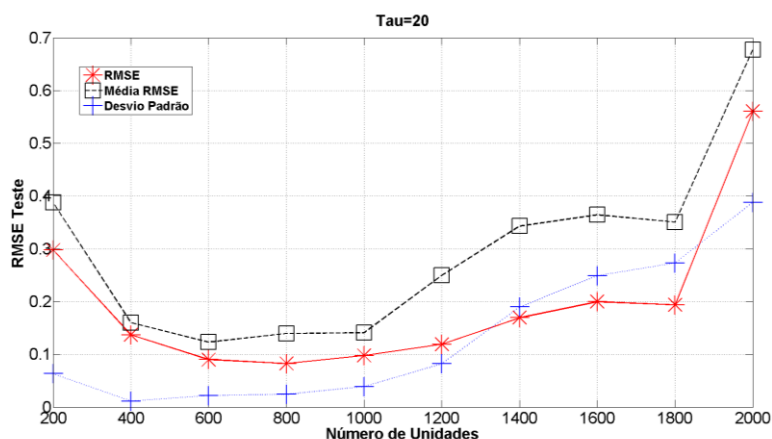
Nos testes para conexões e pesos variáveis para $\tau = 70$ o comportamento observado na figura 20 apresentou um desvio padrão muito baixo (0,0832), o que significa que as 100 redes geradas durante o teste tiveram um desempenho similar. Esse resultado indica que, quando o percentual de conexões é alto, e a série tem comportamento mais caótico, diferentes configurações de conexão com diferentes inicializações de pesos fornecem desempenhos similares e como pode ser observado na figura 20.c onde a diferença que existe entre o melhor RMSE e a média do RMSE tem desempenhos parecidos. Deste modo, é possível gerar diferentes configurações de *Reservoir* (Conexões e Pesos), mantendo os parâmetros (percentual de conexões, raio espectral, fatores de escala nos pesos das conexões de entrada e recorrentes), e ainda assim obter resultados similares, como foi evidenciado neste experimento.

Tabela 18. Resultados conexões fixas e pesos variáveis. (%C = 100)

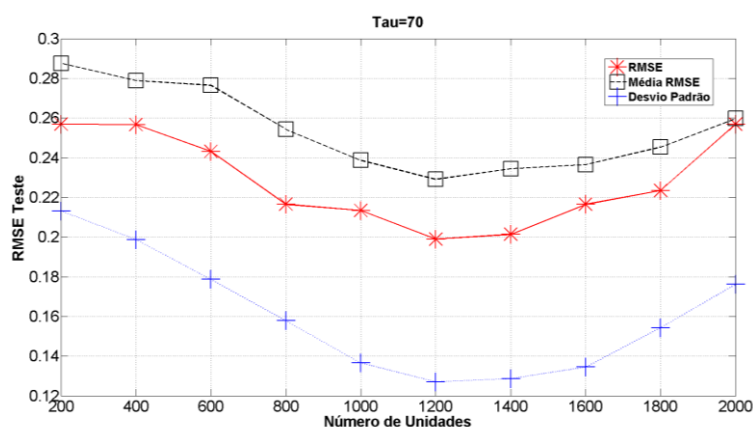
100% de Conexões	200	400	600	800	1000	1200	1400	1600	1800	2000
$\tau = 10$										
RMSE Teste	0,3396	0,1168	0,0748	0,1074	0,1272	0,1306	0,1711	0,1799	0,2659	0,5470
Média Teste	0,5161	0,1536	0,1069	0,2262	0,2219	0,4049	0,5071	0,5202	0,5855	0,8968
Desvio Padrão Teste	0,0916	0,0330	0,0198	0,0877	0,0585	0,1527	0,1922	0,1867	0,2152	0,2998
$\tau = 20$										
RMSE Teste	0,2977	0,1369	0,0902	0,0828	0,0978	0,1192	0,1695	0,1999	0,1940	0,5606
Média Teste	0,3876	0,1599	0,1233	0,1396	0,1410	0,2496	0,3431	0,3646	0,3507	0,6775
Desvio Padrão Teste	0,0636	0,0114	0,0220	0,0243	0,0389	0,0822	0,1898	0,2493	0,2731	0,3879
$\tau = 70$										
RMSE Teste	0,2569	0,2567	0,2431	0,2165	0,2134	0,1989	0,2014	0,2165	0,2234	0,2569
Média Teste	0,2876	0,2789	0,2765	0,2543	0,2387	0,2291	0,2345	0,2365	0,2453	0,2598
Desvio Padrão Teste	0,2132	0,1987	0,1787	0,1578	0,1365	0,127	0,1287	0,1345	0,1543	0,1762



a)



b)



c)

Figura 21. Série Mackey-Glass com conexões fixas e pesos variáveis. (%C = 100)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

Para os testes com conexões fixas e pesos variáveis, não se obteve uma resposta diferente do padrão. O desempenho acompanhou os valores dos modelos base. Neste caso em particular, foi possível observar que o número de unidades afetou o desempenho nas previsões, gerando erros mais altos do que os erros gerados pela mudança no percentual de conexão. Esta conclusão pode ser inferida dos resultados obtidos em geral, onde, desempenhos similares (melhores resultados de cada teste) foram obtidos ao longo dos 5 testes realizados com diferentes percentuais de conexões.

3.2.3.

Fator de escala nos pesos das conexões de entrada

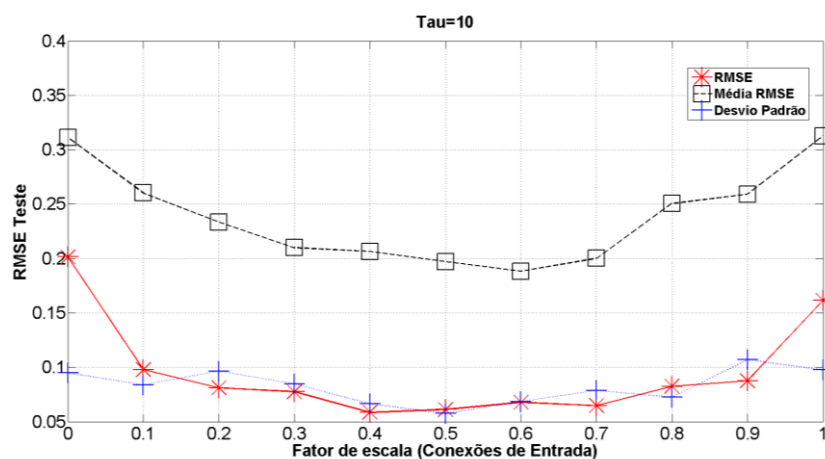
Nesta seção, o parâmetro a ser modificado é o fator de escala dos pesos das conexões de entrada, o qual é usado para modificar (ou não) os pesos das conexões de entrada. Neste caso, o valor utilizado é um fator multiplicativo que varia dentro do intervalo [0,1]. Nesta análise foram avaliados fatores dentro dessa escala com incrementos unitários na primeira casa decimal para cada teste.

Quando o fator utilizado é zero (0), todas as conexões de entradas são suprimidas; já quando é definido com o valor máximo (1,0), os pesos permanecem iguais aos gerados inicialmente.

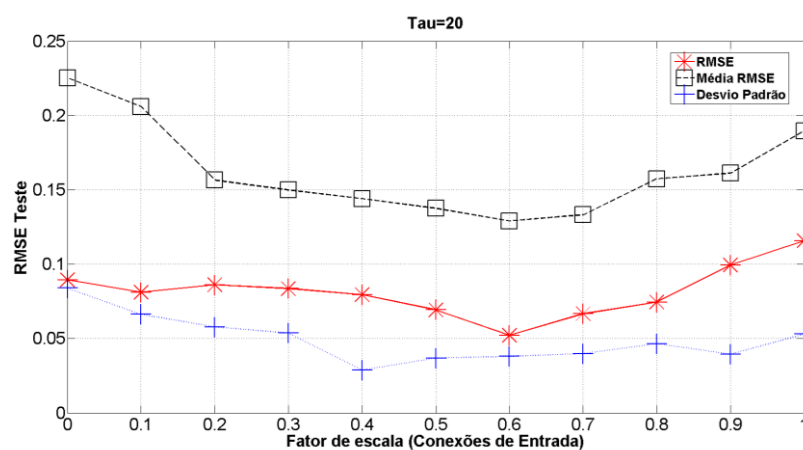
Nas figuras 22 e 23 e nas tabelas 19 e 20 são apresentados os resultados obtidos para as duas possíveis configurações de conexões e pesos.

Tabela 19. Resultados com conexões e pesos variáveis. (sW^e)

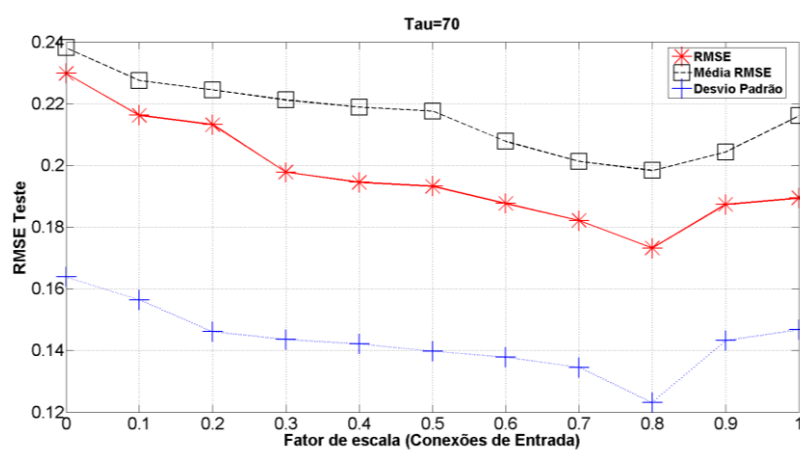
	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$\tau = 10$											
RMSE Teste	0,2015	0,0979	0,0812	0,0776	0,0586	0,0612	0,0678	0,0649	0,0826	0,0877	0,1614
Média Teste	0,3112	0,2603	0,2334	0,2099	0,2064	0,1971	0,1884	0,2001	0,2506	0,2591	0,3128
Desvio Padrão Teste	0,0948	0,0838	0,0965	0,0849	0,0664	0,0576	0,0685	0,0787	0,0726	0,1069	0,0976
$\tau = 20$											
RMSE Teste	0,0895	0,0811	0,0862	0,0836	0,0795	0,0693	0,0522	0,0666	0,0745	0,0995	0,1157
Média Teste	0,2253	0,2059	0,1566	0,1499	0,144	0,1376	0,1291	0,1332	0,1573	0,1612	0,1897
Desvio Padrão Teste	0,084	0,0664	0,0579	0,0536	0,0288	0,0368	0,038	0,0398	0,0466	0,0395	0,053
$\tau = 70$											
RMSE Teste	0,2298	0,2162	0,2132	0,1978	0,1945	0,1932	0,1876	0,1821	0,1732	0,1873	0,1894
Média Teste	0,2381	0,2275	0,2245	0,2212	0,2189	0,2176	0,2078	0,2013	0,1983	0,2043	0,2161
Desvio Padrão Teste	0,1638	0,1564	0,1461	0,1435	0,1421	0,1398	0,1378	0,1345	0,1231	0,1432	0,1467



a)



b)



c)

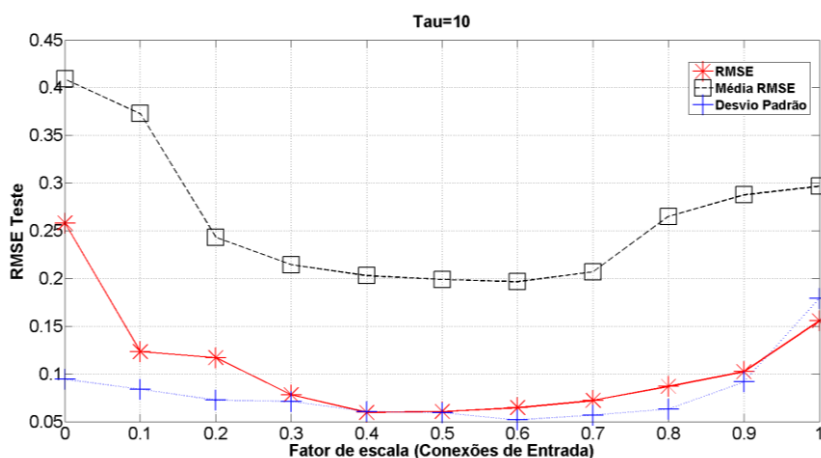
Figura 22. Série Mackey-Glass com conexões e pesos variáveis. (sW^e)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

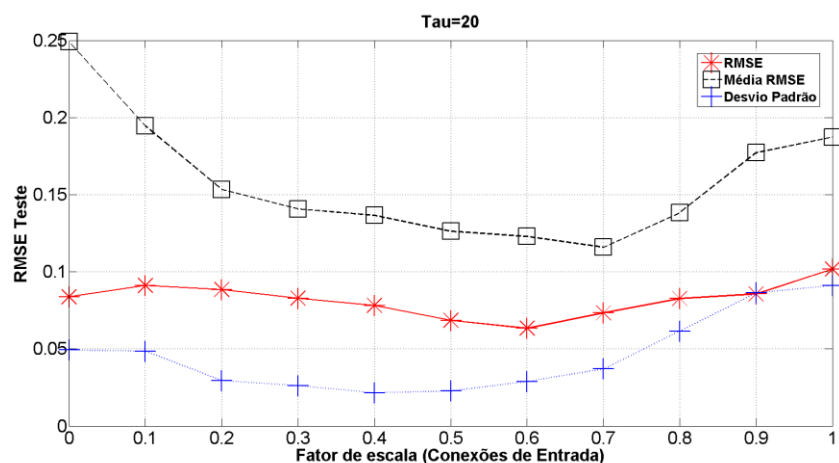
Os resultados obtidos para as 3 séries confirmaram os valores selecionados pelo GA no modelo base ($\tau = 10 \rightarrow 0.4$; $\tau = 20 \rightarrow 0.6$ e $\tau = 70 \rightarrow 0.8$). Analisando o desvio padrão, observa-se que as séries com $\tau = 10$ e $\tau = 20$ tiveram erros que acompanharam o melhor RMSE obtido. Já no caso de $\tau = 70$, o desvio padrão apresenta uma diferença significativa, permitindo concluir que o aumento do fator de escala nas conexões de entrada precisou ser maior para séries mais caóticas.

Tabela 20. Resultados com conexões fixas e pesos variáveis. (sW^e)

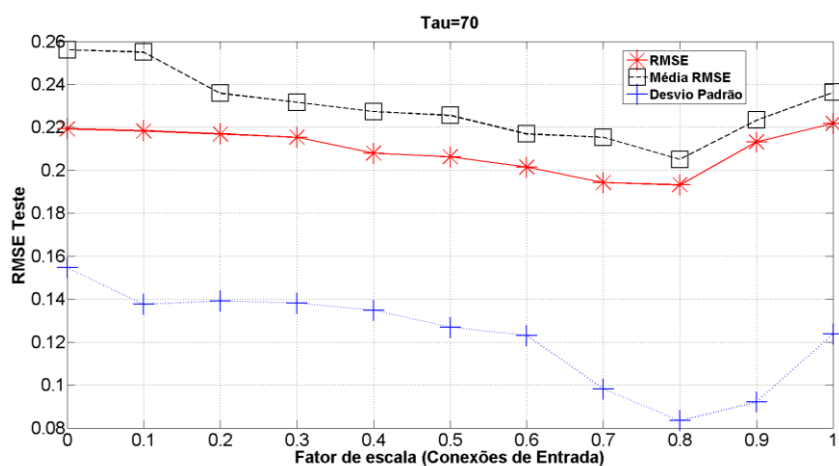
	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$\tau = 10$											
RMSE Teste	0,2581	0,1235	0,117	0,078	0,0597	0,0604	0,0646	0,0721	0,0869	0,1023	0,1556
Média Teste	0,4092	0,3728	0,2432	0,2144	0,203	0,199	0,1965	0,2069	0,265	0,2877	0,2968
Desvio Padrão Teste	0,0945	0,0839	0,0725	0,0713	0,0605	0,0595	0,0516	0,0568	0,0632	0,0916	0,1792
$\tau = 20$											
RMSE Teste	0,0837	0,0912	0,0885	0,0828	0,0782	0,0685	0,0634	0,0733	0,0826	0,0856	0,1017
Média Teste	0,2491	0,1945	0,1533	0,1407	0,1365	0,1263	0,1229	0,1159	0,1383	0,1772	0,1874
Desvio Padrão Teste	0,0494	0,0485	0,0296	0,0262	0,0216	0,0228	0,0289	0,037	0,0615	0,0863	0,0911
$\tau = 70$											
RMSE Teste	0,2193	0,2183	0,2169	0,2153	0,2079	0,2062	0,2014	0,1943	0,1932	0,2132	0,2218
Média Teste	0,2561	0,255	0,2358	0,2317	0,2274	0,2256	0,2169	0,2154	0,2051	0,2234	0,2363
Desvio Padrão Teste	0,1546	0,1376	0,1392	0,1382	0,1348	0,1269	0,1231	0,0981	0,0834	0,0921	0,1238



a)



b)



c)

Figura 23. Série Mackey-Glass com conexões fixas e pesos variáveis. (sW^e)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

Dentro dos testes realizados, modificando o fator de escala nos pesos das conexões de entrada, foi possível constatar que a informação recebida no *Reservoir* é atenuada em menor ou maior grau, dependendo do valor utilizado neste parâmetro.

As informações que estão sendo entregues ao *Reservoir* mediante as conexões estabelecidas no projeto da rede, incluem também as conexões recorrentes, por tanto na próxima subseção a análise dos fatores de escala considerará também o fator de escala nos pesos das conexões recorrentes. A razão para tal conduta é: quando é projetada a rede, estes fatores são os encarregados de manter o equilíbrio das informações recebidas pelo *Reservoir*, o que significa que existe um compromisso no qual a informação do

padrão apresentado à rede deve conter amostras do padrão atual $u(n)$, no caso da camada de entrada, e $u(n - 1)$, no caso da conexão recorrente.

3.2.4.

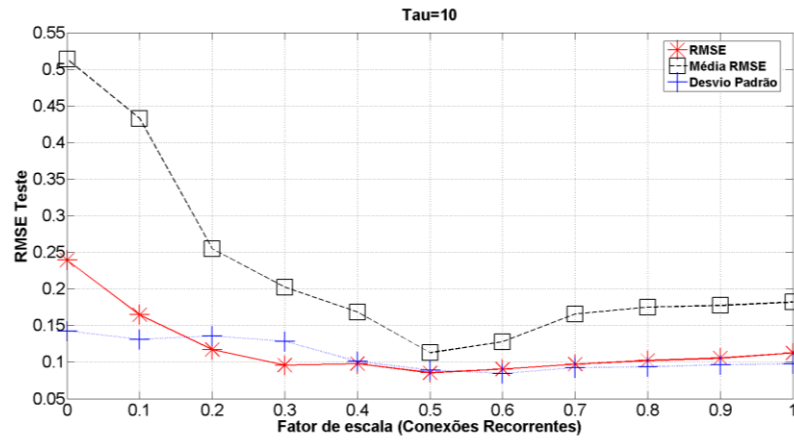
Fator de Escala nas conexões recorrentes

A utilização do fator de escala nos pesos das conexões recorrentes também é um dos parâmetros analisados neste capítulo. Os valores deste fator também variam entre 0 e 1, e, como na seção anterior; quando o valor utilizado é 0 significa que não existem conexões recorrentes, e valor igual a 1 representa os pesos iguais aos criados no momento do projeto da rede.

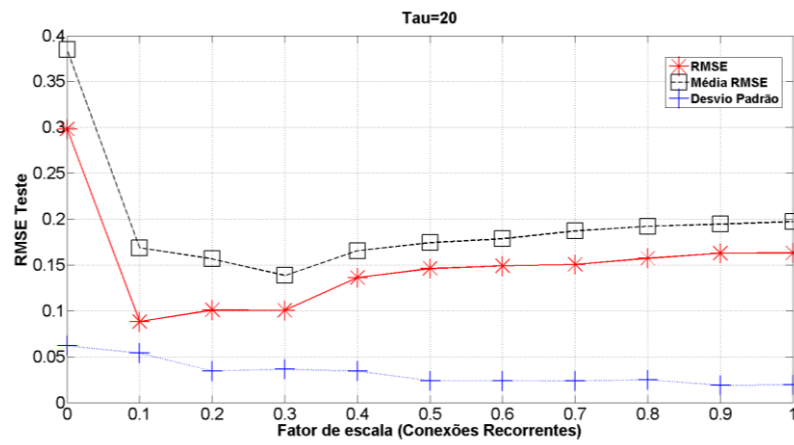
Nas figuras 24 e 25 e nas tabelas 21 e 22 são apresentados os resultados obtidos para ambas as configurações de conexões e pesos.

Tabela 21. Resultados com conexões e pesos variáveis. (sW^{rr})

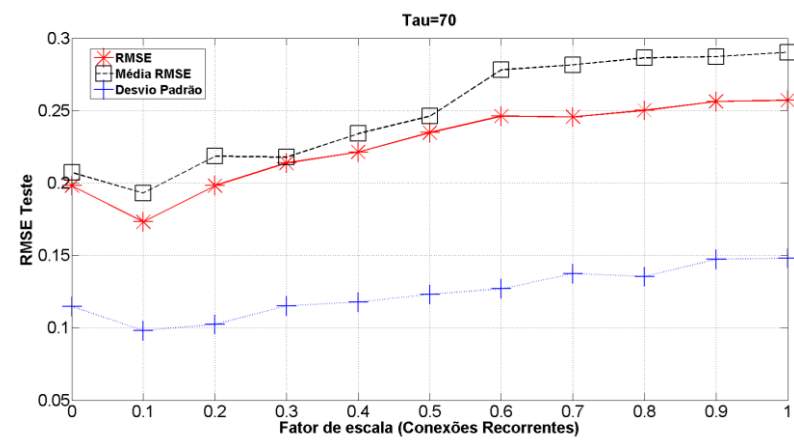
	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$\tau = 10$											
RMSE Teste	0,2391	0,1646	0,1171	0,0959	0,0979	0,0853	0,0905	0,0972	0,1021	0,1054	0,1123
Média Teste	0,5138	0,4326	0,2547	0,2022	0,1682	0,1129	0,1278	0,1658	0,1749	0,1772	0,1821
Desvio Padrão Teste	0,1424	0,1311	0,1356	0,1285	0,1007	0,089	0,0845	0,0925	0,0934	0,0965	0,0976
$\tau = 20$											
RMSE Teste	0,2982	0,0883	0,1009	0,1005	0,1361	0,1462	0,1489	0,1505	0,1572	0,1628	0,1631
Média Teste	0,3846	0,1686	0,1566	0,1387	0,1655	0,1743	0,1786	0,1872	0,1921	0,1945	0,1972
Desvio Padrão Teste	0,0619	0,054	0,0345	0,0363	0,0342	0,0239	0,0239	0,0236	0,0249	0,0189	0,0193
$\tau = 70$											
RMSE Teste	0,1982	0,1732	0,1982	0,2138	0,2213	0,2348	0,2461	0,2457	0,2501	0,2563	0,2571
Média Teste	0,2073	0,1931	0,2185	0,2179	0,2341	0,2461	0,2781	0,2816	0,2865	0,2873	0,2903
Desvio Padrão Teste	0,1147	0,0982	0,1023	0,1152	0,1179	0,1231	0,1269	0,1375	0,1354	0,1473	0,1482



a)



b)



c)

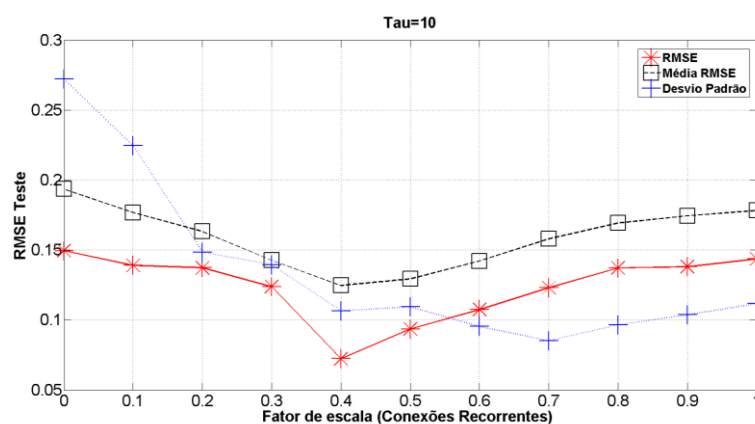
Figura 24. Série Mackey-Glass com conexões e pesos variáveis. (sW^{rr})

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

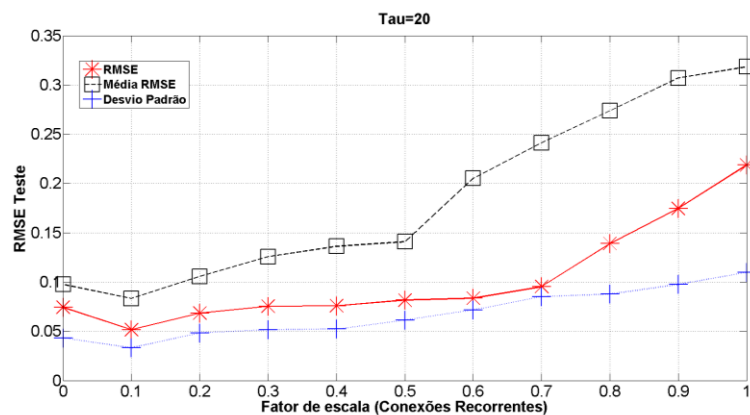
Nos resultados obtidos para os testes de conexões e pesos variáveis é possível observar que os melhores resultados para cada fator de escala esteve muito perto dos valores médios de *RMSE*. Também é possível observar que o desvio padrão esteve mais longe dos valores de *RMSE* e média *RMSE* para a série com $\tau = 70$, o que pode ser entendido pela natureza caótica da própria série.

Tabela 22. Resultados com conexões fixas e pesos variáveis. (sW^{rr})

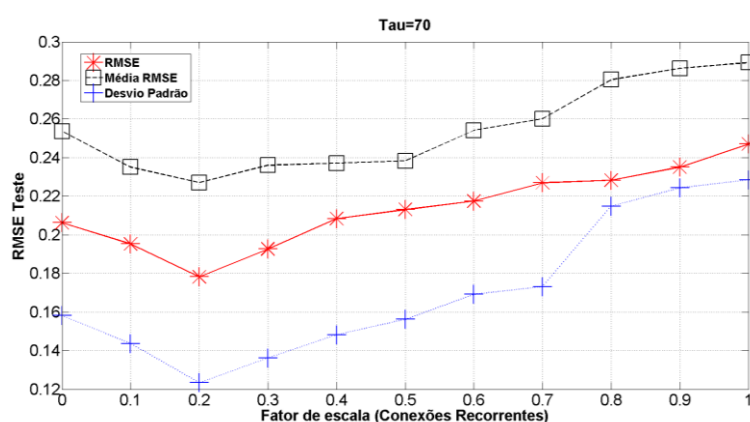
	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$\tau = 10$											
RMSE Teste	0,1494	0,1389	0,1373	0,1237	0,0724	0,0936	0,1074	0,123	0,1372	0,138	0,1436
Média Teste	0,1936	0,1768	0,1632	0,1427	0,1246	0,1293	0,1421	0,1579	0,1692	0,1745	0,1782
Desvio Padrão Teste	0,2722	0,2245	0,1485	0,1395	0,1065	0,1093	0,0954	0,0852	0,0966	0,1037	0,1117
$\tau = 20$											
RMSE Teste	0,0743	0,0517	0,0683	0,0753	0,0759	0,0816	0,0835	0,0953	0,1389	0,1746	0,2187
Média Teste	0,0975	0,0834	0,1057	0,1256	0,1363	0,1408	0,2052	0,2414	0,2738	0,3069	0,3185
Desvio Padrão Teste	0,04329	0,0333	0,0484	0,0517	0,0522	0,0613	0,0716	0,0851	0,0878	0,0976	0,1098
$\tau = 70$											
RMSE Teste	0,2063	0,1953	0,1783	0,1927	0,2084	0,2131	0,2175	0,2269	0,2283	0,2351	0,2471
Média Teste	0,2537	0,2351	0,2271	0,2361	0,2372	0,2384	0,2541	0,2601	0,2804	0,2863	0,2893
Desvio Padrão Teste	0,1582	0,1437	0,1234	0,1362	0,1482	0,1563	0,1692	0,1732	0,2149	0,2243	0,2285



a)



b)



c)

Figura 25. Série Mackey-Glass com conexões fixas e pesos variáveis . (sW^{rr})

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

Na figura 25 são apresentados os resultados para a mudança no fator de escala nos pesos das conexões recorrentes. Onde pode-se concluir que valores pequenos de fatores de escala nas conexões recorrentes são pouco relevantes para gerar estados do *Reservoir* que permitam melhores previsões. Outra observação realizada nesta experiência diz respeito aos valores de erro mais altos, pois estes resultados foram menores do que nas experiências prévias (Número de unidades e percentual de conexões). Assim pode-se assumir que estes parâmetros tem um grau de importância menor no projeto da rede.

3.2.5.

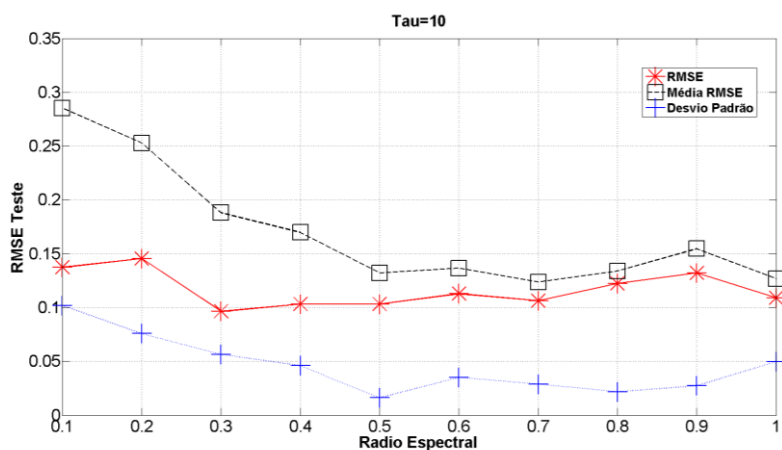
Raio Espectral

O último teste realizado procura identificar a influência do fator de escala utilizado nos pesos do *Reservoir*. De acordo com a teoria básica das redes *ESN*, os pesos do *Reservoir* devem permanecer fixos. No entanto, (Ferreira, 2011) apresentou a utilização do Raio Espectral $\rho(w)$ como uma forma de controlar a não linearidade no comportamento dos estados gerados.

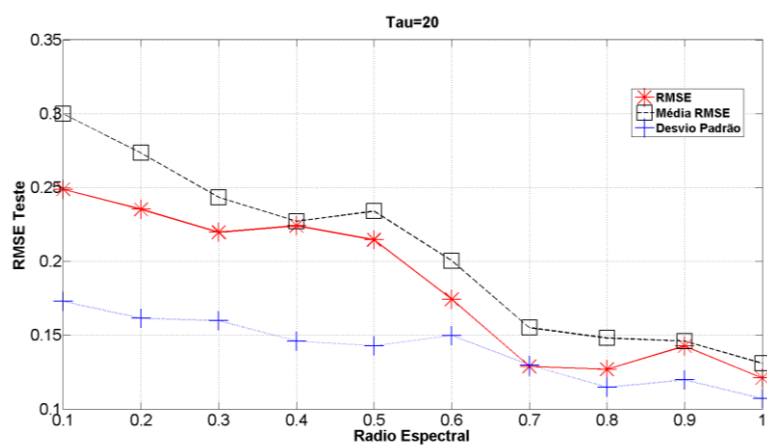
Os valores utilizados para serem testados no modelo começam em 0,1 e terminam em 1, que é o máximo valor permitido. Em cada teste realizado, este fator foi incremento em 0,1. Nas figuras 26 e 27 e nas tabelas 23 e 24 são apresentados os resultados obtidos para cada valor utilizado.

Tabela 23. Resultados com conexões e pesos variáveis. ($\rho(W)$)

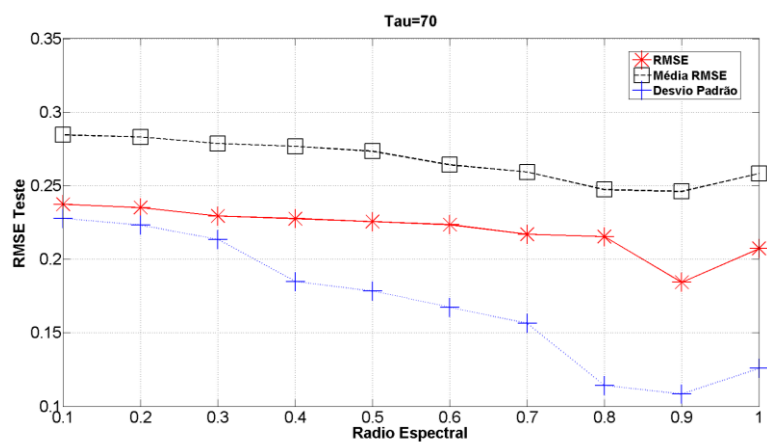
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$\tau = 10$										
RMSE Teste	0,1373	0,1456	0,0964	0,1033	0,1031	0,1128	0,1064	0,1224	0,1321	0,1089
Média Teste	0,2854	0,2529	0,188	0,1698	0,1321	0,1366	0,1237	0,1339	0,1547	0,1269
Desvio Padrão Teste	0,1019	0,0757	0,0565	0,0459	0,0162	0,0351	0,0289	0,0218	0,0272	0,0497
$\tau = 20$										
RMSE Teste	0,2487	0,2352	0,2197	0,2241	0,2146	0,1743	0,1288	0,127	0,1427	0,121
Média Teste	0,2999	0,2734	0,2433	0,2271	0,234	0,2004	0,1551	0,1481	0,1461	0,1309
Desvio Padrão Teste	0,1729	0,1614	0,1599	0,1461	0,1429	0,1498	0,1298	0,1147	0,1198	0,1071
$\tau = 70$										
RMSE Teste	0,2374	0,2351	0,2293	0,2276	0,2256	0,2235	0,2169	0,2154	0,1843	0,2073
Média Teste	0,2845	0,2831	0,2786	0,2768	0,2734	0,2642	0,2593	0,2473	0,2461	0,2583
Desvio Padrão Teste	0,2278	0,2231	0,2134	0,1849	0,1784	0,1673	0,1564	0,1143	0,1083	0,1258



a)



b)



c)

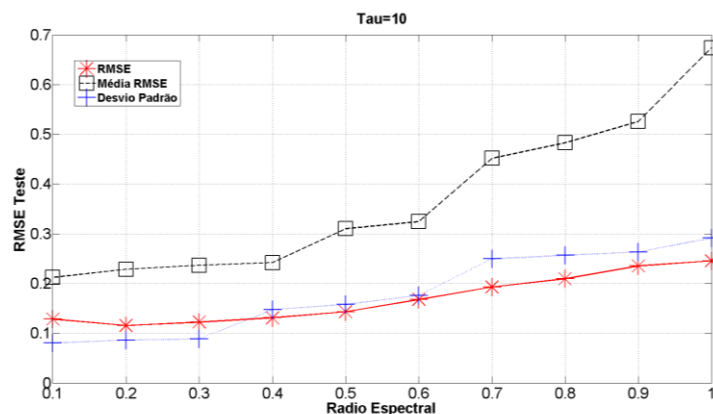
Figura 26. Série Mackey-Glass com conexões e pesos variáveis. ($\rho(W)$)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

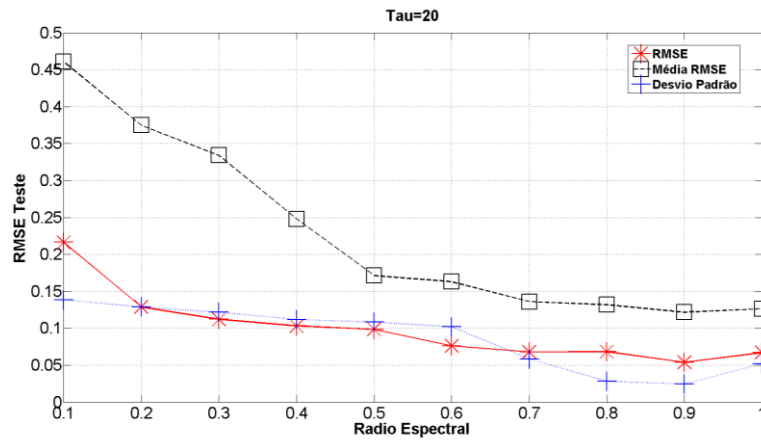
Os resultados obtidos demonstraram que para valores altos do Raio Espectral ($\rho(w)$), os erros obtidos são menores. Neste caso é fácil constatar que estes valores de Raio Espectral mais altos estão próximos ao valor da configuração base $\tau = 10 \rightarrow 0.2$; $\tau = 20 \rightarrow 0.9$ e $\tau = 70 \rightarrow 0.8$. Assim, valores diferentes de $\rho(w)$ do modelo base apresentaram erros mais altos, o que indica uma maior influência deste parâmetro no desempenho. Pode-se concluir que o controle da dinamicidade da atualização dos estados, isto é, o quanto a informação do estado anterior será considerada para o estado seguinte, é determinada mediante a manipulação da intensidade das conexões no *Reservoir*.

Tabela 24. Resultados com conexões fixas e pesos variáveis. ($\rho(W)$)

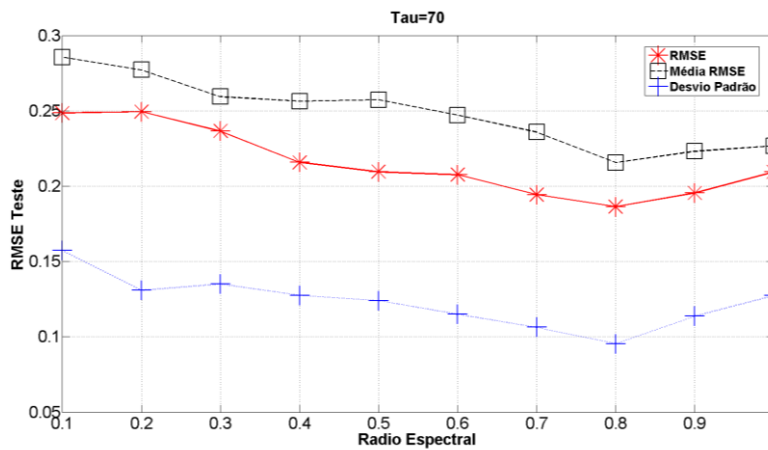
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$\tau = 10$										
RMSE Teste	0,1283	0,1157	0,1224	0,1311	0,143	0,1677	0,1931	0,2095	0,2352	0,2455
Média Teste	0,2124	0,2285	0,2366	0,2422	0,3103	0,3245	0,4516	0,4829	0,5256	0,674
Desvio Padrão Teste	0,0802	0,0862	0,0886	0,1477	0,1578	0,1761	0,2498	0,2569	0,2636	0,2915
$\tau = 20$										
RMSE Teste	0,2164	0,1288	0,112	0,1032	0,0985	0,0761	0,0679	0,0681	0,0539	0,0671
Média Teste	0,4608	0,3748	0,334	0,248	0,1714	0,163	0,136	0,1318	0,1218	0,1263
Desvio Padrão Teste	0,1384	0,1286	0,1218	0,1117	0,1082	0,1019	0,0585	0,0282	0,0248	0,0519
$\tau = 70$										
RMSE Teste	0,2485	0,2494	0,2365	0,2158	0,2095	0,2076	0,1943	0,1864	0,1954	0,2094
Média Teste	0,2856	0,2771	0,2594	0,2563	0,2573	0,2471	0,2358	0,2157	0,2231	0,2268
Desvio Padrão Teste	0,1573	0,1308	0,135	0,1276	0,1241	0,1149	0,1063	0,0954	0,1137	0,1274



a)



b)



c)

Figura 27. Série Mackey-Glass com conexões fixas e pesos variáveis. ($\rho(W)$)

a) $\tau = 10$. b) $\tau = 20$. c) $\tau = 70$.

Finalmente, a análise modificando o valor do Raio Espectral para conexões fixas e pesos variáveis foi realizada e, segundo as figuras 26 e 27, apresentou resultados que concordam com a teoria (Jaeger, 2010). Neste caso é possível alterar os pesos das conexões do *Reservoir* utilizando um fator de escala. Entretanto, como a função de ativação sendo utilizada nos neurônios do *Reservoir* é a tangente hiperbólica, este parâmetro não deve ultrapassar o valor 1, como discutido no Capítulo 2, pois as saídas desses neurônios podem saturar e, conseqüentemente, apresentar um comportamento alternando valores 0 e 1.

3.3.

Análise e discussão dos resultados

Como observação geral, pode-se concluir que, em relação aos resultados obtidos utilizando os dois tipos de configuração de topologia de rede, isto é, a primeira com conexões e pesos variáveis, e a segunda com conexões fixas e pesos variáveis, os resultados apresentaram uma variação baixa, isto é, na maioria das experiências os resultados apresentam um comportamento similar. Esta pequena variação ocorre devido às mudanças de pesos e conexões, fazendo com que, apesar da rede conservar outros parâmetros iguais, a mudança de topologia gere resultados levemente diferentes. Embora estes resultados apresentem este tipo de comportamento, é possível identificar globalmente qual seria a configuração ótima, assim como a relação com os modelos base para cada série.

A sensibilidade na configuração dos parâmetros pode ser percebida nos resultados obtidos. A simples mudança de um parâmetro faz com que os resultados mudem para valores de previsão não desejáveis. Como parâmetros mais críticos foram identificados o número de unidades no *Reservoir*, as conexões do *Reservoir* e os pesos das conexões do *Reservoir*. Esta identificação se baseia em comportamentos observados nos testes, onde, por exemplo, foi possível identificar redes com parâmetros gerais inalterados mas com respostas diferentes (comportamento caótico) o que pode ser causado pela inicialização de pesos. Também foi possível identificar redes com configurações similares de parâmetros gerais apresentando respostas similares (comportamento de atrator).

Uma outra conclusão observada é a relação de complemento que geram em conjunto o fator de escala dos pesos das conexões de entrada e fator de escala dos pesos das conexões recorrentes, onde, valores altos (perto de 1) em qualquer um dos dois fatores é acompanhado por um valor baixo (perto de 0) no fator complementar da configuração, isto é, para valores altos no fator de escala dos pesos de entrada, o valor do fator de escala nos pesos das conexões recorrentes é baixo e vice versa. Uma interpretação possível para este resultado é o nível de contribuição que cada uma das conexões tem na construção dos estados que permitem mapear a série, onde se pode dizer que estes dois parâmetros tem uma relação complementar.

No próximo capítulo será apresentado o modelo proposto, que procura aproveitar a otimização de parâmetros gerais, além realizar uma análise baseada na interpretação dos estados gerados pelo *Reservoir*.

4.

Modelo Híbrido *Echo State Network-Genetic Algorithm-Separation Ratio Graph*

Este capítulo apresenta a proposta de um novo sistema híbrido neuro-evolucionário, denominado *Echo State Network-Genetic Algorithm-Separation Ratio Graph* (ESN-GA-SRG), o qual otimiza os parâmetros gerais de uma ESN utilizando um algoritmo genético. Além disso, é realizada uma análise dos estados gerados pelo *Reservoir*, utilizando o método *Separation Ratio Graph*, com o intuito de determinar a configuração adequada do *Reservoir* (conexões e pesos).

Além de oferecer uma maior compreensão do projeto das ESNs, a análise de sensibilidade realizada no capítulo anterior permitiu identificar quais são os parâmetros referentes ao *Reservoir* que, ao serem alterados, geram maiores variações de desempenho no modelo. Os parâmetros que se mostraram mais sensíveis foram: o número de unidades no *Reservoir*, o percentual de conexões do *Reservoir* e os pesos das conexões do *Reservoir*. Esta constatação indicou a necessidade de se aumentar o foco de pesquisa no entendimento das ESNs sobre tais parâmetros.

Na procura de alternativas que permitissem avaliar o comportamento do *Reservoir* no desempenho geral da rede, no Capítulo 2 foram apresentados modelos que permitem avaliar os estados internos gerados durante a apresentação dos padrões. Desta forma, é possível realizar uma análise dos estados gerados utilizando o coeficiente de separação ou o coeficiente de *Lyapunov*, que visa a avaliar se a topologia gerada é adequada para a tarefa escolhida, neste caso previsão *Multi-Step*.

Inicialmente foram estudados diferentes modelos que avaliassem o comportamento dos estados gerados quando apresentados os padrões ao *Reservoir*, e que tivessem uma sinergia com o modelo original proposto pelo Jagger. Neste caso, foram encontradas uma variedade de propostas (Goodman, 2006), (Norton, 2008), (Schrauwen, 2008), que se eventualmente melhoravam o desempenho final, nenhuma delas realizava um estudo aprofundado dos estados quando gerados, além de carecer de um modelo que adaptasse estas propostas ao modelo básico das ESNs.

A ideia de integrar estes modelos de avaliação do *Reservoir* ao Algoritmo Genético, utilizado para otimização, pretende facilitar o projeto e permitir melhorar a resposta em tarefas de previsão de séries temporais *Multi-Step*, quando redes do tipo ESNs são usadas.

A utilização de Algoritmos Genéticos para a otimização de parâmetros globais na configuração da rede ESN foi empregada em trabalhos anteriores (Xu, 2005), (Ferreira, 2011), (Velasco, 2014), com resultados bastante satisfatórios.

As seções a seguir apresentam a visão geral do modelo híbrido proposto e detalha seus respectivos módulos separadamente.

4.1.

Visão Geral

A Figura 28 apresenta a visão geral do modelo **Echo State Network-Genetic Algorithm-Separation Ratio Graph (ESN-GA-SRG)** proposto neste trabalho.

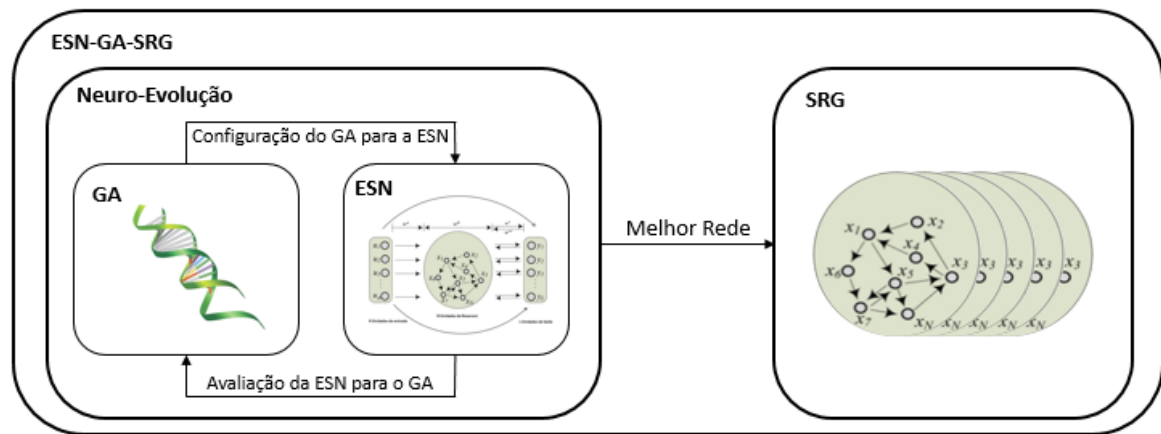


Figura 28. Diagrama completo do modelo proposto.

Na estrutura global do modelo proposto destacam-se duas partes, as quais são:

1. **NEURO-EVOLUÇÃO:** Este módulo é utilizado para definir a melhor configuração de parâmetros globais de uma rede *ESN* para previsão de séries temporais. Para esta tarefa é utilizado um algoritmo genético com n número de gerações, que tem como objetivo determinar os valores de: Número de unidades ($\#U$), Percentual de Conexão ($\%C$), Raio Espectral $\rho(W)$, Fator de escala dos pesos das conexões de entrada (sW_{in}) e Fator de escala dos pesos das conexões recorrentes (sW^{rr}). Além desses parâmetros, o algoritmo genético também determina o tamanho da janela de entrada (Jan), isto é, o número de valores da série histórica que devem ser apresentados à entrada da *ESN*. Como função aptidão é utilizada uma métrica de erro na previsão do conjunto de validação.
2. **SRG:** Após a determinação da melhor configuração de rede *ESN* pelo módulo de Neuro-Evolução, o módulo *SRG* irá determinar qual o melhor *Reservoir* para a topologia gerada pela Neuro-Evolução. Para realizar este procedimento de escolha do *Reservoir*, são criadas n redes *ESN* com a melhor topologia do módulo de Neuro-Evolução. Para cada uma dessas n redes, são apresentados os padrões de entradas

(que previamente são emparelhados utilizando um algoritmo de vizinhança) e coletados os estados gerados pelos neurônios dos n *Reservoirs*. Finalmente, será selecionado aquele *Reservoir* que apresentar a melhor avaliação. Este procedimento é, portanto, um ajuste fino da rede, implementando uma busca local realizada para n inicializações do *Reservoir*.

A seguir, nas próximas 2 seções, serão apresentados em maior detalhe cada um dos módulos que compõem o modelo proposto.

4.2.

Modulo Neuro-Evolução

Como explicado na seção anterior este modulo se encarrega de definir a melhor configuração de parâmetros globais mediante a utilização de Algoritmos Genéticos para a rede ESN.

Este tipo de técnica inteligente, tem sido utilizada amplamente na otimização de problemas, é indicada para os casos nos quais os modelos determinísticos não têm capacidade de considerar todas as variáveis, dada a complexidade do espaço de busca gerado, ou em problemas para os quais modelos matemáticos são difíceis ou impossíveis de serem obtidos. A inspiração dos procedimentos aplicados neste tipo de algoritmo vem do princípio Darwiniano de seleção natural (Goldberg, 1999), onde as soluções são codificadas, submetidas a processos de reprodução, e avaliadas para determinar seu desempenho.

O procedimento completo que deve ser realizado é apresentado na figura 29.

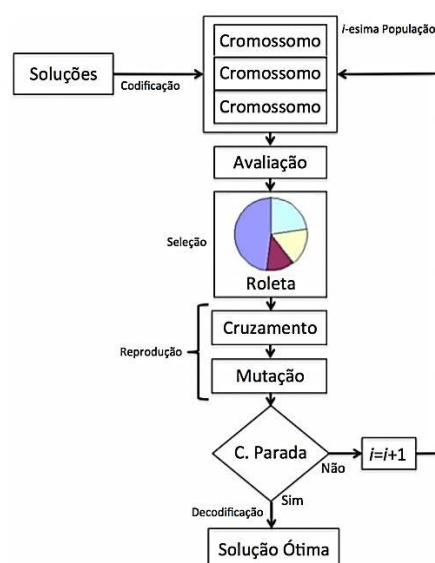


Figura 29. Fluxograma do Algoritmo Genético.

Nas próximas seções serão apresentados os métodos empregados para cada etapa do Algoritmo Genético usado para otimizar a topologia da *ESN*.

4.2.1.Representação e Codificação

A representação é fundamental na modelagem de um Algoritmo Genético, pois nesta etapa é definida a estrutura do cromossomo, com os respectivos genes que o compõem, de maneira que este seja capaz de descrever todas as variáveis e condições do problema dentro do espaço de busca (Velasco, 2014). Tem-se diferentes tipos de representação e cada uma é utilizada em função do problema abordado. No caso de otimização de parâmetros para redes *ESNs*, os possíveis valores que cada um dos genes do cromossomo pode assumir são valores numéricos. Outra possível opção é a utilização de uma codificação binária, mas, devido à grande quantidade de bits necessários para representar cada parâmetro, o espaço de busca definido pela codificação seria grande demais. Segundo (Michelewicz, 1996), a representação binária apresenta desempenho pobre quando aplicada a problemas com alta dimensionalidade.

Na representação numérica, o tamanho do cromossomo é definido pelo número de parâmetros a serem otimizados. Para o problema abordado, o primeiro gene representa o tamanho da Janela, o segundo o número de unidades no *Reservoir*, o terceiro gene o percentual de conexão, o quarto gene representa o Raio Espectral e os dois últimos genes representam o fator de escala nos pesos das conexões de entrada e nos pesos das conexões recorrentes, respectivamente. Na figura 30 é apresentada a composição do cromossomo, onde *Jan* corresponde ao tamanho da Janela, *#U* é o número de unidades no *Reservoir*, *%C* corresponde ao percentual de conexão, $\rho(W)$ é Raio Espectral e sW^e e sW^{rr} são os fatores de escala dos pesos das conexões de entrada e conexões recorrentes respectivamente.

<i>Jan</i>	<i>#U</i>	<i>%C</i>	$\rho(W)$	sW^e	sW^{rr}
------------	-----------	-----------	-----------	--------	-----------

Figura 30. Composição do Cromossomo.

4.2.2.Inicialização e Seleção

O processo de inicialização da população gera soluções de forma aleatória, considerando que, como qualquer processo evolutivo, é necessário que exista diversidade, por isso é importante que a população inicial cubra a maior área possível do espaço de busca.

No processo de seleção podem ser aplicados diversos métodos, alguns oferecem vantagens dependendo da natureza do problema. A seguir é apresentada a informação referente ao método utilizado neste trabalho, que é seleção por roleta, se destacando a vantagem que oferece a existência de uma tendência a selecionar os indivíduos mais aptos segundo (Mitchell, 1998).

- Roleta: os indivíduos são selecionados de acordo com a sua aptidão normalizada em relação à função objetivo. A cada indivíduo associa-se uma probabilidade de seleção proporcional à sua aptidão. Portanto, os indivíduos com melhor aptidão têm maiores chances de serem escolhidos; já os indivíduos com menor aptidão têm menor probabilidade de serem selecionados.

4.2.3.Reprodução

O procedimento de reprodução, que permite a geração de novos indivíduos imitando o processo natural, está dividido em 3 estágios: o primeiro é o acasalamento, onde pares de indivíduos são selecionados; o segundo e terceiro estágio, conhecidos como cruzamento e mutação, realizam a recombinação e modificação dos pares de indivíduos selecionados, gerando novos elementos dentro da população. O código genético é, portanto, passado dos pais para os filhos nesta etapa. Existem diferentes métodos, ou operadores, amplamente estudados e aplicados a diferentes problemas na literatura (Lacerda, 2008). A escolha destes operadores está relacionada com o tipo de representação e codificação.

O operador de cruzamento selecionado para representação numérica é o cruzamento de um ponto. Na figura 31 é apresentado um exemplo do procedimento realizado neste operador.

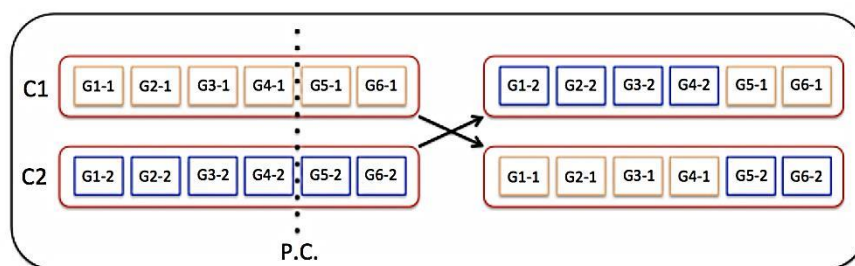


Figura 31. Exemplo de cruzamento de um ponto.

O Ponto de cruzamento (P.C.) também é selecionado aleatoriamente, e posteriormente é realizada a recombinação para serem criados os filhos que contêm parte do código genético do pai (C1) e da mãe (C2).

O método adotado para o operador de mutação é o uniforme, onde é sorteado aleatoriamente um gene do cromossomo e este é alterado para outro valor aleatório dentro dos limites definidos para cada variável. No caso desta tese, os valores limites para os genes apresentados na Figura 30 são aqueles apresentados na Tabela 4.

4.2.4. Função Aptidão

Como função aptidão do algoritmo genético é utilizada uma métrica de erro para avaliar o desempenho da rede neural ESN configurada pelo cromossomo específico. Neste trabalho foi adotado o *sMAPE* obtido do conjunto de validação, por ser a métrica utilizada no estudo de casos. Entretanto, qualquer outra métrica de erro pode ser utilizada, como o *Mean Absolute Percentage Error* (MAPE) ou o *Root Mean Squared Error* (RMSE). A equação 4.1 apresenta a métrica de erro do presente trabalho.

$$sMAPE_{Validação} (\%) = \frac{1}{N} \sum_{t=1}^N \frac{\|y_t - y'_t\|}{(y_t + y'_t)/2} * 100 \quad (4.1)$$

Onde, y_t é o valor real, y'_t é o valor previsto e N é o número de observações.

4.3. Separation Ratio Graph

O modelo *SRG* proposto em (Gibbons, 2010) foi aplicado a tarefas de classificação como apresentado no Capítulo 2, no presente trabalho se realiza uma extensão deste modelo para tarefas de previsão *multi-step* de series temporais adotando os procedimentos do modelo original.

A utilização deste modelo permite avaliar se os estados (*Echoes*) gerados pelo *Reservoir*, durante a apresentação dos padrões, são interpretados de forma correta, i.e. entradas similares(vizinhos mais próximos) geram respostas(estados) similares.

Se a razão entre a distância das respostas com relação à distância das entradas for perto de 1, pode-se interpretar que o *Reservoir* tem uma resposta ideal.

O *Reservoir* apresenta comportamentos indesejáveis em dois casos: o comportamento está na parte caótica (entradas similares-respostas diferentes) ou comportamento de um atrator (entradas diferentes-respostas iguais).

Esse módulo, portanto, é utilizado para escolher o *Reservoir* com comportamento mais próximo do ideal, dentre n inicializações possíveis para a ESN com parâmetros gerais otimizados gerada pelo módulo de Neuro-Evolução..

Como foi apresentado no capítulo 2, é realizado um agrupamento em pares para todo o conjunto de padrões, isto é, mediante um algoritmo de agrupamento e vizinhança são

selecionados pares de vizinhos mais próximos e apresentados ao *Reservoir*, após este processo são coletados os estados resultantes de cada par de padrões apresentados.

A equação 4.2 é utilizada para identificar qual é o melhor *Reservoir*, desta forma é obtido o coeficiente SRG para cada *Reservoir* gerado, e finalmente é selecionado como melhor aquele que tenha o valor do coeficiente SRG mais baixo.

$$SRG = \sum_{p=1}^N |||dif_p(u)|| - ||dif_p(x)||| \quad (4.2)$$

Onde $dif_p(u)$ é a distancia entre o dois de padrões que conformam o par p , $dif_p(x)$ é a distancia entre os dois estados gerados pelos padrões que conformam o par p e N é o número total de pares de padrões gerados pelo algoritmo de agrupamento e vizinhança.

Para interpretar os resultados obtidos por cada *Reservoir*, é realizada uma regressão linear utilizando as distâncias dos vetores de entradas e as distâncias dos estados gerados. A melhor configuração é obtida a partir da reta gerada após a regressão, onde os parâmetros que compõem a equação da reta são avaliados.

Segundo a equação 4.3 os valores a serem observados são m que é o coeficiente angular e b que é o ponto de corte do eixo y .

$$y = mx + b \quad (4.3)$$

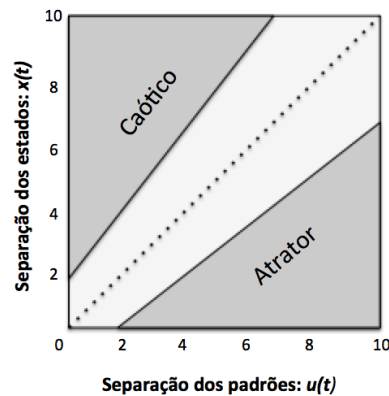


Figura 32. Região objetivo e Zonas caóticas e de atratores.

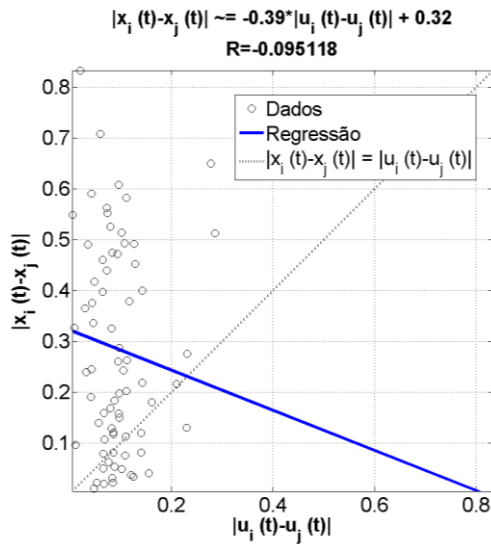
Para que o *Reservoir* apresente um comportamento ideal, a reta resultante da regressão linear deve estar sobre a linha pontilhada segundo a figura 32, isto é, $m = 1$ e $b = 0$. Portanto, o critério de seleção do melhor *Reservoir* é através da regressão que resulte em m mais próximo 1 e b mais próximo de 0.

Adicionalmente se tem o coeficiente de correlação R calculado através da equação 4.4, que determina a relação linear entre a distância dos vetores de entrada e a distância dos estados gerados.

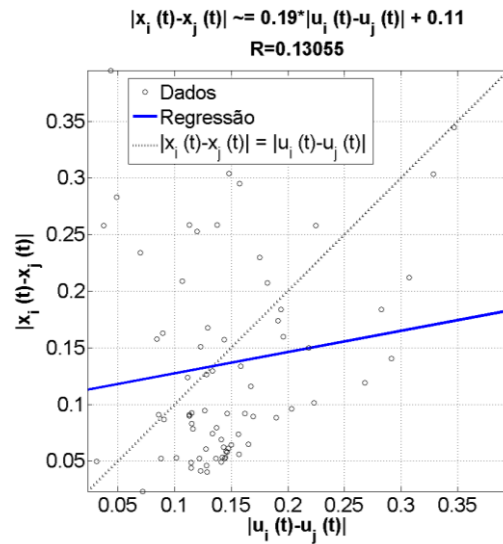
$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum (x_i - \bar{x})^2)(\sum (y_i - \bar{y})^2)}} \quad (4.4)$$

Na Figura 33 são apresentados 4 exemplos de regressão obtida para 4 *Reservoir* diferentes, no eixo vertical se tem a distância de entre as entradas $|x_i(t) - x_j(t)|$ (padrões) e no eixo horizontal estão as distancias $|u(t) - u(t)|$ entre os estados gerados.

É possível observar no exemplo da figura 33.**a**, que o comportamento do *Reservoir* não é o esperado, pois ele apresenta um coeficiente de correlação negativo. No exemplo da figura 33.**b**, o resultado tem um coeficiente de correlação positivo, mas a dispersão dos dados é grande. Já no exemplo da figura 33.**c**, o coeficiente de correlação está muito perto do valor ideal, mas, a posição dos dados está longe da linha pontilhada a qual indica o comportamento ideal. Finalmente, no exemplo da figura 33.**d**, tem-se o melhor comportamento dos 4 exemplos apresentados. Neste caso, o valor do coeficiente angular não é o melhor (mais próximo de 1) dos 4 exemplos apresentados, além o ponto de corte do eixo y não é 0, mas, é possível observar que a reta resultante da regressão realizada fica mais perto da linha pontilhada, por tanto esta configuração do *Reservoir* apresenta o melhor comportamento.



a)



b)

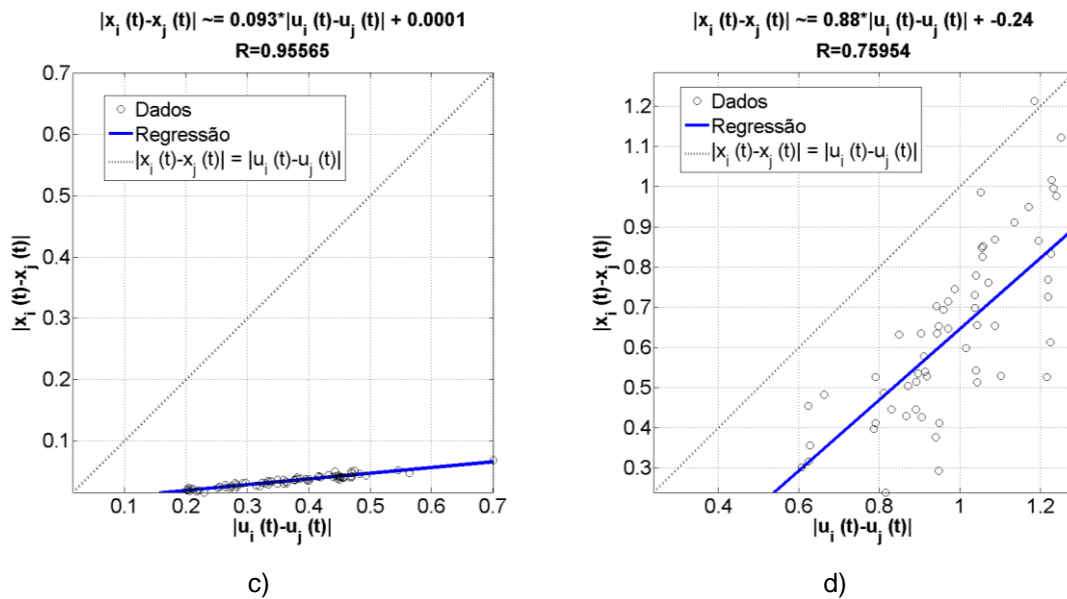


Figura 33. Exemplos de 4 Regressões obtidas entre os padrões e estados.

No caso da figura 33.a, o comportamento do *Reservoir* se encontra na zona caótica, isto é, para vetores de entrada similares (distância de separação pequena) a resposta do *Reservoir* apresenta estados diferentes (distância de separação grande).

Adicionalmente pode-se concluir no caso do exemplo da figura 33.c, que o *Reservoir* está na zona de atratores, e tem um coeficiente de regressão próximo de 1, lembrando que a teoria especifica que, no caso de atratores, a resposta do *Reservoir* tem estados muito parecidos (distância de separação pequena), quando apresentados vetores de entrada diferentes (distância de separação grande), mas a reta resultante da regressão encontra-se longe da linha pontilhada que representa o comportamento ideal.

A seguir serão apresentados conceitos adicionais à operação do SRG, que foram utilizados na seleção da melhor topologia do *Reservoir*, e o procedimento de geração e seleção dos estados utilizados no modelo SRG.

4.3.1. Topologias de Conexão

Para a interpretação das topologias de conexão no *Reservoir* é utilizada uma Matriz de tamanho $N \times N$, sendo N o número de unidades. Valores diferentes de zero, em qualquer posição da matriz, representam os pesos das conexões internas do *Reservoir*, e as posições com valores iguais a zero constituem ausência de conexão.

Na equação 4.5 é apresentado um exemplo da interpretação utilizada para as conexões e pesos no *Reservoir*.

$$W^R = \begin{pmatrix} W_{(1,1)}^R & \dots & W_{(1,N)}^R \\ \vdots & \ddots & \vdots \\ W_{(N,1)}^R & \dots & W_{(N,N)}^R \end{pmatrix} \quad (4.5)$$

Essa é uma matriz simétrica, dado que os valores dos pesos entre um neurônio i e um neurônio j são iguais ($W_{(j,i)}^R = W_{(i,j)}^R$); a diagonal da matriz representa a conexão do neurônio com ele mesmo, onde i é índice da linha e j é o índice da coluna.

4.3.2. Agrupamento e Vizinhaça

Como mencionado anteriormente o agrupamento e vizinhaça dos padrões é necessário, pois os padrões devem ser agrupados em pares para assim serem apresentados ao *Reservoir* e posteriormente coletados os estados gerados.

No modelo proposto o agrupamento e vizinhaça é realizado utilizando o algoritmo *k-vizinhos mais próximos* (*k-nearest neighbors – knn*), onde a distância euclidiana é utilizada como métrica de proximidade. A equação 4.6 apresenta um exemplo para dois padrões i e j .

$$d_{i,j}(u_i, u_j) = d_{j,i}(u_j, u_i) = \sqrt{(u_i - u_j)^2} \quad (4.6)$$

No modelo proposto k é igual a 1, pois somente é procurado 1 vizinho mais próximo para cada padrão, isto é, para cada padrão são obtidas as distancias com todos os padrões disponíveis, posteriormente é selecionado como vizinho mais próximo o padrão que obtenha o menor valor de distância, na tabela 25 é apresentado um exemplo da matriz de distancias para 3 padrões.

Tabela 25. Matriz de distancias entre padrões.

d	u_1	u_2	u_3
u_1	0	$\sqrt{(u_1 - u_2)^2}$	$\sqrt{(u_1 - u_3)^2}$
u_2	$\sqrt{(u_2 - u_1)^2}$	0	$\sqrt{(u_2 - u_3)^2}$
u_3	$\sqrt{(u_3 - u_1)^2}$	$\sqrt{(u_3 - u_2)^2}$	0

O procedimento de geração de estados deve cumprir com a condição de que todos os estados x são gerados com um padrão u partindo do mesmo lugar, isto é, não é considerado o estado imediatamente anterior como histórico para gerar o estado atual.

No próximo capítulo serão apresentados os *benchmarks* utilizados para os estudos de casos e os resultados obtidos na previsão *Multi-Step* utilizando o modelo proposto.

5.

Resultados

Este capítulo apresenta os diversos experimentos e testes realizados com o modelo proposto *ESN-GA-SRG*, comparando seus resultados com os modelos *ESNs* e *ESN-GA*. Além da comparação com modelos baseados em *ESN*, também foi feita uma comparação com modelos estatísticos tradicionais (Box, 1976) e baseados em comitês de previsores, descritos em (Valle, 2012) para previsão *Multi-Step*. Finalmente, apresenta-se uma classificação comparativa dos resultados obtidos pelo modelo proposto com os resultados oficiais da *NN3 Forecasting Competition*.

Como métrica de avaliação de desempenho é utilizado o *sMAPE* para previsão dos conjuntos de dados fora da amostra. A escolha desta métrica deveu-se ao fato de ser a utilizada na *NN3 Forecasting Competition*.

5.1.

Estudo de Casos

Este trabalho utilizou, como estudo de casos, a versão reduzida (11 séries) da ***NN3 Forecasting Competition for Neural Networks & Computational Intelligence***, e 36 séries da ***M3 Forecasting Competition***, selecionadas segundo: amostragem, assimetria, sazonalidade e estacionaridade. Algumas das razões desta escolha são:

1. Competição oficial de previsão *Multi-Step*, de nível mundial, com participação dos principais institutos de pesquisa;
2. Composta por séries temporais diversas, que permitem avaliar, de forma completa, o desempenho de modelos propostos para tais fins;
3. Acesso à relação de resultados finais da competição para todas as séries propostas;
4. Participação de um dos mais importantes pesquisadores na área de *Reservoir Computing*, o Prof. Herbert Jaeger, que se classificou na segunda posição global dentre os resultados para a versão reduzida do *NN3 Forecasting Competition*.

Nas tabelas 26 a 29 são apresentadas informações das séries utilizadas, como o tamanho da base total, a quantidade de dados utilizados para treino e os tamanhos dos conjuntos de validação e teste, os quais foram definidos em função do regulamento da

competição. Além disso, as tabelas indicam o coeficiente de assimetria que foi apresentado no Capítulo 3.

Tabela 26. Informações gerais das 11 séries NN3 utilizadas.

Série	Dados totais	Conjunto de treino	Conjunto de Validação	Conjunto de Teste	Assimetria
101	144	126	18	18	-0.0985
102	144	126	18	18	-0.1476
103	144	126	18	18	1.3085
104	133	115	18	18	-0.7074
105	144	126	18	18	-0.1508
106	144	126	18	18	0.2233
107	144	126	18	18	0.2815
108	133	115	18	18	1.7009
109	141	123	18	18	1.3751
110	144	126	18	18	4.4008
111	144	126	18	18	0.8792

Tabela 27. Informações gerais das 12 séries M3 com amostragem anual.

Amostragem Anual					
Série	Dados totais	Conjunto de treino	Conjunto de Validação	Conjunto de Teste	Assimetria
33	20	8	6	6	-1,4141
50	20	8	6	6	-0,6796
70	20	8	6	6	-1,2893
105	20	8	6	6	0,6033
119	20	8	6	6	7,7543
162	21	9	6	6	-0,6440
171	47	35	6	6	-0,0091
200	43	31	6	6	104525
351	20	8	6	6	-1,3609
564	25	13	6	6	2,8818
599	25	13	6	6	-1,1126
642	36	24	6	6	-0,9847

Tabela 28. Informações gerais das 12 séries M3 com amostragem trimestral.

Amostragem trimestral					
Série	Dados totais	Conjunto de treino	Conjunto de Validação	Conjunto de Teste	Assimetria
671	44	28	8	8	0,0443
677	44	28	8	8	-1,1469
692	44	28	8	8	-0,0068
713	44	28	8	8	0,1067
730	43	27	8	8	-1,0847
758	44	28	8	8	-0,8037
773	44	28	8	8	-0,9884
829	44	28	8	8	-0,3549
847	43	27	8	8	23,3153
854	59	43	8	8	-0,6452
865	64	48	8	8	-1,0614
896	56	40	8	8	-0,7616

Tabela 29. Informações gerais das 12 séries M3 com amostragem mensal.

Amostragem Mensal					
Série	Dados totais	Conjunto de treino	Conjunto de Validação	Conjunto de Teste	Assimetria
1430	69	33	18	18	1,0424
1460	69	33	18	18	2,1892
1725	126	110	18	18	0,2113
1900	140	124	18	18	0,0985
1906	134	118	18	18	-0,8738
1912	144	128	18	18	-1,0965
1996	144	128	18	18	-0,2653
2001	144	128	18	18	-0,4481
2013	144	128	18	18	-0,9488
2016	144	128	18	18	-0,6650
2089	144	128	18	18	-1,0064
2579	134	118	18	18	-0,9716

Conforme mencionado, utiliza-se o “*Symmetric mean absolute percentage error*” (*sMAPE*) como métrica de avaliação de desempenho das previsões realizadas pelo modelo *ESN-GA-SRG*. Esta métrica foi a escolhida para coincidir com a métrica utilizada na competição *NN3*. Os valores de erro são obtidos utilizando a equação 5.1.

$$sMAPE (\%) = \frac{1}{N} \sum_{t=1}^N \frac{|y_t - y'_t|}{(y_t + y'_t)/2} * 100 \quad (5.1)$$

Onde, N é o número total de valores previstos, y_t é o valor real e y'_t é o valor previsto.

A tabela 30 apresenta os limites utilizados na configuração do cromossomo, assim como a precisão utilizada para cada parâmetro.

Tabela 30. Limites dos parâmetros do Cromossomo.

Parâmetro	Símbolo	Valor Mínimo	Valor Máximo	Precisão
Janela	Jan	1	36	1
Número de Unidades	$\#U$	100	2000	1
Percentual de Conexão	$\%C$	20	100	1
Raio Espectral	$\rho(W)$	0.1	1	0.1
Fator de Escala nos pesos das Conexões de entrada	sW^e	0	1	0.1
Fator de Escala nos pesos das Conexões recorrentes	sW^{rr}	0	1	0.1

5.2.

Resultados para a versão reduzida NN3

A seguir são apresentados os resultados obtidos nos experimentos realizados com o modelo proposto *ESN-GA-SRG* e a sua comparação com os modelos *ESN* clássico e *ESN-GA*. Na tabela 31 é apresentada a média dos *sMAPE* para os conjuntos de teste das 11 séries da competição, em cada um dos modelos.

Tabela 31. Resultados comparativos para as séries NN3 com os modelos utilizados.

Método	<i>sMAPE</i> _Teste
<i>ESN</i>	15,95
<i>ESN-GA</i>	14,55
<i>ESN-GA-SRG</i>	11,51

A tabela 32 apresenta os resultados de *sMAPE* para cada uma das 11 séries, permitindo comparar os resultados obtidos nas 11 séries individualmente.

Tabela 32. Resultados comparativos para cada série NN3.

Série	<i>sMAPE</i>								
	<i>ESN</i>			<i>ESN-GA</i>			<i>ESN-GA-SRG</i>		
	Treino	Val.	Teste	Treino	Val.	Teste	Treino	Val.	Teste
1	2,96	3,32	3,71	2,89	3,12	3,15	2,15	2,39	2,82
2	13,43	14,57	14,82	12,65	12,89	13,14	8,70	9,41	9,52
3	34,71	34,98	35,21	25,33	26,03	27,16	25,33	26,03	27,16
4	16,46	17,15	17,24	11,21	11,60	12,32	9,59	10,36	10,87
5	3,68	3,91	4,03	2,88	3,27	3,32	2,31	2,44	2,52
6	3,29	3,38	3,65	2,45	2,64	3,28	1,87	2,04	2,19
7	5,72	6,16	6,49	5,82	6,38	6,56	3,87	4,11	4,14
8	27,57	27,88	28,14	28,85	29,32	29,69	19,57	19,80	20,43
9	11,76	12,13	12,25	10,19	11,31	11,63	6,87	7,33	8,07
10	33,82	34,56	35,16	33,22	33,84	34,44	26,65	26,79	27,79
11	13,67	13,88	14,72	14,35	15,22	15,36	9,54	10,94	11,10

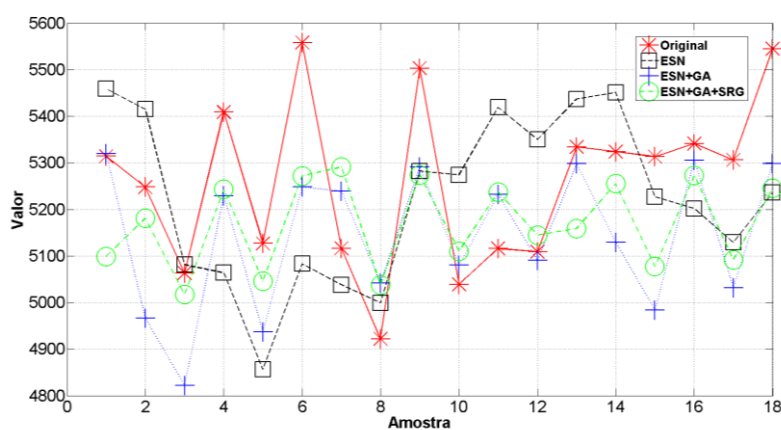
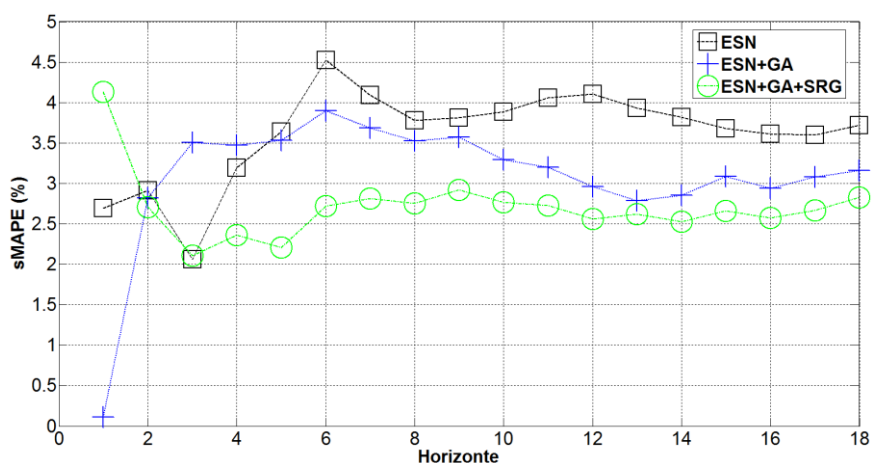
Para cada uma das 11 séries, são apresentados os parâmetros de configuração finais do modelo *ESN-GA-SRG*, os resultados da previsão do conjunto fora da amostra (teste) para os modelos *ESN*, *ESN-GA* e *ESN-GA-SRG*, a evolução do *sMAPE* ao longo do horizonte de previsão (18 passos a frente) e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG.

Na tabela 33 são apresentados os parâmetros de configuração da rede *ESN* para a série NN3-101.

Tabela 33. Parâmetros finais de configuração para a série *NN3-101*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	381
Janela	14
Percentagem de conexão	30%
Raio Espectral	0.6
Fator de escala das conexões de entrada	0.3
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.7

Nas figuras 34 a 36 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do SMAPE a longo do horizonte de previsão e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série *NN3-101*.

Figura 34. Resultados Série *NN3-101* fora da amostra.Figura 35. Evolução do *sMAPE* a longo do horizonte de previsão (Série *NN3-101*).

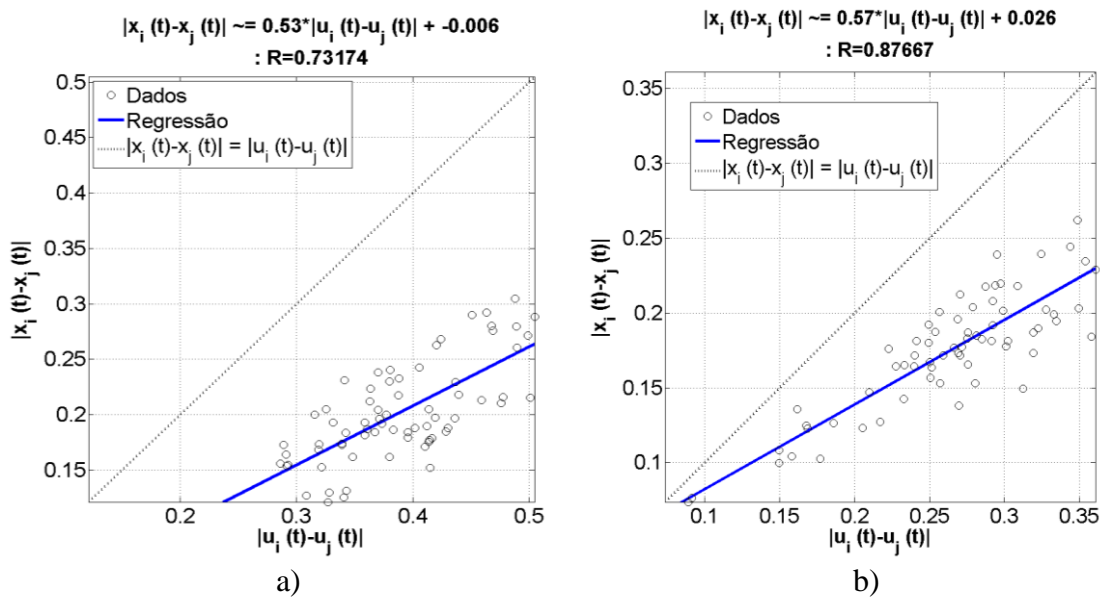


Figura 36. a) Regressão obtida antes do SRG. b) Regressão obtida após do SRG.

Na figura 34 é possível observar que a previsão realizada pelo modelo *ESN-GA-SRG* acompanha melhor a série original, na figura 35 pode ser observado o comportamento do *sMAPE* acumulado, onde ao longo do horizonte de previsão mostra que o comportamento *multi-step* no modelo *ESN-GA-SRG* tem uma melhor evolução do que os outros modelos utilizados, e no caso das regressões lineares apresentadas na figura 36 é possível observar que a reta resultante para o comportamento do *Reservoir* após do método SRG está mais perto da linha pontilhada e o coeficiente de dispersão está mais próximo de 1.

Na tabela 34 são apresentados os parâmetros de configuração da rede ESN para a série NN3-102.

Tabela 34. Parâmetros finais de configuração para a série NN3-102.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	412
Janela	12
Porcentagem de conexão	52%
Raio Espectral	0.5
Fator de escala das conexões de entrada	0.5
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.4

Nas figuras 36 a 38 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do *SMAPE* a longo do horizonte de previsão e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série NN3-102.

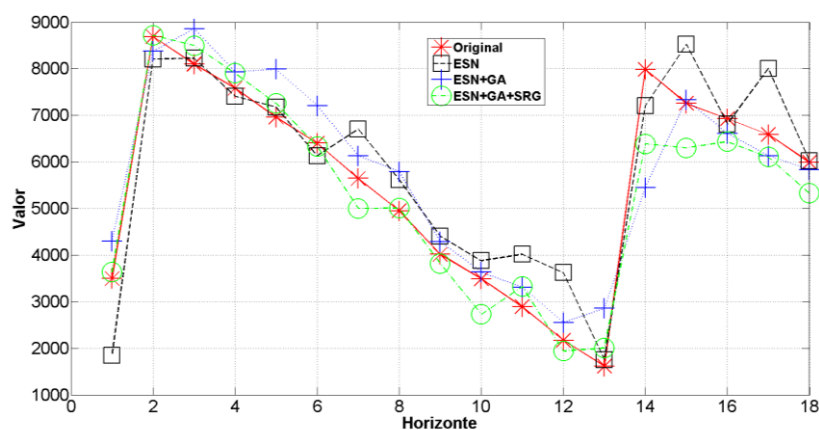


Figura 36. Resultados Série NN3-102 fora da amostra.

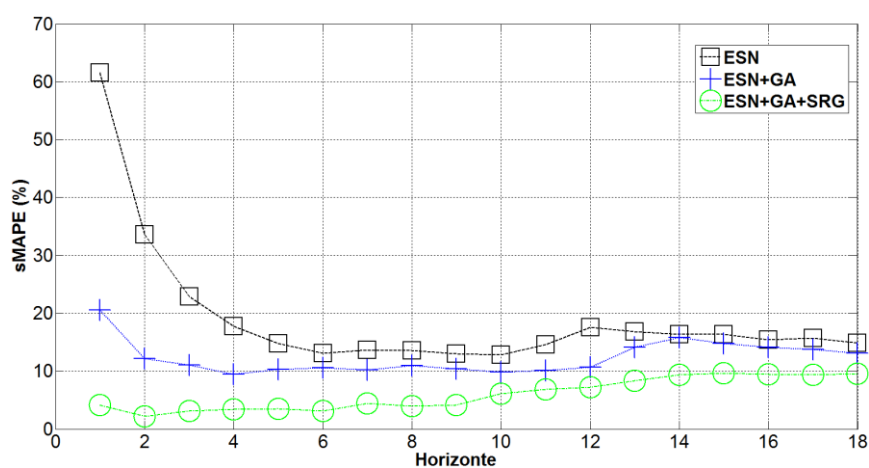
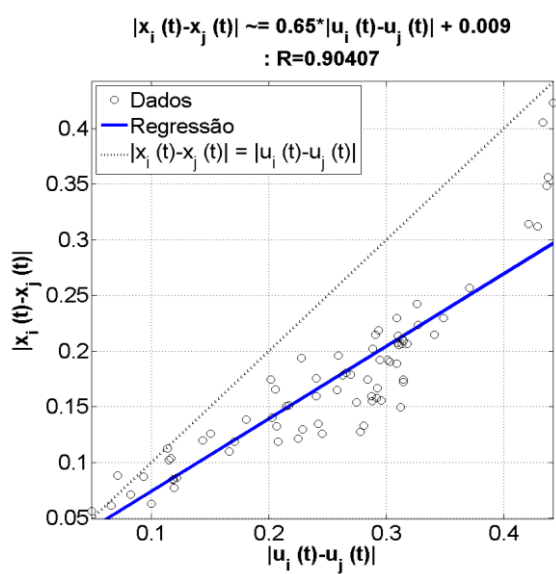
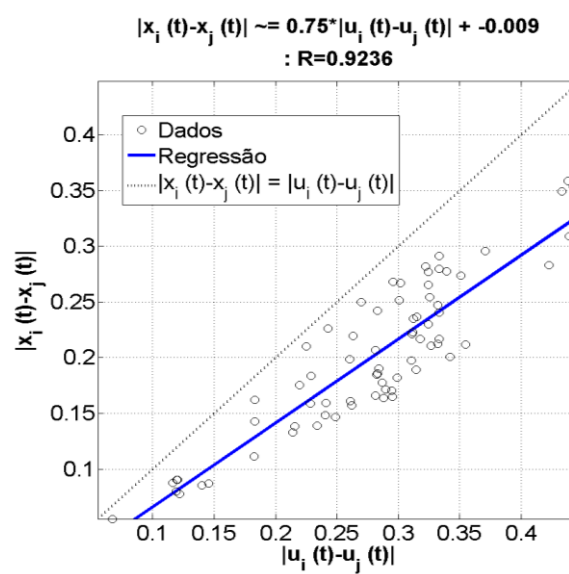


Figura 37. Evolução do sMAPE a longo do horizonte de previsão (Série NN3-102).



a)



b)

Figura 38. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Nas observações que podem ser geradas dos resultados para a Série *NN3-102*, segundo os resultados na figura 36, pode-se concluir que a previsão gerada pelos 3 modelos utilizados conseguiu acompanhar os valores originais, mas, novamente, o modelo *ESN-GA-SRG* conseguiu a melhor previsão. No caso da evolução do *Smape*, os 3 modelos mantiveram a mesma tendência ao longo do horizonte de previsão, isto é, o modelo *ESN-GA-SRG* teve o melhor começo dos 3 modelos comparados e se manteve assim durante todo horizonte.

Na figura 38 é possível observar que os estados gerados mantiveram um comportamento similar nos dois casos (antes e depois do modelo SRG), mais, é possível observar que o coeficiente de dispersão está mais próximo de 1, assim como a reta resultante da regressão está mais próxima da linha pontilhada.

Na tabela 35 são apresentados os parâmetros de configuração da rede ESN para a série *NN3-103*.

Tabela 35. Parâmetros finais de configuração para a série *NN3-103*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	365
Janela	12
Porcentagem de conexão	45%
Raio Espectral	0.4
Fator de escala das conexões de entrada	0.7
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.4

Nas figuras 39 a 41 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do *SMAPE* a longo do horizonte de previsão e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série *NN3-103*.

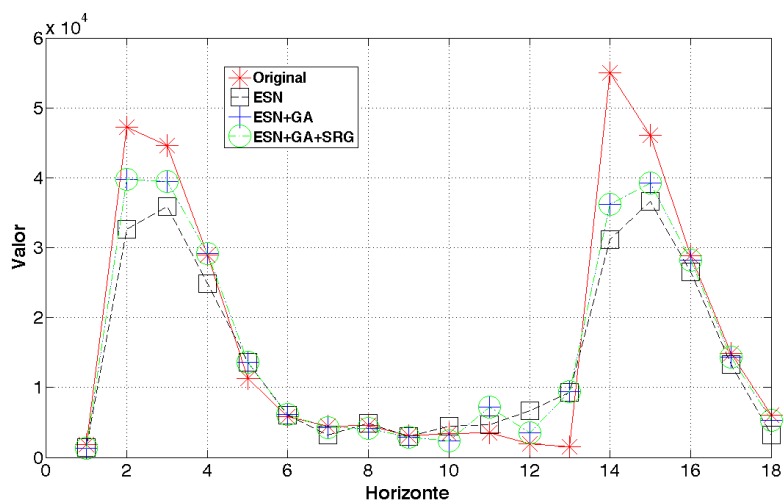


Figura 39. Resultados Série *NN3-103* fora da amostra.

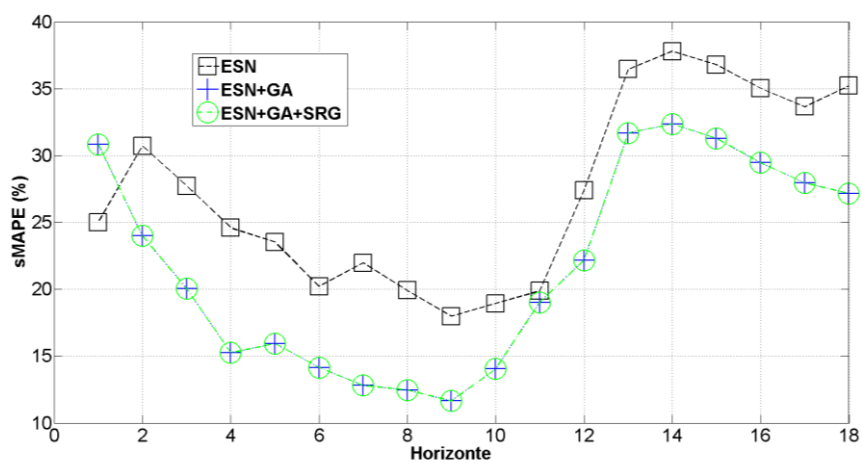


Figura 40. Evolução do *sMAPE* a longo do horizonte de previsão (Série *NN3-103*).

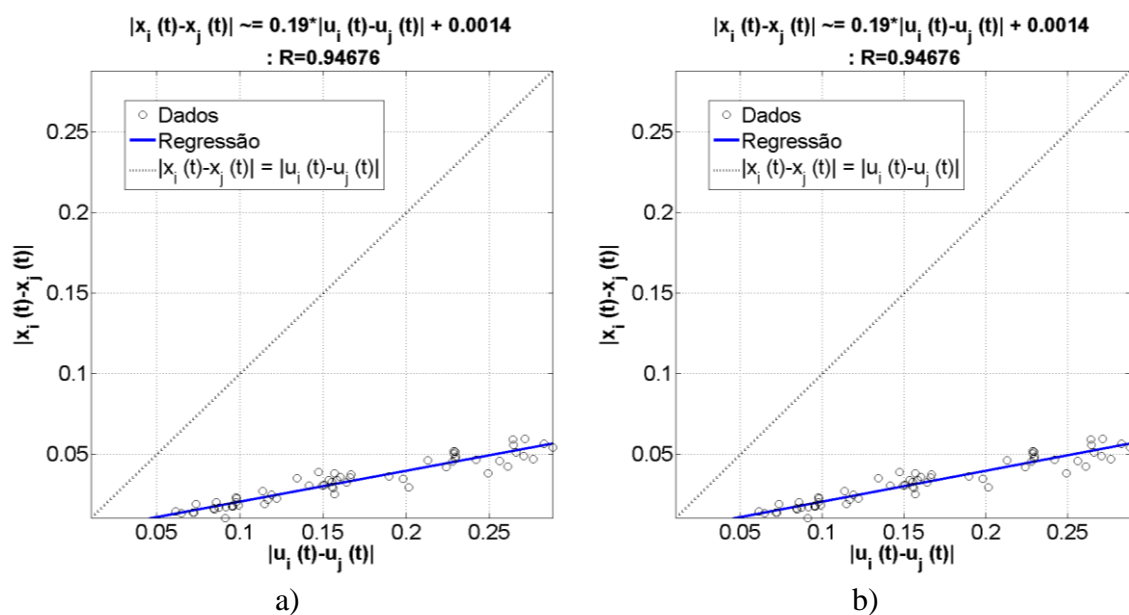


Figura 41. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Nos resultados obtidos para a Série *NN3-103*, o modelo *ESN-GA-SRG* não conseguiu uma configuração melhor para o *Reservoir* do que a obtida pelo GA. Neste caso é conservada a resposta gerada pelo modelo *ESN-GA* como a melhor previsão.

Na evolução do *sMAPE* ao longo do horizonte de previsão é possível observar que o erro teve um aumento significativo exatamente onde o comportamento da série mudou de forma abrupta. Para as regressões obtidas é possível identificar que o comportamento do *Reservoir* teve uma clara tendência para atrator, pois a distância dos estados gerados é pouca em comparação com a distância dos padrões.

Na tabela 36 são apresentados os parâmetros de configuração da rede ESN para a série *NN3-104*.

Tabela 36. Parâmetros finais de configuração para a série *NN3-104*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	637
Janela	9
Porcentagem de conexão	27%
Raio Espectral	0.6
Fator de escala das conexões de entrada	0.5
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.4

Nas figuras 42 a 44 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do *SMAPE* a longo do horizonte de previsão e a regressão

linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série NN3-104.

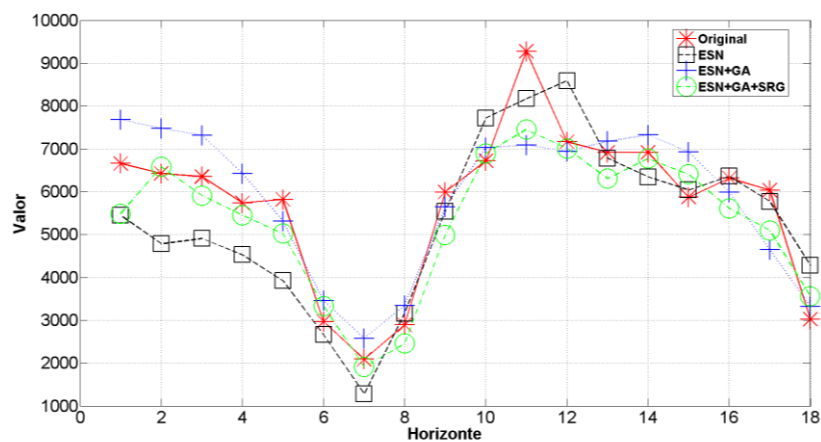


Figura 42. Resultados Série NN3-104 fora da amostra.

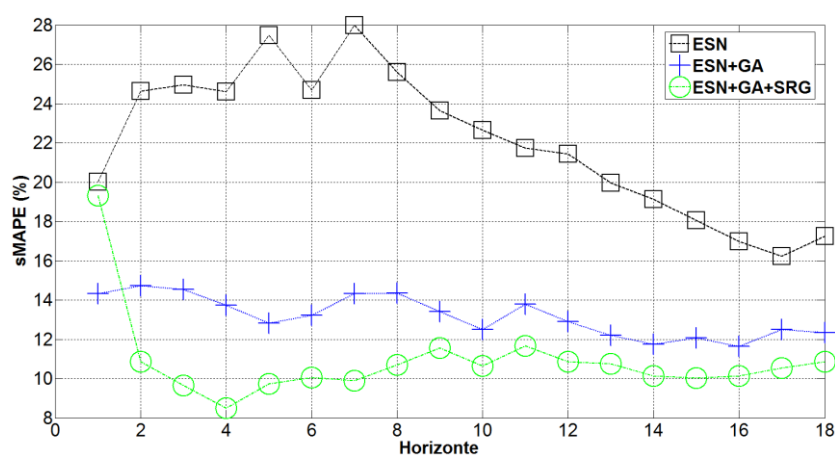


Figura 43. Evolução do sMAPE a longo do horizonte de previsão (Série NN3-104).

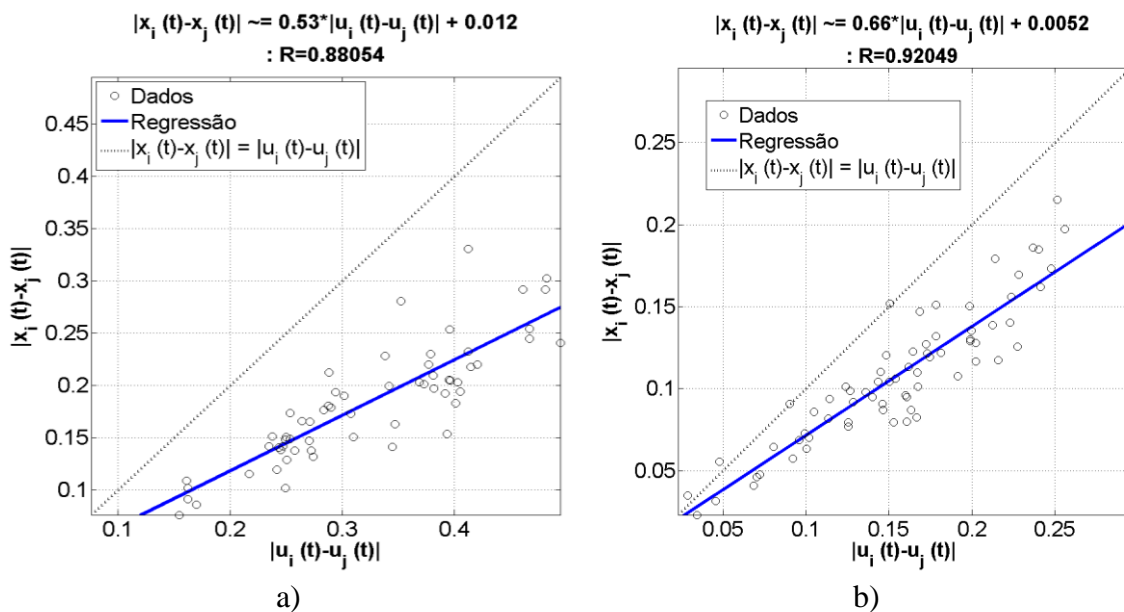


Figura 44. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Os resultados para a Série *NN3-104* apresentaram novamente o modelo *ESN-GA-SRG* como o vencedor. O modelo *ESN-GA* realizou uma previsão aceitável mas, nos primeiros valores da previsão, não teve bons resultados o que acabou prejudicando o *sMAPE* final.

Na evolução do *sMAPE* ao longo do horizonte de previsão é possível observar que o modelo *ESN-GA-SRG* teve um começo não tão bom, mas, conseguiu melhorar ao longo da previsão e manter o menores erros acumulados.

Na figura 44 é possível observar que a distribuição dos estados manteve uma boa relação e o coeficiente de correlação está mais próximo de 1 após de utilizar o modelo SRG, gerando uma dispersão menor nas distancias dos estados.

Na tabela 37 são apresentados os parâmetros de configuração da rede ESN para a série *NN3-105*.

Tabela 37. Parâmetros finais de configuração para a série *NN3-105*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	271
Janela	13
Porcentagem de conexão	65%
Raio Espectral	0.7
Fator de escala das conexões de entrada	0.3
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.8

Nas figuras 45 a 47 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do SMAPE a longo do horizonte de previsão e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série NN3-105.

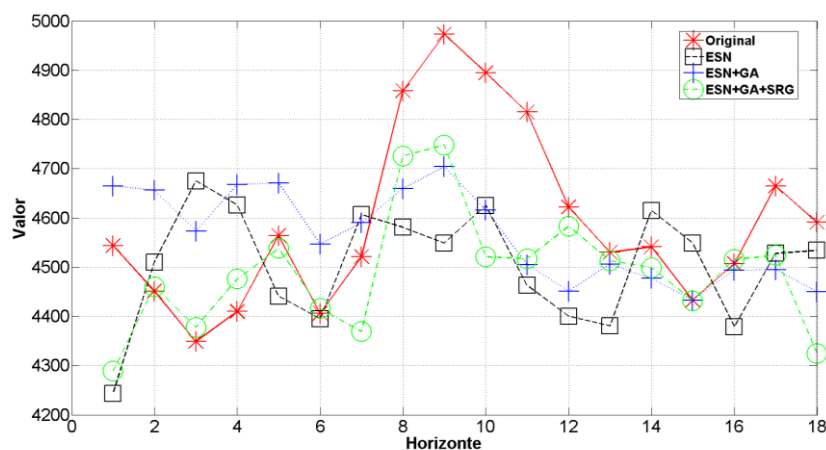


Figura 45. Resultados Série NN3-105 fora da amostra.

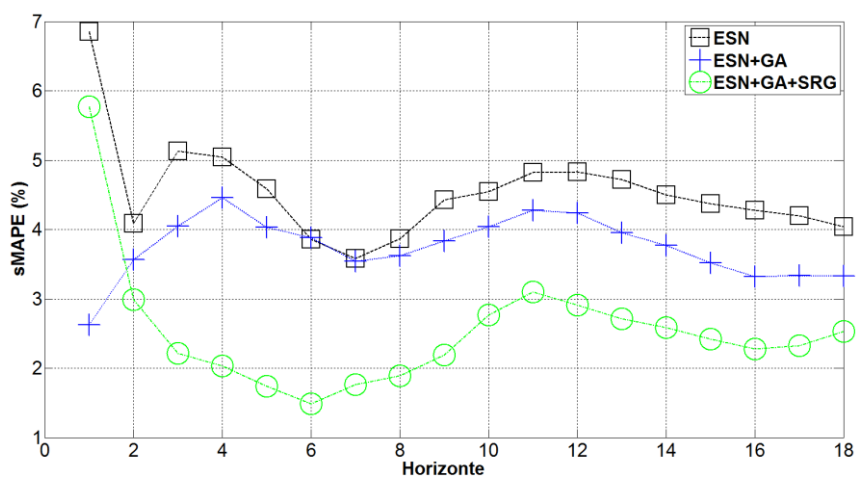


Figura 46. Evolução do sMAPE a longo do horizonte de previsão (Série NN3-105).

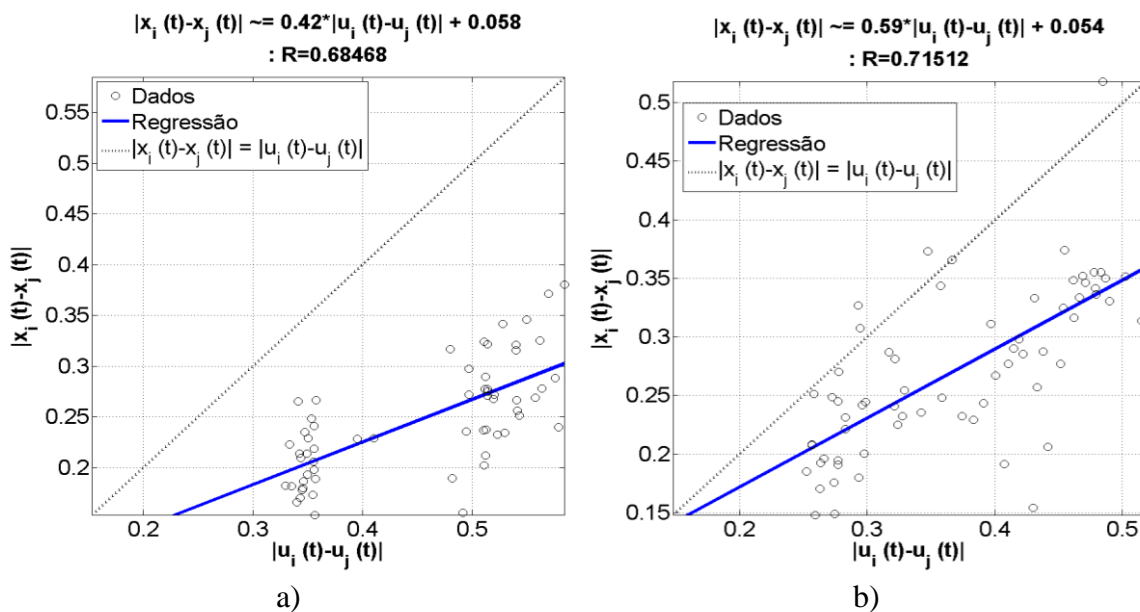


Figura 47. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

O comportamento do modelo *ESN-GA-SRG* na previsão da Série *NN3-105* apresentou uma notória dificuldade para realizar a previsão dos passos 8 a 11 nos três modelos onde pode ser observado o incremento do erro. Na evolução do *sMAPE* ao longo da previsão os 3 modelos apresentaram comportamentos similares, isto é, errando nos mesmos pontos, mas, a diferença está no valor do erro, onde o modelo *ESN-GA-SRG* conseguiu atingir os menores valores de erro ao longo da previsão.

Na regressão linear pode ser observado que a dispersão dos estados é maior quando comparada com os resultados das series anteriores, o que pode ser entendível ao observar o comportamento da previsão para o modelo *ESN-GA-SRG*.

Na tabela 38 são apresentados os parâmetros de configuração da rede ESN para a série *NN3-106*.

Tabela 38. Parâmetros finais de configuração para a série *NN3-106*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	368
Janela	8
Porcentagem de conexão	72%
Raio Espectral	0.8
Fator de escala das conexões de entrada	0.6
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.6

Nas figuras 48 a 50 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do SMAPE a longo do horizonte de previsão e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série NN3-106.

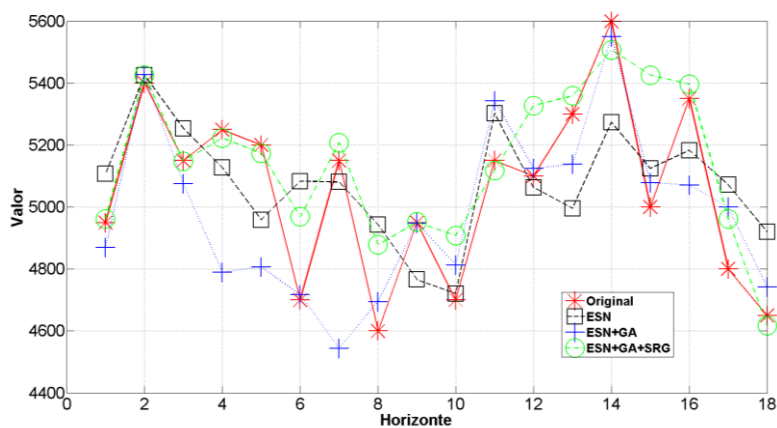


Figura 48. Resultados Série NN3-106 fora da amostra.

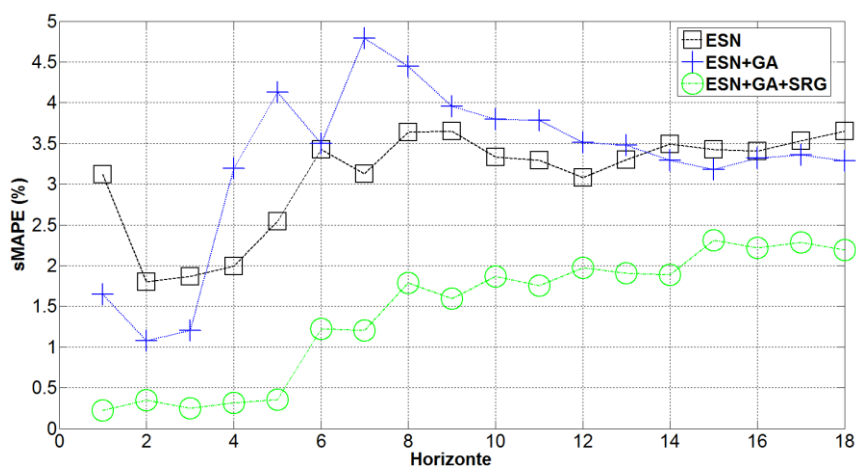


Figura 49. Evolução do sMAPE a longo do horizonte de previsão (Série NN3-106).

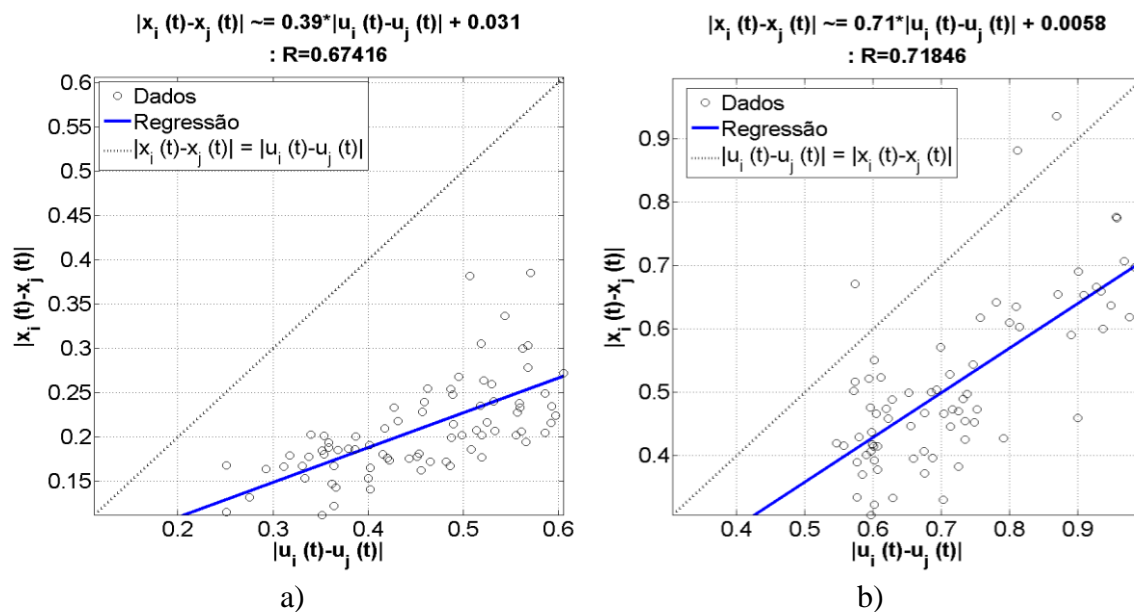


Figura 50. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Nos resultados obtidos é possível observar que o modelo *ESN-GA-SRG* teve o melhor resultado. A Série *NN3-106* tem algumas variações que não foram previstas por nenhum dos modelos utilizados. A evolução do *sMAPE* apresentou um comportamento onde o modelo *ESN-GA-SRG* teve um começo muito bom mas no decorrer da previsão piorou um pouco, mesmo assim conseguiu obter os melhores resultados.

No caso da regressão linear pode ser observado na figura 50 que a dispersão teve uma leve melhora, mais, o coeficiente angular teve uma melhora significativa aproximando a reta resultante na regressão ao comportamento ideal.

Na tabela 39 são apresentados os parâmetros de configuração da rede ESN para a série *NN3-107*.

Tabela 39. Parâmetros finais de configuração para a série *NN3-107*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	405
Janela	10
Porcentagem de conexão	89%
Raio Espectral	0.4
Fator de escala das conexões de entrada	0.7
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.2

Nas figuras 51 a 53 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do *SMAPE* a longo do horizonte de previsão e a regressão

linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série NN3-107.

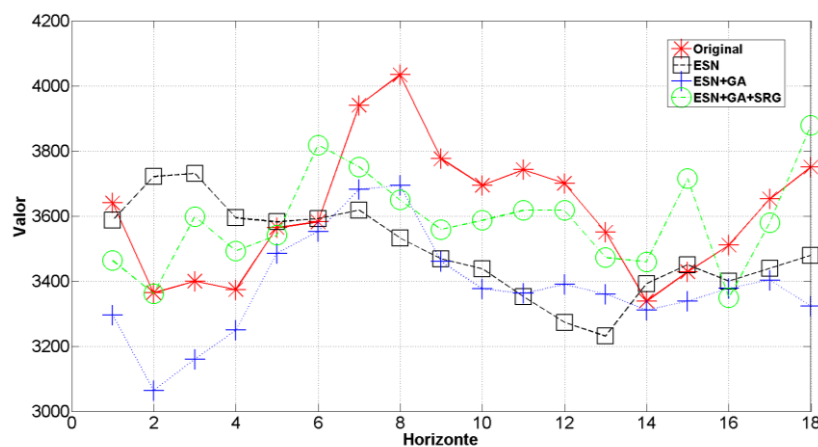


Figura 51. Resultados Série NN3-107 fora da amostra.

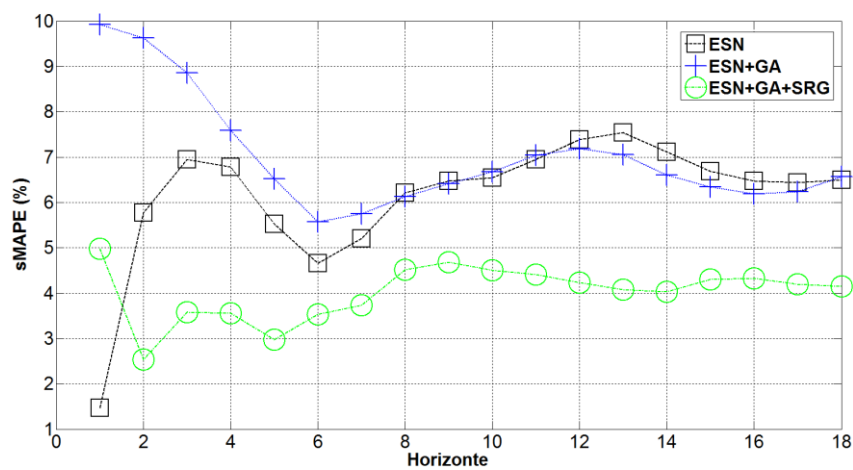
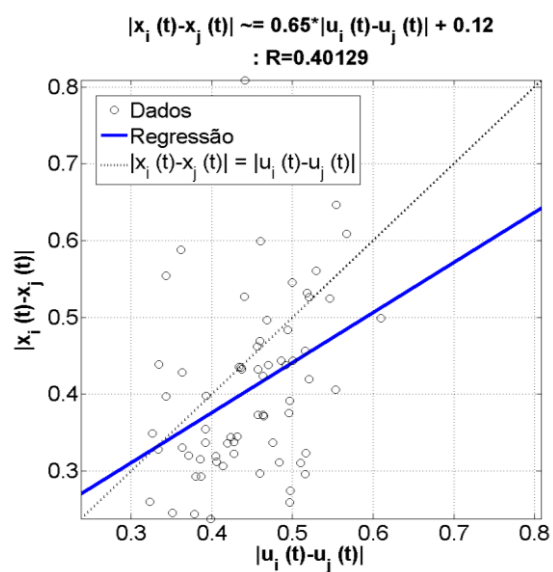
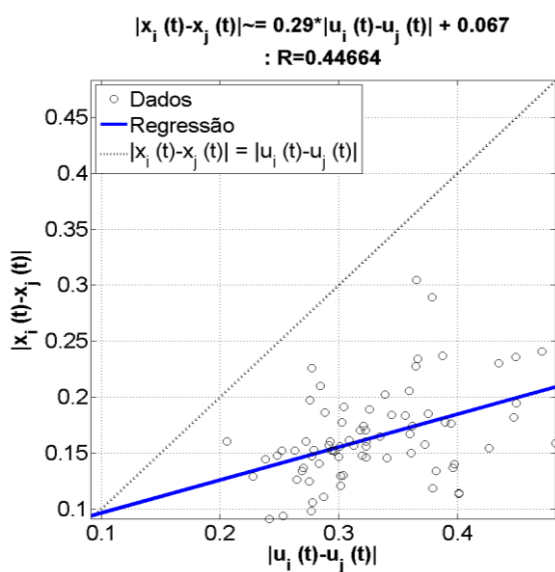


Figura 52. Evolução do sMAPE a longo do horizonte de previsão (Série NN3-107).



a) b)
 Figura 53. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Para a Série *NN3-107* os resultados do modelo *ESN-GA-SRG* foram também os melhores, conseguindo acompanhar melhor os valores originais do que os outros 2 modelos. A evolução do *sMAPE* teve um comportamento particular para o modelo *ESN-GA* pois acabou acompanhado a evolução do modelo *ESN*, no caso do modelo *ESN-GA-SRG* o comportamento do *SMAPE* foi o melhor após dos passo 2 e se manteve durante todo o horizonte de previsão.

No caso da regressão obtida pode-se observar uma dispersão grande, esta dispersão poderia ser a responsável pelo erro em pontos da série com valores máximos onde não foi possível obter previsões acertadas.

Na tabela 40 são apresentados os parâmetros de configuração da rede *ESN* para a série *NN3-108*.

Tabela 40. Parâmetros finais de configuração para a série *NN3-108*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	195
Janela	9
Porcentagem de conexão	34%
Raio Espectral	0.7
Fator de escala das conexões de entrada	0.4
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.5

Nas figuras 54 a 56 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do *SMAPE* a longo do horizonte de previsão e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método *SRG* para a série *NN3-108*.

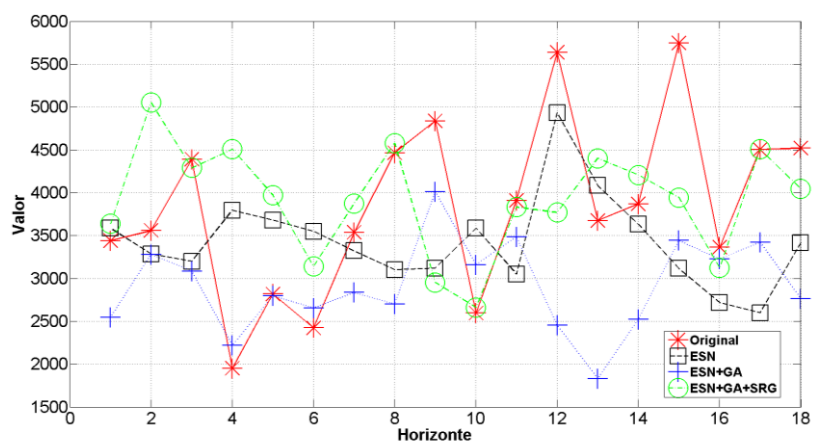


Figura 54. Resultados Série NN3-108 fora da amostra.

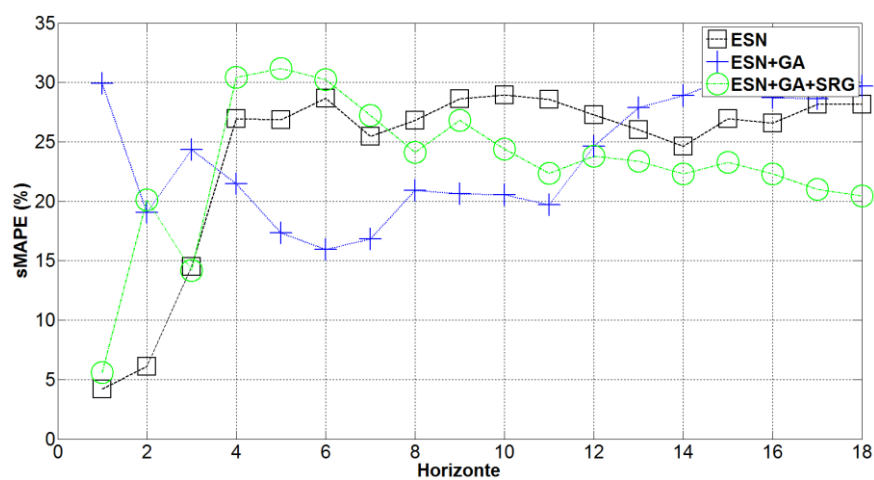


Figura 55. Evolução do sMAPE a longo do horizonte de previsão (Série NN3-108).

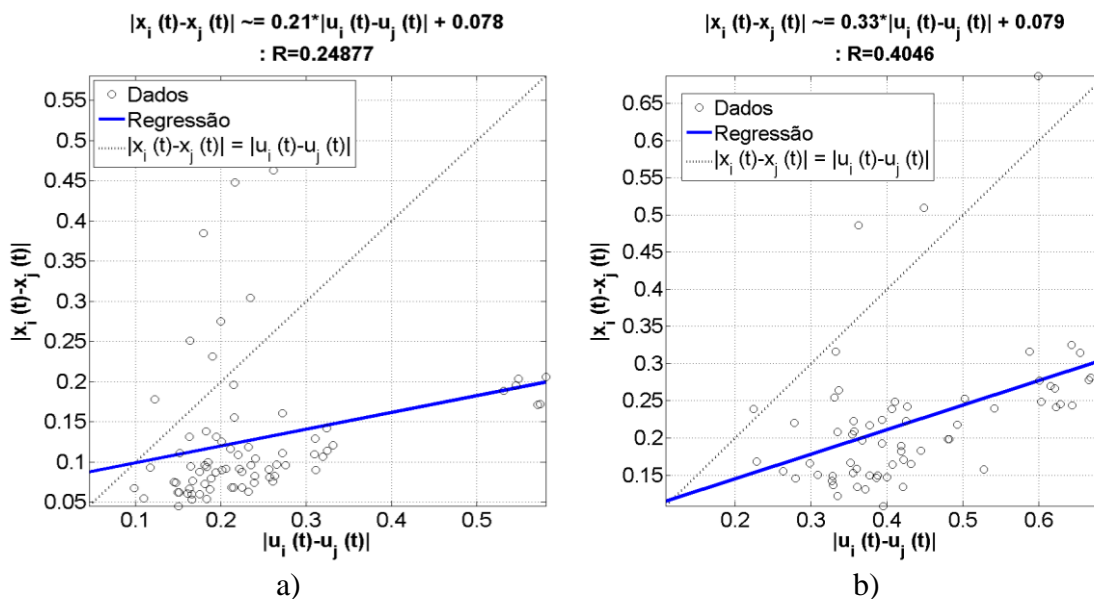


Figura 56. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Os resultados obtidos mostram que para a Série *NN3-108* se teve muita dificuldade no trabalho de realizar a previsão pelos 3 modelos utilizados. No modelo *ESN-GA-SRG* os primeiros valores previstos não conseguiram acompanhar os valores originais e no final, novamente, foi observado que o acompanhamento na previsão foi pobre.

A evolução do *sMAPE* apresentou uma melhora significativa (10%) após do passo 4 na previsão até o final do horizonte, após o passo 12 o modelo *ESN-GA-SRG* conseguiu obter o menor *SMAPE* dos modelos comparados e ao final ser o modelo vencedor. A regressão linear obtida tem uma leve tendência a identificar o comportamento atrator do *Reservoir*, pois a distância dos estados gerados é pouca em comparação com a distância dos padrões.

Na tabela 41 são apresentados os parâmetros de configuração da rede ESN para a série *NN3-109*.

Tabela 41. Parâmetros finais de configuração para a série *NN3-109*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	748
Janela	6
Porcentagem de conexão	23%
Raio Espectral	0.8
Fator de escala das conexões de entrada	0.7
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.1

Nas figuras 57 a 59 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do SMAPE a longo do horizonte de previsão e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série NN3-109.

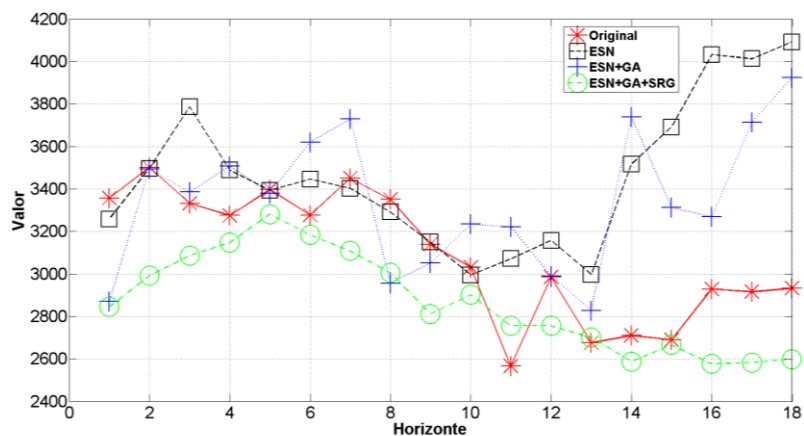


Figura 57. Resultados Série NN3-109 fora da amostra.

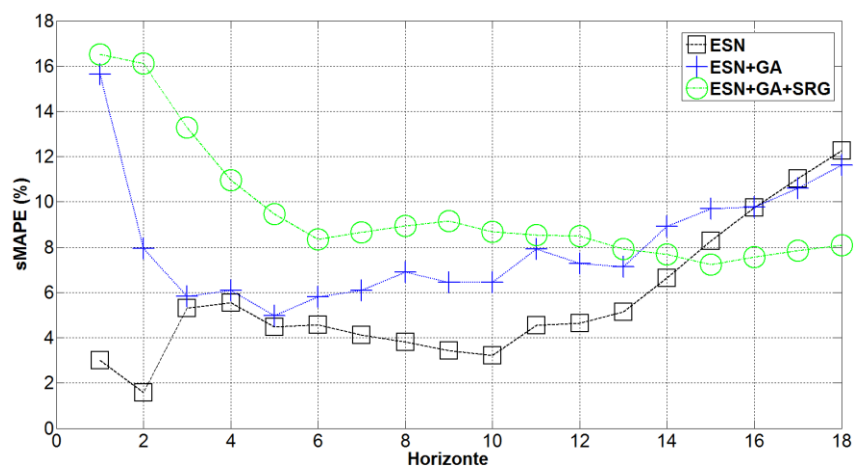


Figura 58. Evolução do sMAPE a longo do horizonte de previsão (Série NN3-109).

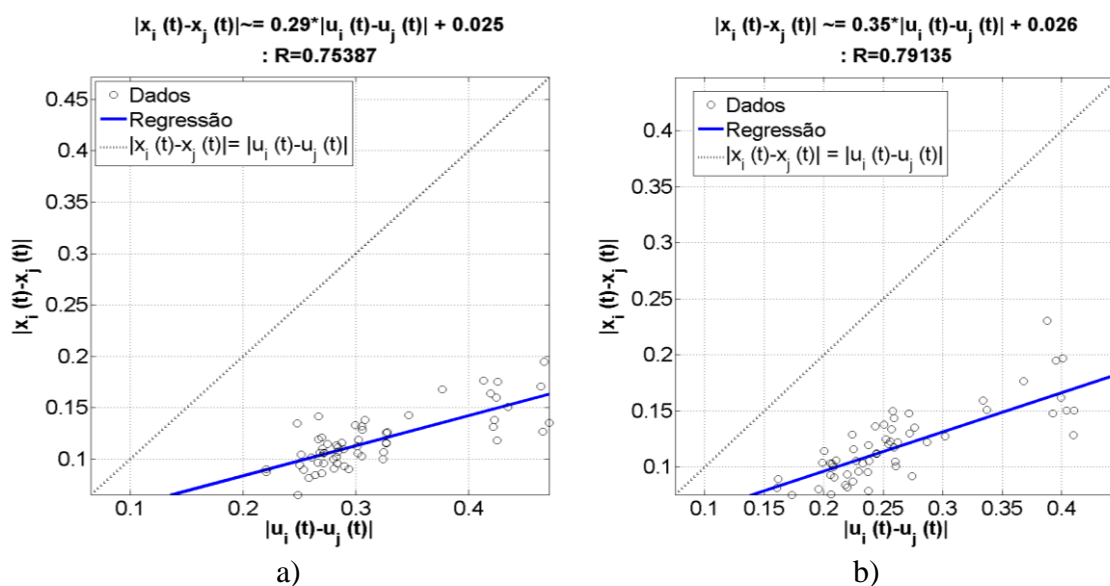


Figura 59. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Nos resultados obtidos para a Série *NN3-109*, o *ESN-GA-SRG* obteve a melhor previsão, sendo que foi o modelo que conseguiu acompanhar os valores originais. Os outros 2 modelos obtiveram resultados razoáveis (*ESN*=12,25 e *ESN+GA*=11.63), embora as previsões dos pontos 13 para frente tenham dado valores que estiveram longe dos valores da série original.

A evolução do *sMAPE* apresenta a vantagem do modelo *ESN-GA-SRG*, o qual conseguiu melhorar o erro obtido nos dois primeiros passos, após do passo 6 se observa que o erro se manteve ao longo da previsão conseguindo melhorar as previsões dos outros modelos a partir do passo 15.

Na regressão linear obtida para o modelo *ESN-GA-SRG* os estados gerados novamente estiveram perto da zona de atratores, mas, com um coeficiente de correlação próximo de 1, no caso do coeficiente angular se teve uma melhora leve que permitiu estar mais perto do comportamento ideal na reta resultante na regressão.

Na tabela 42 são apresentados os parâmetros de configuração da rede ESN para a série *NN3-110*.

Tabela 42. Parâmetros finais de configuração para a série *NN3-110*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	437
Janela	12
Porcentagem de conexão	60%
Raio Espectral	0.4
Fator de escala das conexões de entrada	0.5
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.6

Nas figuras 60 a 62 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do SMAPE a longo do horizonte de previsão e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série NN3-110.

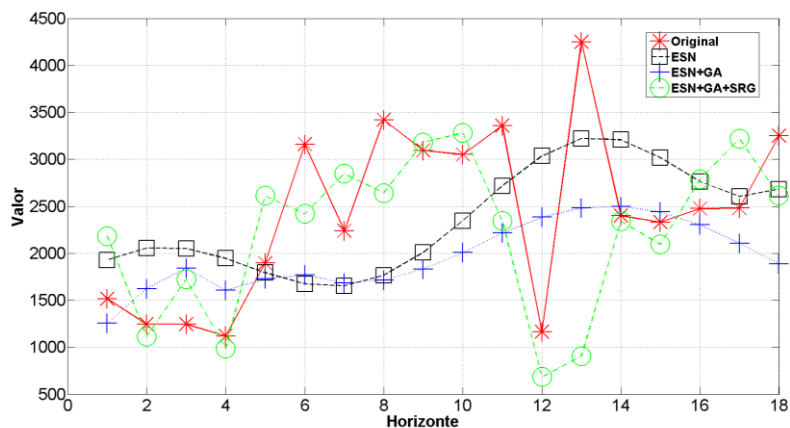


Figura 60. Resultados Série NN3-110 fora da amostra.

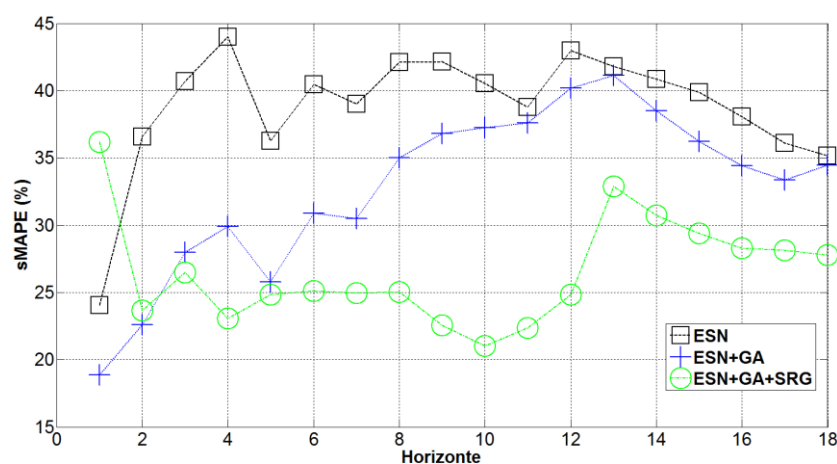


Figura 61. Evolução do sMAPE a longo do horizonte de previsão (Série NN3-110).

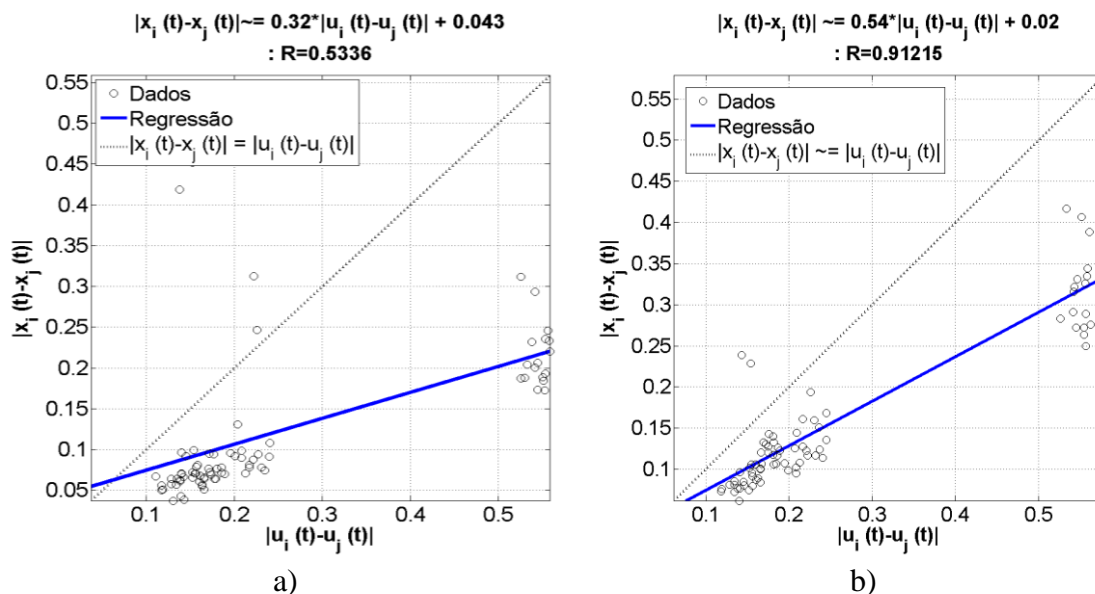


Figura 62. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Para a Série *NN3-110* os resultados apresentaram um comportamento onde o modelo *ESN-GA-SRG* acompanhou os valores originais. Já os outros 2 modelos não apresentaram bons resultados. Na evolução do *sMAPE* os últimos passos apresentaram erros maiores que nos passos iniciais; entretanto, os outros 2 modelos apresentaram comportamentos similares, mas com magnitudes maiores.

Na figura 62 é possível observar que os estados ficaram agrupados em duas seções, o que pode explicar o fato de ter conseguido acompanhar a diferença de magnitude do passo 12 na série original.

Na tabela 43 são apresentados os parâmetros de configuração da rede ESN para a série *NN3-111*.

Tabela 43. Parâmetros finais de configuração para a série *NN3-111*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	331
Janela	11
Porcentagem de conexão	49%
Raio Espectral	0.9
Fator de escala das conexões de entrada	0.3
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.8

Nas figuras 63 a 65 são apresentados os resultados da previsão para o conjunto fora da amostra (teste), a evolução do *SMAPE* a longo do horizonte de previsão e a regressão

linear obtida para o *Reservoir* antes e depois de utilizar o método SRG para a série NN3-111.

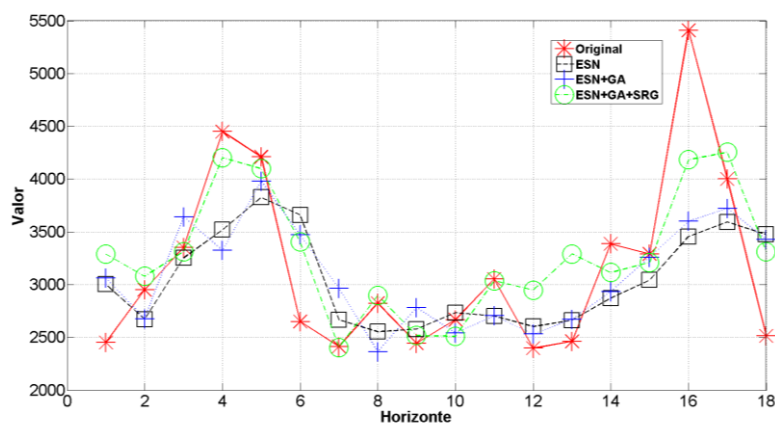


Figura 63. Resultados Série NN3-111 fora da amostra.

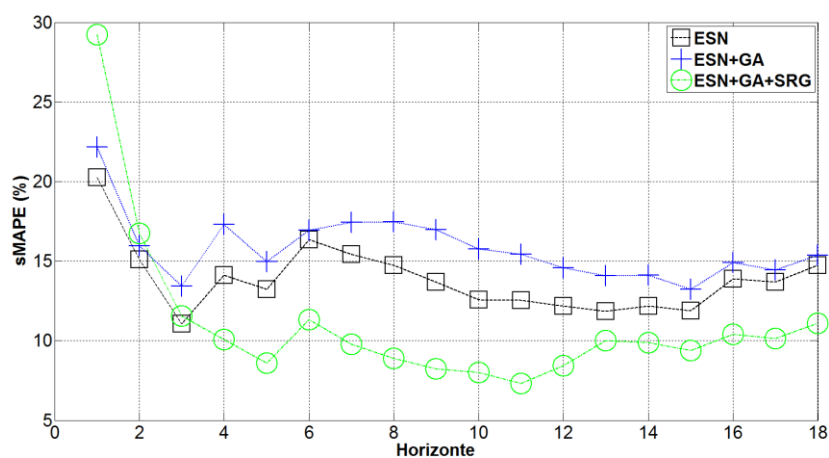


Figura 64. Evolução do sMAPE a longo do horizonte de previsão (Série NN3-111).

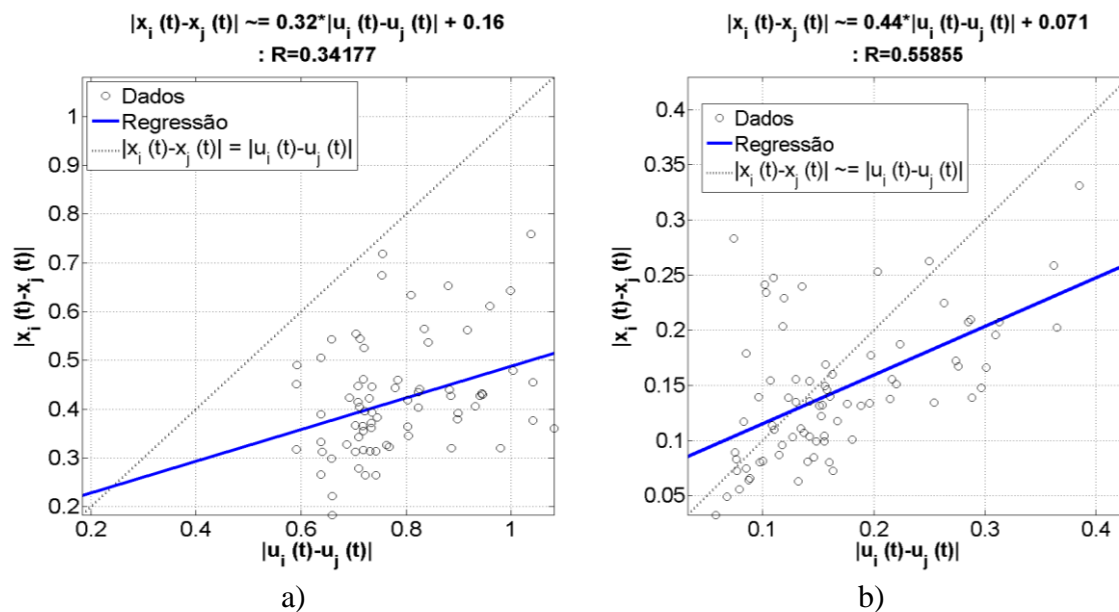


Figura 65. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Para a Série NN3-111 o modelo *ESN-GA-SRG* também apresentou o melhor resultado. Em geral, os outros 2 modelos conseguiram acompanhar o comportamento da série, mas nos valores máximos somente o modelo proposto conseguiu previsões mais adequadas. O comportamento da evolução do *sMAPE* foi estável para os 3 modelos, sendo que os 3 começaram com previsões ruins; entretanto, ao longo do horizonte de previsão o modelo *ESN-GA-SRG* conseguiu manter os valores de erros mais baixos e, por conseguinte, o menor erro global.

Na interpretação dos estados gerados pelo *Reservoir*, observou-se na regressão linear uma dispersão maior do que em outras séries, onde alguns estados tiveram um comportamento caótico que pode ser explicado pela diferença nos primeiros pontos da previsão.

Em geral pode-se concluir que a utilização do modelo *ESN-GA-SRG* melhorou a resposta dos outros modelos utilizados em 10 das 11 séries que compõem o estudo de caso. Adicionalmente, a interpretação do comportamento do *Reservoir* durante as previsões realizadas foi mais clara, isto é, os resultados dos estados permitiram compreender de forma mais completa o comportamento das redes *ESN*.

Em uma comparação posterior, foram considerados os resultados obtidos por outros modelos estatísticos como: Holt-Winters Multiplicativo (HW), Regressão Harmônica (REG), Decomposição clássica (DEC), o Experimento HW/REG/DEC (Método: *AFTER* para janela expansiva & pesos estáticos) e *Box & Jenkins* (BJ). Estes resultados foram publicados em

(Valle, 2012), onde foi proposto o modelo *Neural Expert Weighting* (NEW) que também foi incluído na tabela comparativa a seguir.

Tabela 44. Resultados comparativos para as séries NN3.

Método	SMAPE
HW	15,08
REG	16,68
DEC	16,25
HW/REG/DEC	13,71
BJ	15,03
NEW	13,84
ESN-GA-SRG	11,51

Na tabela 45 estão discriminados os resultados de *sMAPE* para cada uma das 11 séries individualmente. Conforme pode ser verificado, o modelo *ESN-GA-SRG* conseguiu ser o vencedor em 4 séries e no *sMAPE* médio das 11 séries.

Tabela 45. Resultados para cada uma das séries.

Série	HW	REG	DEC	HW/REG/DEC	BJ	NEW	ESN-GA-SRG
1	1,80	2,67	2,61	1,95	1,99	2,64	2,82
2	30,63	24,55	29,01	24,77	12,64	22,05	9,52
3	29,60	33,81	29,16	33,81	44,01	34,91	27,16
4	6,11	9,37	13,78	8,13	4,96	6,41	10,87
5	1,53	10,22	10,66	1,62	2,67	8,28	2,52
6	5,38	3,62	4,10	4,08	5,31	6,07	2,19
7	5,10	3,48	3,29	3,02	5,87	3,94	4,14
8	36,29	29,54	29,62	30,35	29,58	25,83	20,43
9	7,77	26,88	16,69	6,71	7,02	8,62	8,07
10	30,48	27,55	28,62	27,25	33,61	23,02	27,79
11	11,14	11,79	11,23	11,79	17,75	10,49	11,10

Na figura 66 são apresentados os resultados de *sMAPE* para cada série, permitindo comparar os resultados obtidos nas 11 séries.

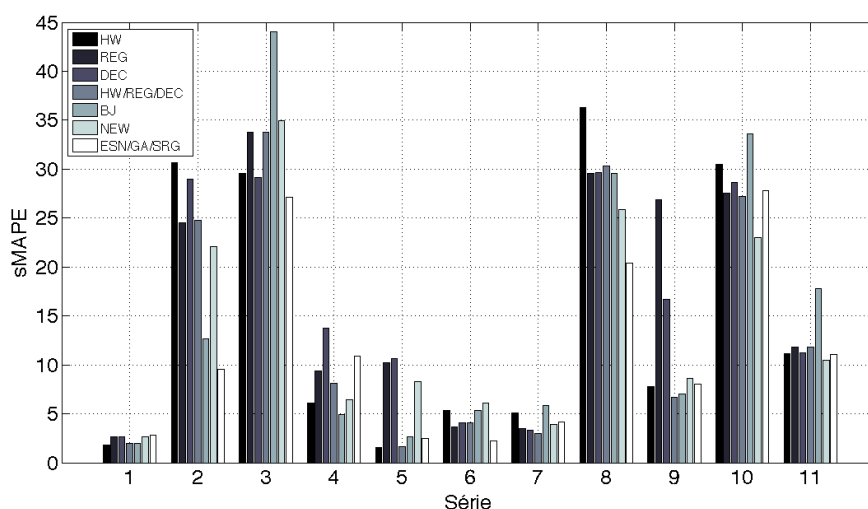


Figura 66. Diagrama de barras comparativas dos resultados obtidos.

Na tabela 46 são apresentados os resultados dos primeiros 10 colocados da competição *NN3-Forecasting*, permitindo observar que o modelo proposto teria conseguido a melhor posição. É importante destacar que o segundo colocado (Ilies, Jäger, Kosuchinas, Rincon, Sakenas, Vaskevcius), que obteve um *sMAPE* de 14,26, utilizou redes *ESN* com uma metodologia onde as previsões realizadas por diferentes redes *ESN* foram ponderadas para a previsão final.

Tabela 46. Resultados dos primeiros 10 lugares da competição *NN3*.

Posição	Modelo	sMAPE
	<i>ESN-GA-SRG</i>	11,51
1	Yan	13,68
2	Ilies, Jäger, Kosuchinas, Rincon, Sakenas, Vaskevcius	14,26
3	Chan, Yao	14,46
4	Yousefi, Miromeni, Lucas	14,49
5	Ahmed, Atiya, Gayar, El-Shishiny	14,52
6	Flores, Anaya, Ramirez, Morales	15,00
7	Adeodato, Vasconcelos, Arnaud, Chunha, Monteiro	15,10
8	Luna, Soares, Ballini	15,35
9	Theodosiou, Swamy	16,19
10	Hwang, Song, Kasabov	16,31

5.2.1.

Resultados testes de significância estatística para as séries *NN3*.

Com o intuito de avaliar de forma estatística os resultados obtidos nas previsões *multi-step* para cada uma das 11 séries foram utilizados testes paramétricos e não-paramétricos apresentados abaixo. Maiores informações e discussões com relação aos testes realizados podem ser encontradas em (Conover, 1980).

Teste	Classificação
T-Student	Paramétrico
Sync	Não-paramétrico
Wilcoxon (Signed Rank)	Não-paramétrico
Friedman	Não-paramétrico
Post-Hoc	Não-paramétrico

O valor estabelecido para nível de significância, identificado como *p-value*, foi o 5%, e a comparação dos diferentes modelos é realizada utilizando o *sMAPE* acumulado para cada passo do horizonte de previsão.

Nas tabelas 47 a 49 são apresentados os resultados para os testes *t-student*, *Sync* e *Wilcoxon*, onde é possível constatar que o modelo proposto conseguiu ter diferenças estatisticamente significativas na maioria dos casos, quando a comparação é realizada par a par.

Tabela 47. Resultados do teste *t-student* para as séries *NN3*.

Teste t				
Modelo	<i>p-valor</i>	Lim. Inferior	Lim. Superior	Conclusão
ESN	4,95e-13	-4,9258	-3,9252	ESN+GA+SRG é melhor
ESN+GA	7,05e-10	-2,6911	-1,9020	ESN+GA+SRG é melhor
NEW	0,2132	-1,6480	0,3955	Equivalente
TRAD	4,09e-8	-3,5469	-2,2412	ESN+GA+SRG é melhor
HW	0,0522	-1,9417	0,0103	Equivalente
BJ	5,86e-8	-4,5670	-2,8510	ESN+GA+SRG é melhor

Tabela 48. Resultados do teste *sync* para as séries *NN3*.

Teste sinal				
Modelo	<i>p-valor</i>	Lim. Inferior	Lim. Superior	Conclusão
ESN	7,62e-6	-5,0435	-4,0349	ESN+GA+SRG é melhor
ESN+GA	7,62e-6	-2,8911	-1,7094	ESN+GA+SRG é melhor
NEW	0,0309	-1,6666	-0,0227	ESN+GA+SRG é melhor
TRAD	1,44e-4	-3,4366	-2,7927	ESN+GA+SRG é melhor
HW	0,2379	-2,2997	0,6573	Equivalente
BJ	1,44e-4	-4,1143	-3,2727	ESN+GA+SRG é melhor

Tabela 49. Resultados do teste *Wilcoxon* para as séries *NN3*

Teste Wilcoxon				
Modelo	<i>p-valor</i>	Lim. Inferior	Lim. Superior	Conclusão
ESN	1,80e-4	-4,9358	-4,1027	ESN+GA+SRG é melhor
ESN+GA	1,80e-4	-2,7995	-1,8844	ESN+GA+SRG é melhor
NEW	0,0407	-1,6666	-0,0016	ESN+GA+SRG é melhor
TRAD	2,53e-4	-3,5662	-2,6785	ESN+GA+SRG é melhor
HW	0,0451	-2,1617	0,0187	ESN+GA+SRG é melhor
BJ	2,14e-4	-4,6085	-3,2181	ESN+GA+SRG é melhor

Na tabela 50 é apresentado o resultado obtido para o teste de Friedman, onde são comparados de forma simultânea os 6 modelos selecionados. É possível observar que o modelo proposto conseguiu obter a primeira posição na comparação.

Tabela 50. Resultados para o teste Friedman.

Teste Friedman							
Modelo	ESN+GA+SRG	NEW	Tradicional	HW	BJ	ESN	ESN+GA
h=1	5	7	6	1	4	2	3
h=2	3	7	4	1	5	2	6
h=3	3	7	5	2	4	1	6
h=4	3	7	4	2	5	1	6
h=5	2	7	4	3	5	1	6
h=6	2	7	4	3	5	1	6
h=7	1	7	4	2	6	3	5
h=8	1	7	4	3	5	2	6
h=9	1	6	4	3	5	2	7
h=10	1	6	4	3	5	2	7
h=11	1	6	4	3	5	2	7
h=12	1	6	4	3	5	2	7
h=13	1	6	4	2	5	3	7
h=14	1	6	4	2	5	3	7
h=15	1	7	5	2	4	3	6
h=16	1	7	3	2	4	5	6
h=17	1	7	3	2	4	5	6
h=18	1	7	3	2	4	6	5
Soma	30	120	73	41	85	46	109
Média	1,6667	6,6667	4,0556	2,2778	4,7222	2,5556	6,0556
Posição	1	7	4	2	5	3	6

Na tabela 51 é apresentado o resultado obtido para o teste de Post-Hoc, onde novamente são comparados todos os modelos de forma simultânea, mas, desta vez, é obtido o nível de significância para todos os modelos selecionados. Novamente é possível observar que o modelos proposto conseguiu obter diferenças estatisticamente significativas para 4 dos 6 modelos.

Tabela 51. Resultados para o teste Post-Hoc.

Teste Pos-Hoc						
Modelo	Modelo	Lim. Inferior	Dif. Médias	Lim. Superior	p-Valor	Conclusão
ESN+GA+SRG	NEW	-7,1230	-5,0000	-2,8770	0,0000	ESN+GA+SRG é melhor
ESN+GA+SRG	Tradicional	-4,5119	-2,3889	-0,2659	0,0158	ESN+GA+SRG é melhor
ESN+GA+SRG	HW	-2,7341	-0,6111	1,5119	0,9797	Equivalente
ESN+GA+SRG	BJ	-5,1786	-3,0556	-0,9325	0,0004	ESN+GA+SRG é melhor
ESN+GA+SRG	ESN	-3,0119	-0,8889	1,2341	0,8808	Equivalente
ESN+GA+SRG	ESN+GA	-6,5119	-4,3889	-2,2659	0,0000	ESN+GA+SRG é melhor
NEW	Tradicional	0,4881	2,6111	4,7341	0,0053	Equivalente
NEW	HW	2,2659	4,3889	6,5119	0,0000	Equivalente
NEW	BJ	-0,1786	1,9444	4,0675	0,0981	Equivalente
NEW	ESN	1,9881	4,1111	6,2341	0,0000	Equivalente
NEW	ESN+GA	-1,5119	0,6111	2,7341	0,9797	Equivalente
Tradicional	HW	-0,3453	1,7778	3,9008	0,1707	Equivalente
Tradicional	BJ	-2,7897	-0,6667	1,4564	0,9685	Equivalente
Tradicional	ESN	-0,6230	1,5000	3,6230	0,3627	Equivalente
Tradicional	ESN+GA	-4,1230	-2,0000	0,1230	0,0802	Equivalente

HW	BJ	-4,5675	-2,4444	-0,3214	0,0122	HW é melhor
HW	ESN	-2,4008	-0,2778	1,8453	0,9997	Equivalente
HW	ESN+GA	-5,9008	-3,7778	-1,6547	0,0000	HW é melhor
BJ	ESN	0,0436	2,1667	4,2897	0,0419	Equivalente
BJ	ESN+GA	-3,4564	-1,3333	0,7897	0,5128	Equivalente
ESN	ESN+GA	-5,6230	-3,5000	-1,3770	0,0000	ESN é melhor

5.3.

Resultados para as séries selecionadas da competição *M3*

A seguir são apresentados os resultados obtidos nos experimentos realizados com o modelo proposto *ESN-GA-SRG* e a sua comparação com os 6 modelos melhor posicionados para cada conjunto de séries selecionadas da competição *M3*.

Inicialmente, a base global disponível foi separada em 3 grupos segundo a amostragem: anual, trimestral e mensal com horizontes de previsão de 6, 8 e 18 passos a frente respectivamente. Dentro de cada um destes 3 grupos, foram selecionadas 12 séries para cada grupo em função de características como assimetria, estacionariedade e sazonalidade, conforme apresentado nas tabelas 52 a 54.

Tabela 52. Séries selecionadas para o grupo de amostragem anual.

Anual		Assimetria		
		Baixa	Media	Alta
Estacionariedade	Sazonalidade			
Não	Não	M3_171	M3_105	M3_119
Não	Sim	M3_050	M3_033	M3_564
Sim	Não	M3_162	M3_642	M3_070
Sim	Sim	M3_599	M3_351	M3_200

Tabela 53. Séries selecionadas para o grupo de amostragem trimestral.

Trimestral		Assimetria		
		Baixa	Media	Alta
Estacionariedade	Sazonalidade			
Não	Não	M3_896	M3_730	M3_847
Não	Sim	M3_758	M3_773	M3_865
Sim	Não	M3_692	M3_713	M3_854
Sim	Sim	M3_671	M3_829	M3_677

Tabela 54. Séries selecionadas para o grupo de amostragem mensal.

Mensal		Assimetria		
		Baixa	Media	Alta
Estacionariedade	Sazonalidade			
Não	Não	M3_1996	M3_1430	M3_1460
Não	Sim	M3_1725	M3_2089	M3_1912
Sim	Não	M3_2001	M3_2016	M3_2579
Sim	Sim	M3_1900	M3_1906	M3_2013

Na tabela 55 são apresentados os resultados comparativos, em média, para as 12 séries selecionadas para cada grupo de amostragem. É possível observar que, segundo a métrica *Smape*, o modelo proposto conseguiu o menor erro na previsão *multi-step*, no caso particular do ForecastPro que é uma plataforma computacional foram utilizados diversos modelos como Exponential Smoothing/Box&Jenkins/Poisson.

Tabela 55. Resultados comparativos para as séries M3.

Anual		Trimestral		Mensal	
Método	sMAPE	Método	sMAPE	Método	sMAPE
ForecastPro	17,05	ForecastPro	4,42	ForcX	14,88
Flors_Pearc2	17,20	ForcX	5,06	SMARTFCS	15,19
AutoBox2	17,38	Holt	5,38	Arama	15,88
Flors_Pearc1	17,38	THETA	5,64	B_J_Auto	16,19
SMARTFCS	17,39	B_J_Auto	5,73	ForecastPro	16,44
THETA	17,40	Arama	5,96	Auto_ANN	16,68
ESN-GA-SRG	8,88	ESN-GA-SRG	3,86	ESN-GA-SRG	11,73

Nas tabelas 56 a 58 apresentam-se os resultados de *sMAPE* para cada uma das 12 séries, permitindo comparar os resultados obtidos nas 12 séries individualmente.

Tabela 56. Resultados comparativos para cada série M3 anual.

sMAPE							
Série	ForecastPro	Flors_Pearc2	AutoBox2	Flors_Pearc1	SMARTFCS	THETA	ESN-GA-SRG
33	32,80	12,56	34,91	7,48	18,96	12,55	6,26
50	13,73	12,22	15,23	12,22	26,28	21,66	7,32
70	9,59	16,20	9,92	10,48	8,41	8,71	8,17
105	21,54	11,17	7,55	15,66	15,21	18,91	5,52
119	31,66	29,53	24,88	32,03	29,66	31,34	26,41
162	29,31	19,24	16,88	16,50	30,82	21,06	14,55
171	3,14	31,94	17,17	38,46	17,58	26,09	3,10
200	14,85	12,88	33,14	14,77	12,79	13,88	11,24
351	17,38	22,06	20,56	22,06	12,32	22,59	8,46
564	12,31	11,31	13,31	11,31	8,29	9,48	8,42
599	5,28	13,76	4,01	14,12	15,18	8,92	2,14
642	12,96	13,46	10,98	13,46	13,25	13,65	5,09

Tabela 57. Resultados comparativos para cada série M3 trimestral.

<i>sMAPE</i>							
Série	ForecastPro	ForcX	Holt	THETA	B_J_Auto	Ararma	ESN-GA-SRG
671	3,91	1,84	3,03	2,83	8,51	3,59	1,42
677	1,09	5,31	4,06	4,14	1,09	1,78	1,09
692	4,16	5,55	3,78	3,79	3,65	5,94	3,35
713	4,39	5,71	4,08	5,49	3,96	3,88	3,46
730	3,71	3,72	5,30	6,66	6,04	5,47	3,28
758	3,68	3,40	2,60	6,56	3,47	9,08	2,17
773	8,79	8,65	17,83	7,09	8,62	9,89	6,48
829	3,55	7,53	3,85	10,84	6,98	3,09	5,74
847	11,72	10,57	10,24	10,69	16,95	16,95	11,04
854	3,20	2,93	3,06	2,87	2,89	2,70	2,70
865	2,59	2,78	4,02	4,13	4,01	6,13	3,49
896	2,29	2,85	2,83	2,65	2,66	3,14	2,16

Tabela 58. Resultados comparativos para cada série M3 mensal.

<i>sMAPE</i>							
Série	ForcX	SMARTFCS	Ararma	B_J_Auto	ForecastPro	Auto_ANN	ESN-GA-SRG
1430	15,98	15,49	25,38	13,96	15,12	16,26	13,51
1460	50,54	51,39	63,25	63,25	56,99	48,82	38,52
1725	13,94	9,65	8,57	9,62	8,74	11,83	8,18
1900	2,51	4,20	3,39	2,93	2,80	3,04	2,49
1906	3,91	3,80	4,20	4,35	3,90	6,20	3,90
1912	6,40	3,96	6,86	9,20	6,39	10,24	5,14
1996	6,40	7,32	6,23	5,79	7,87	6,30	4,95
2001	18,87	19,56	18,02	17,19	17,06	21,33	17,02
2013	5,97	4,58	4,78	7,71	8,49	5,87	5,84
2016	2,68	3,30	4,12	2,07	7,85	4,43	1,55
2089	21,65	25,72	15,99	22,26	32,48	33,69	12,33
2579	29,77	33,30	29,83	36,02	34,75	32,20	27,38

Nas tabelas 59 a 61 são apresentadas as comparações dos resultados obtidos pelo modelo proposto na previsão para o conjunto fora da amostra e a melhor previsão disponível na base de dados da conferencia **M3**. Originalmente são disponibilizadas as previsões de 22 modelos testados para a base de dados **M3** e, desses 22 modelos, foi selecionada a melhor previsão segundo a métrica **sMAPE**.

No caso da primeira coluna os valores entre parêntesis representam a assimetria de cada série e na última coluna se tem a posição que conseguiu a modelo ESN-GA-SRG em comparação com os 22 modelos que tem disponibilizadas as previsões para a base **M3**, isto é, se o valor é 1 quer dizer que o modelo proposto conseguiu superara as previsões dos 22 modelos, se o valor for 3, quer dizer que teve 2 modelos que conseguiram gerar melhores previsões que o modelo proposto.

Tabela 59. Comparação do *sMAPE* para as séries anuais.

Series Anuais (horizonte=6)			
Assimetria Baixa			
Série	sMape_t	Melhor	Pos ESN+GA+SRG
M3_0171 (-0,0091)	3,0979	3,1470	1
M3_0050 (-0,6796)	7,3201	7,7307	1
M3_0162 (-0,6440)	14,5461	16,5073	1
M3_0599 (-1,1126)	2,1429	1,8553	2
Assimetria Média			
M3_0105 (0,6033)	5,5191	6,7810	1
M3_0033 (-1,4141)	6,2559	5,1345	2
M3_0642 (-0,9847)	5,0895	6,6080	1
M3_0351 (-1,3609)	8,4590	8,0671	2
Assimetria Alta			
M3_0119 (7,7543)	26,4079	24,8895	3
M3_0564 (2,8818)	8,4158	5,36	3
M3_0070 (-1,2893)	8,1687	8,1081	3
M3_0200 (1,4525)	11,2445	12,7912	1

Tabela 60. Comparativa do *sMAPE* para as séries trimestrais.

Series Trimestrais (horizonte=8)			
Assimetria Baixa			
Série	sMape_t	Melhor (22)	Pos ESN+GA+SRG
M3_0896 (-0,7616)	2,1560	2,2646	1
M3_0758 (-0,8037)	2,1736	2,1802	1
M3_0692 (-0,0068)	3,3453	3,6473	1
M3_0671 (0,0443)	1,4196	1,5117	1
Assimetria Media			
M3_0730 (-1,0847)	3,2801	3,3553	1
M3_0773 (-0,9884)	6,4759	7,0256	1
M3_0713 (0,1067)	3,4628	3,6846	1
M3_0829 (-0,3549)	5,7433	3,0902	4
Assimetria Alta			
M3_0847 (23,3153)	11,0396	10,2407	6
M3_0865 (-1,0614)	3,4853	1,9497	6
M3_0854 (-0,6452)	2,7044	2,6995	2
M3_0677 (-1,1469)	1,0908	1,0895	2

Tabela 61. Comparativa do *sMAPE* para as séries mensais.

Series Mensais (horizonte=18)			
Assimetria Baixa			
Série	sMape_t	Melhor (22)	Pos ESN+GA+SRG
M3_1996 (-0,2653)	4,9509	5,7892	1
M3_1725 (0,2113)	8,1807	8,0084	2
M3_2001 (-0,4481)	17,0218	16,4626	2
M3_1900 (0,0985)	2,4895	2,5028	1
Assimetria Media			
M3_1430 (1,0424)	13,5065	12,9857	2
M3_2089 (-1,0064)	12,3327	15,9869	1
M3_2016 (-0,6650)	1,5519	2,0673	1

M3_1906 (-0,8738)	3,8981	3,7980	3
-------------------	--------	--------	---

Assimetria Alta			
M3_1460 (2,1892)	38,5206	48,8221	1
M3_1912 (-1,0965)	5,1391	3,7447	5
M3_2579 (-0,9716)	27,3807	29,7661	1
M3_2013 (-0,9488)	5,8351	4,5847	4

5.3.1.

Resultados para as séries do grupo anual da competição M3

Para cada uma das séries, as tabelas 62 a 74, apresentam os parâmetros de configuração finais do modelo *ESN-GA-SRG* para as séries do grupo anual e as figuras 67 a 92 apresentam os resultados da previsão do conjunto fora da amostra (teste) para o modelo proposto em cada uma das séries selecionadas e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG.

Tabela 62. Parâmetros finais de configuração para a série M3-0033.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	240
Janela	6
Porcentagem de conexão	90%
Raio Espectral	1
Fator de escala das conexões de entrada	0.6
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.6

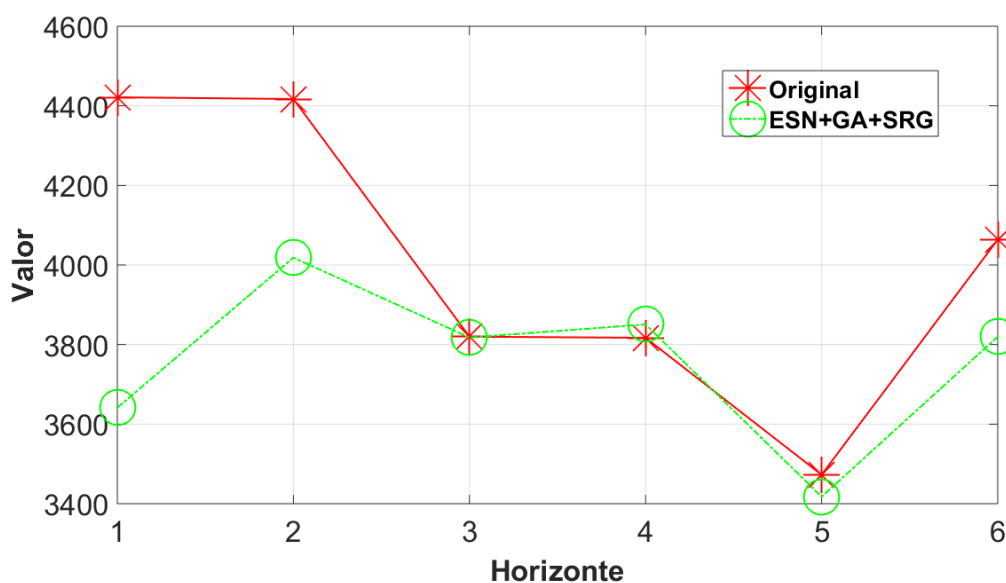


Figura 67. Resultados Série M3-0033 fora da amostra.

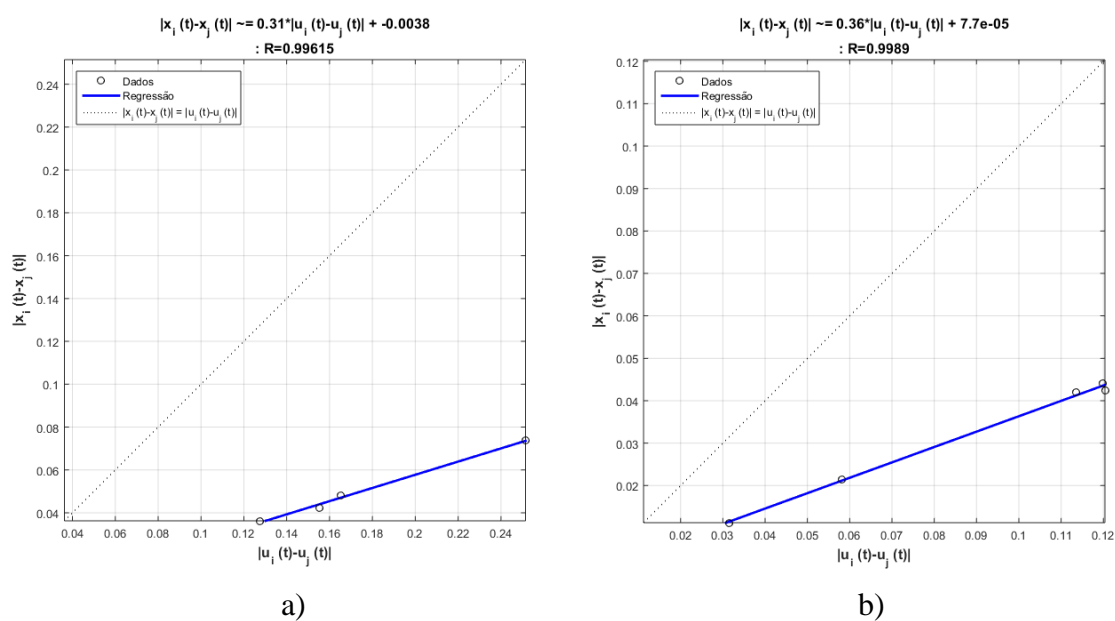


Figura 68. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 63. Parâmetros finais de configuração para a série *M3-0050*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	110
Janela	6
Porcentagem de conexão	20%
Raio Espectral	0.7
Fator de escala das conexões de entrada	0.7
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.2

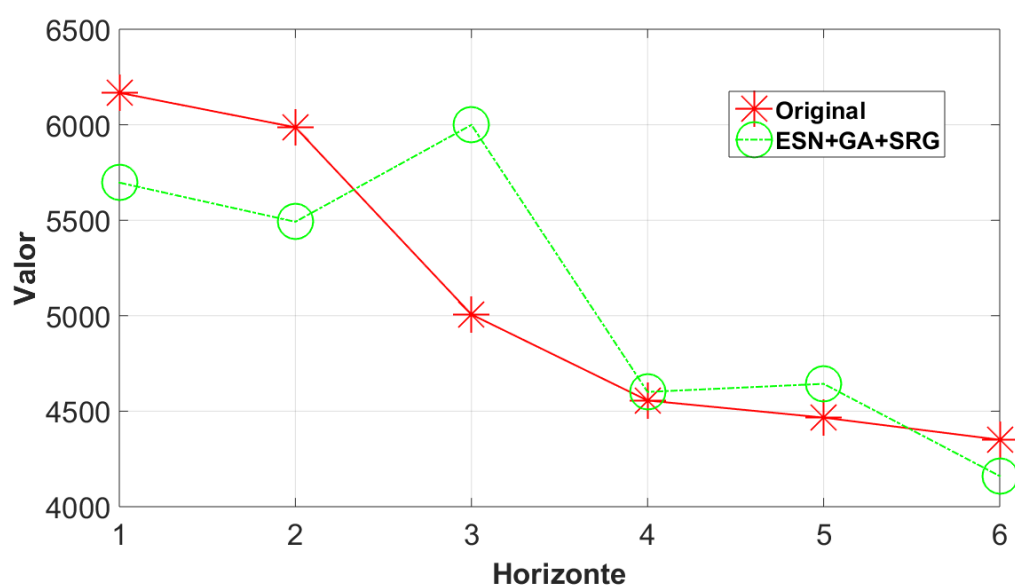


Figura 69. Resultados Série *M3-0050* fora da amostra.

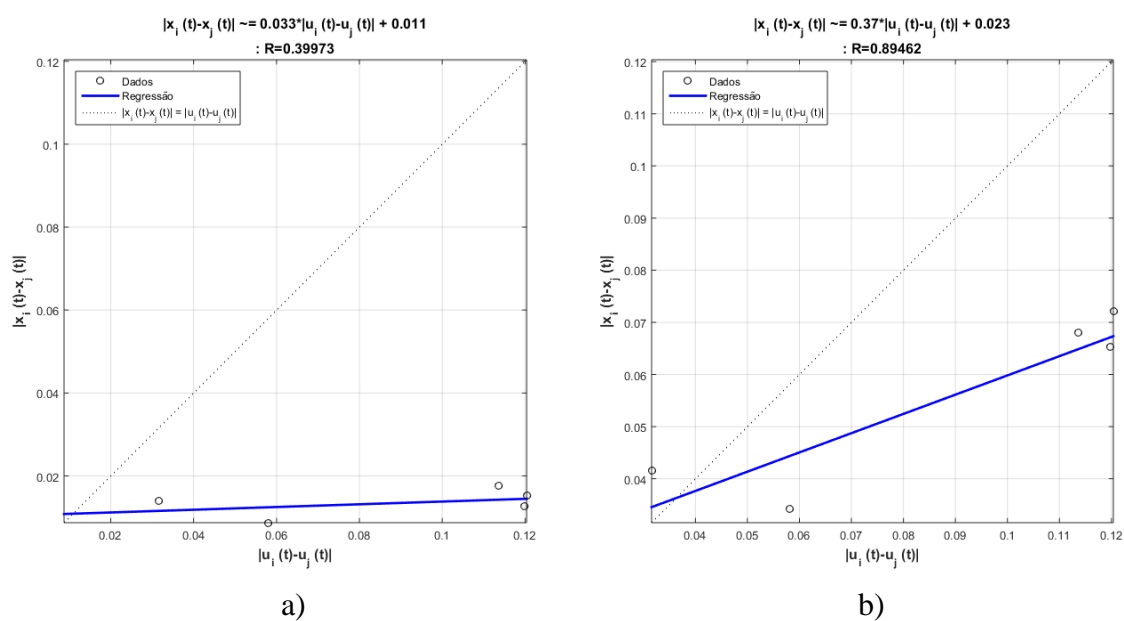


Figura 70. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 64. Parâmetros finais de configuração para a série M3-0070.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	120
Janela	5
Porcentagem de conexão	20%
Raio Espectral	0.9
Fator de escala das conexões de entrada	0.6
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.3

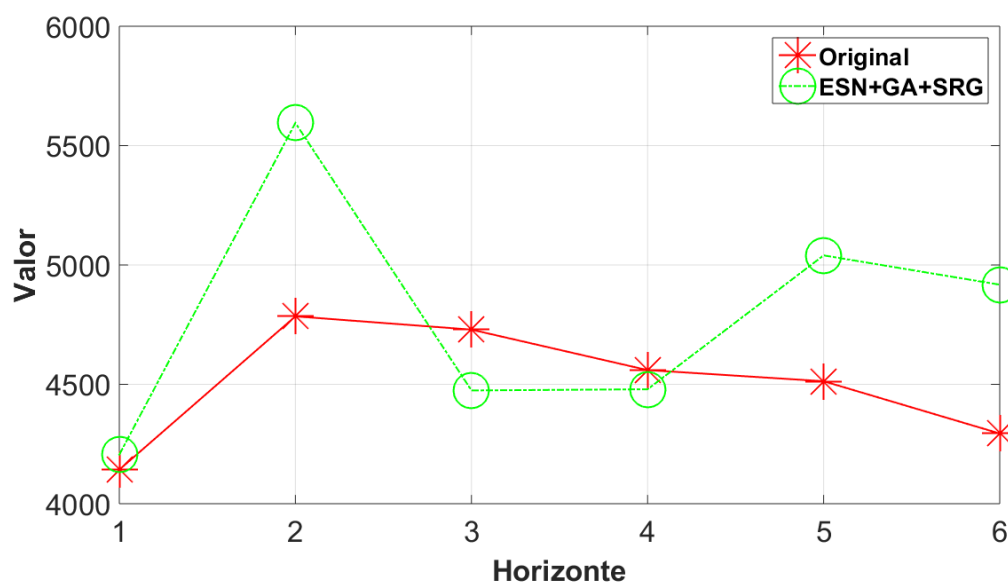


Figura 71. Resultados Série M3-0070 fora da amostra.

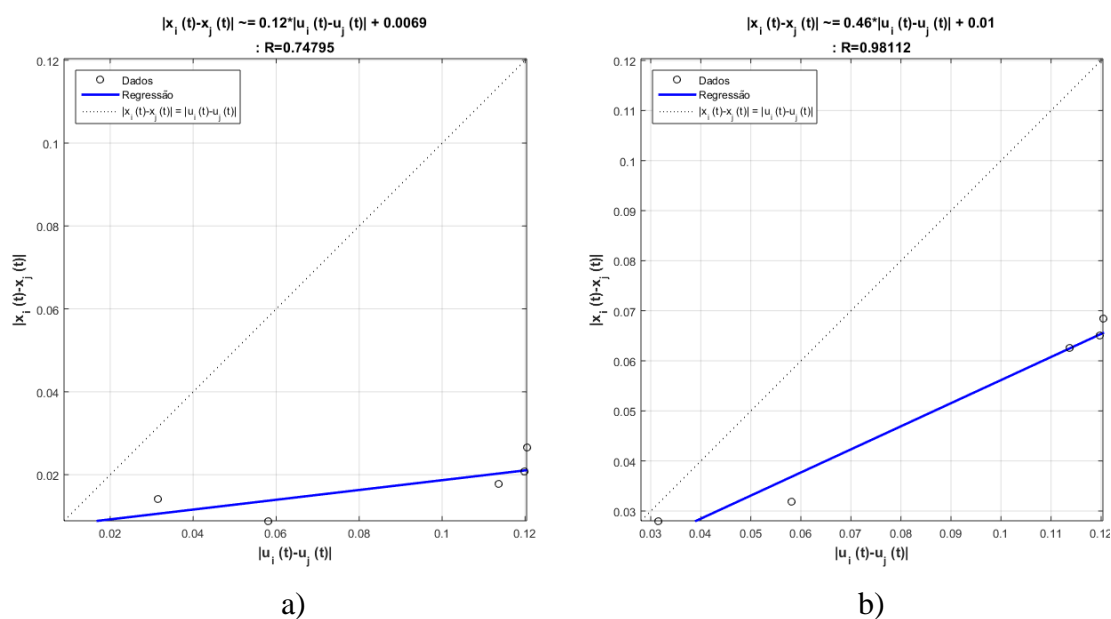


Figura 72. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 65. Parâmetros finais de configuração para a série M3-0105.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	186
Janela	4
Porcentagem de conexão	23%
Raio Espectral	0.8
Fator de escala das conexões de entrada	0.5
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.4

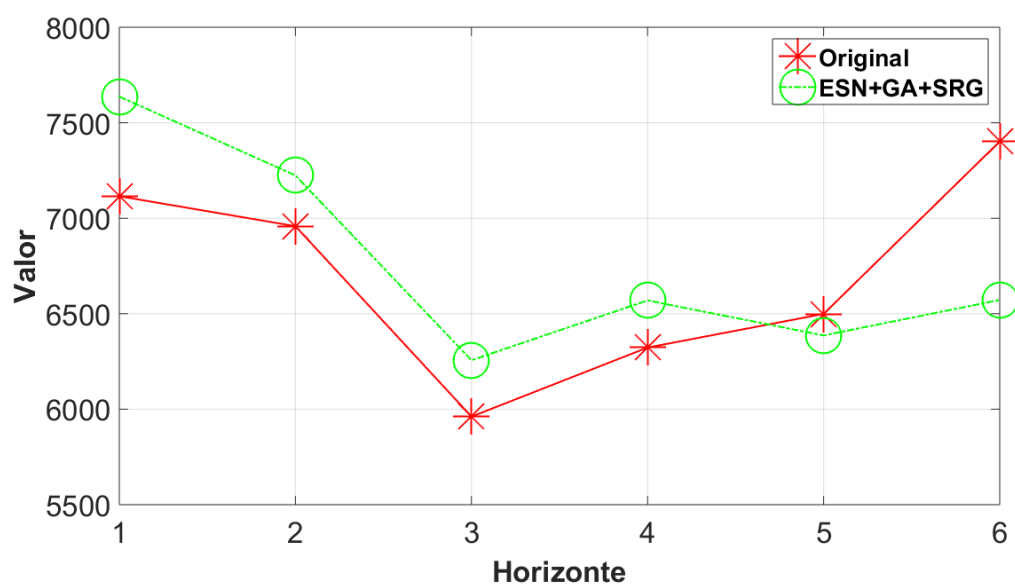


Figura 73. Resultados Série M3-0105 fora da amostra.

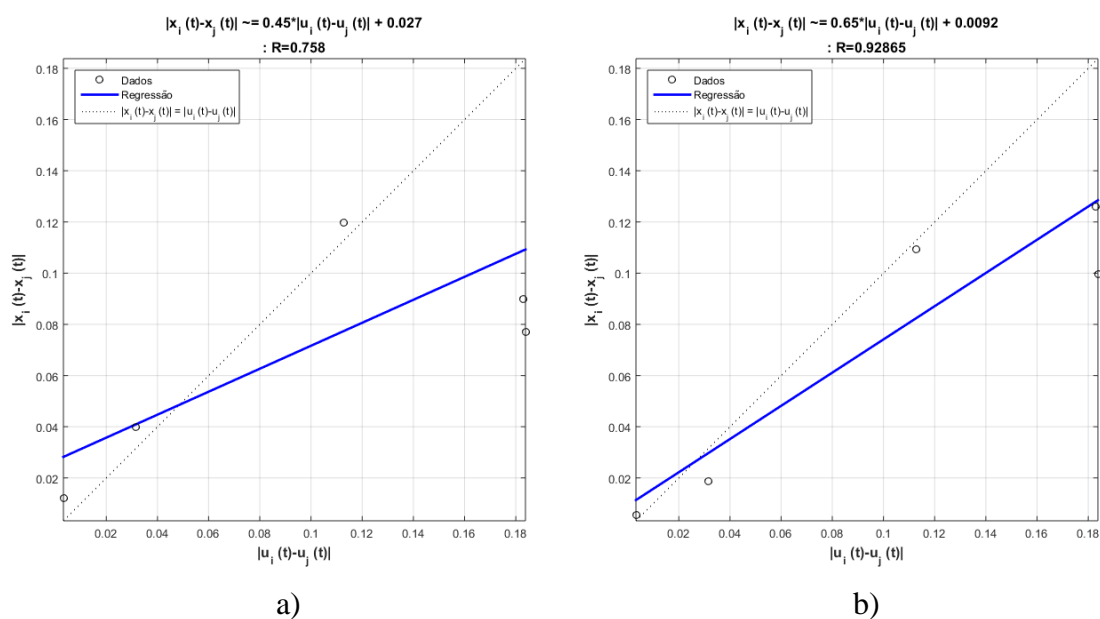


Figura 74. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 66. Parâmetros finais de configuração para a série M3-0116.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	231
Janela	4
Porcentagem de conexão	20%
Raio Espectral	0.9
Fator de escala das conexões de entrada	0.3
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.6

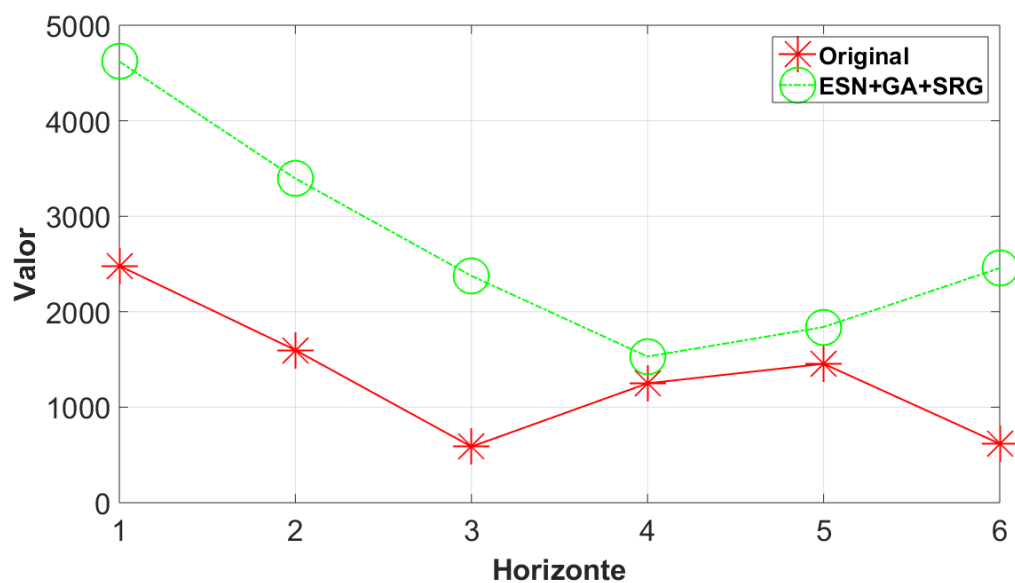


Figura 75. Resultados Série M3-0116 fora da amostra.

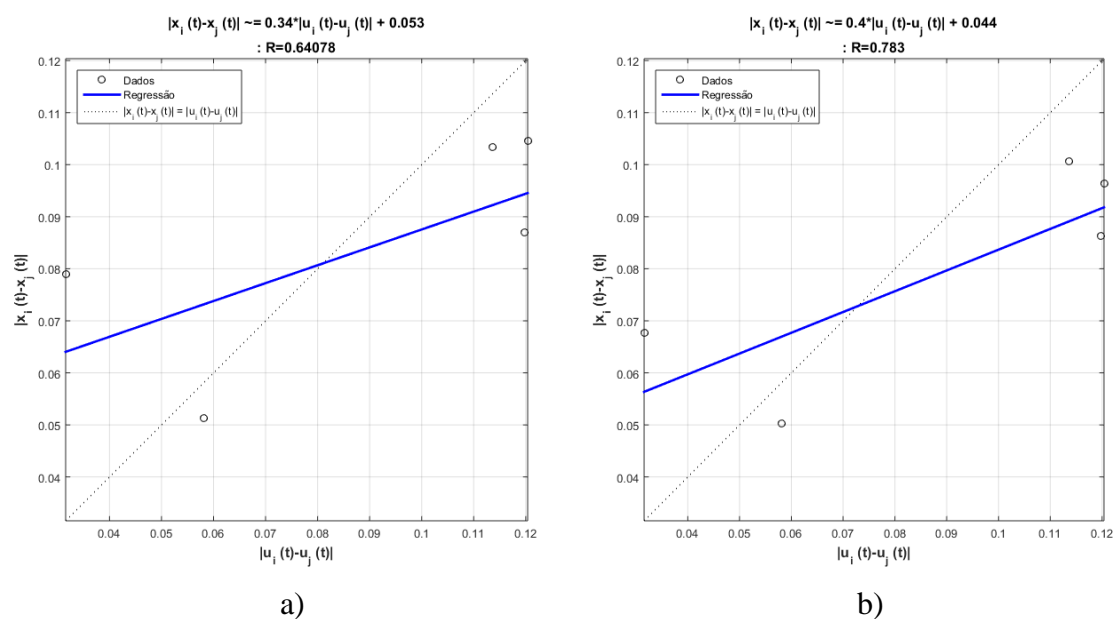


Figura 76. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 67. Parâmetros finais de configuração para a série M3-0119.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	287
Janela	5
Percentagem de conexão	90%
Raio Espectral	0.2
Fator de escala das conexões de entrada	0.8
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.3

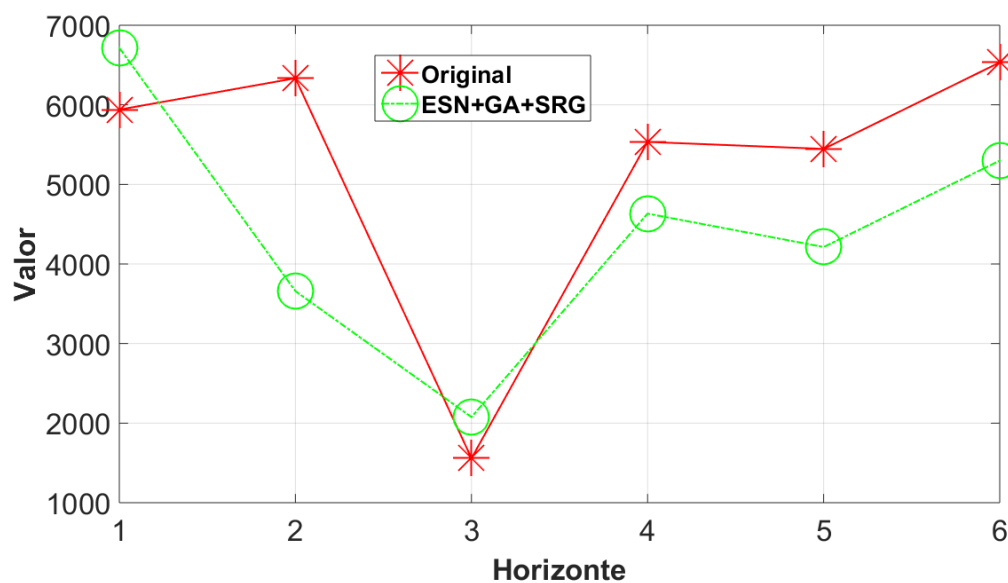


Figura 77. Resultados Série M3-0119 fora da amostra.

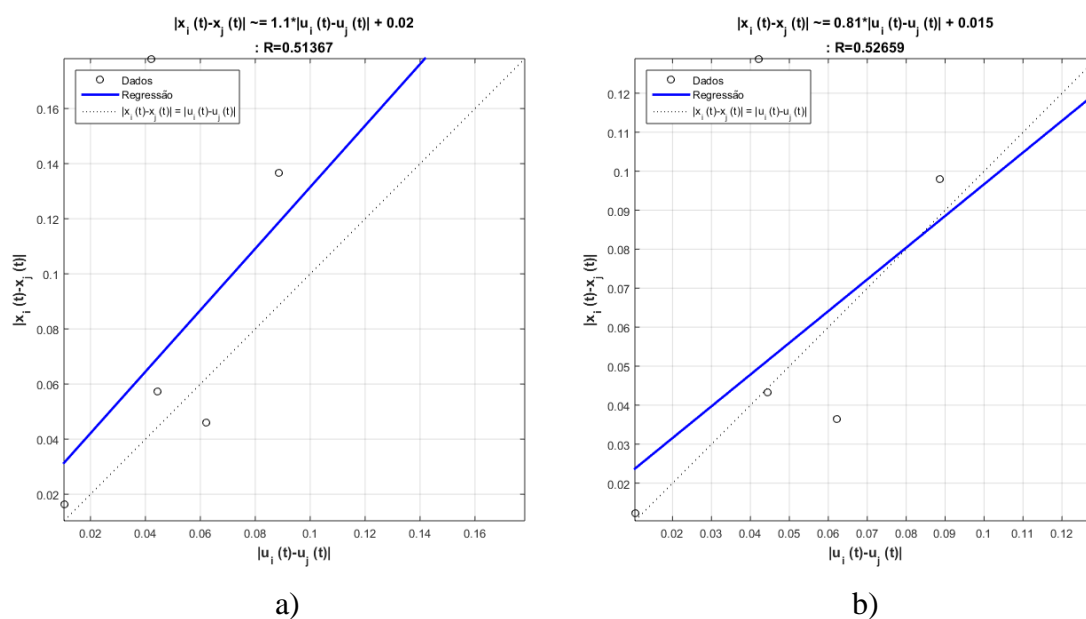


Figura 78. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 68. Parâmetros finais de configuração para a série M3-0162.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	308
Janela	6
Porcentagem de conexão	30%
Raio Espectral	0.7
Fator de escala das conexões de entrada	0.1
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.9

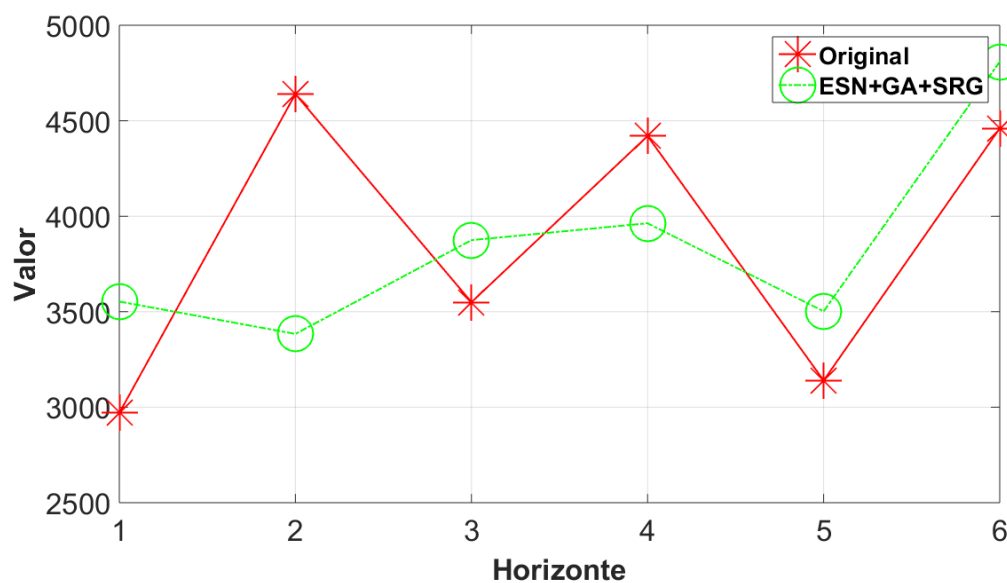


Figura 79. Resultados Série M3-0162 fora da amostra.

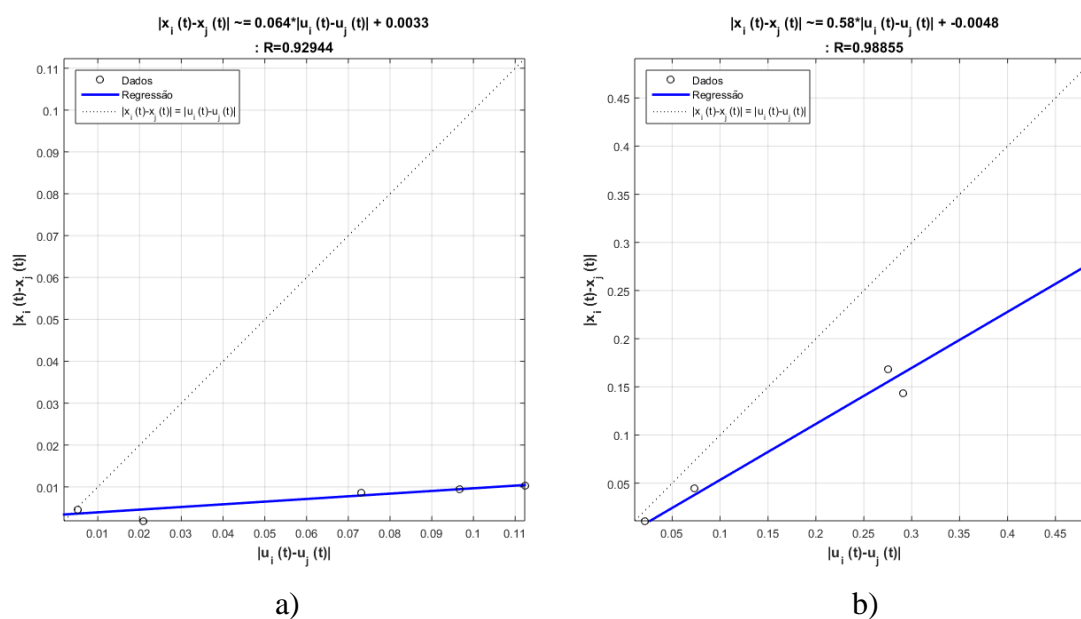


Figura 80. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 69. Parâmetros finais de configuração para a série M3-0171.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	213
Janela	6
Porcentagem de conexão	25%
Raio Espectral	0.6
Fator de escala das conexões de entrada	0.7
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.4

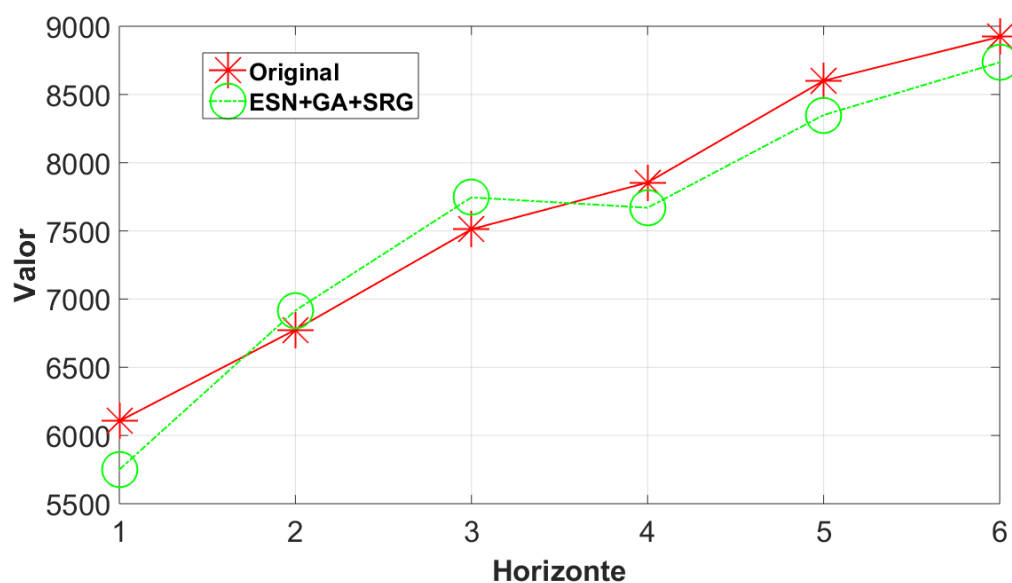


Figura 81. Resultados Série M3-0171 fora da amostra.

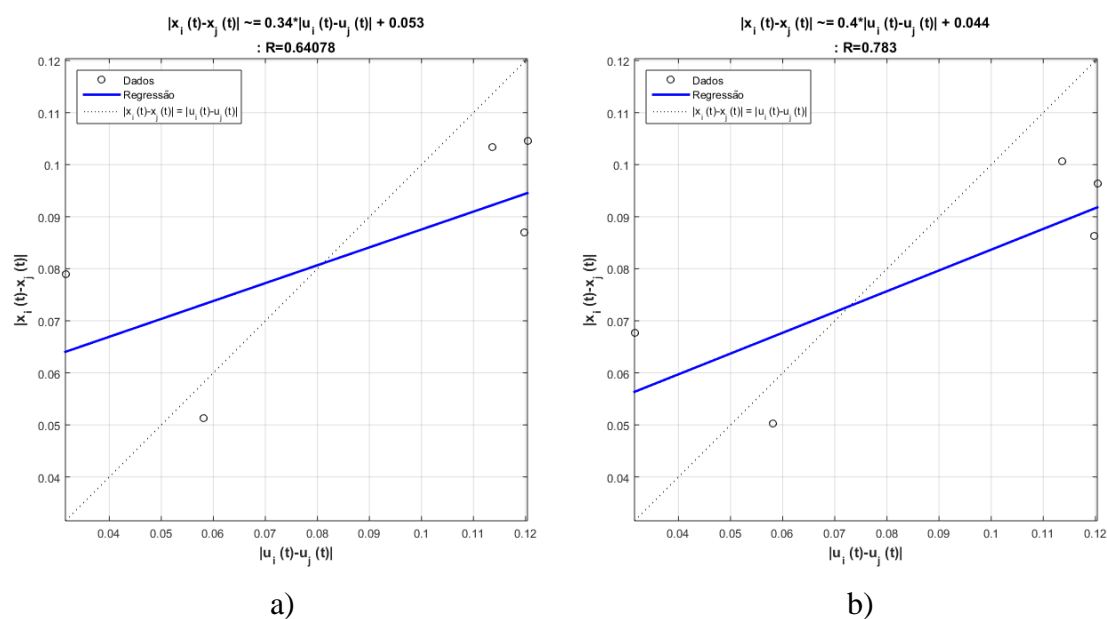


Figura 82. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 70. Parâmetros finais de configuração para a série M3-0200.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	279
Janela	7
Porcentagem de conexão	21%
Raio Espectral	0.6
Fator de escala das conexões de entrada	0.6
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.5

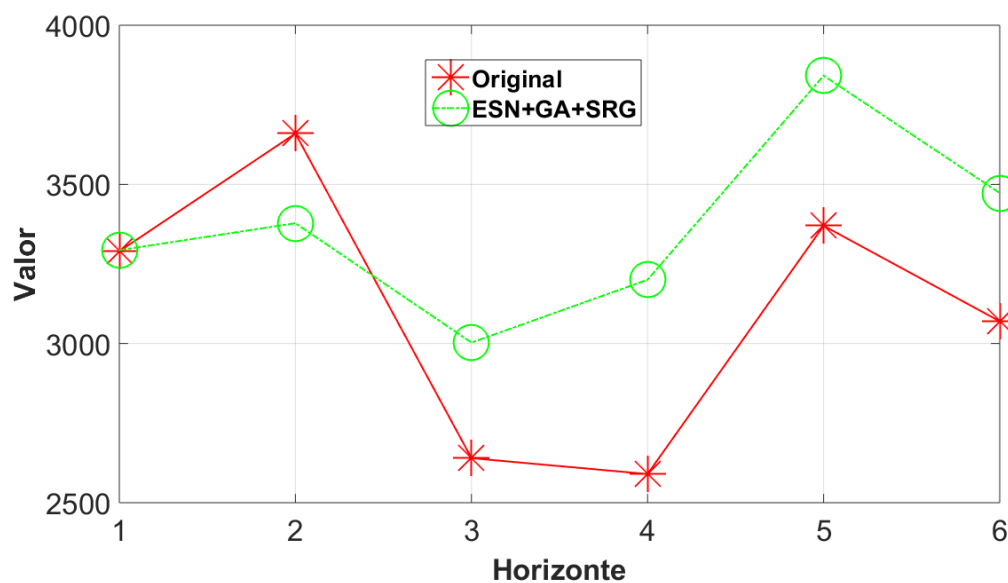


Figura 83. Resultados Série M3-0200 fora da amostra.

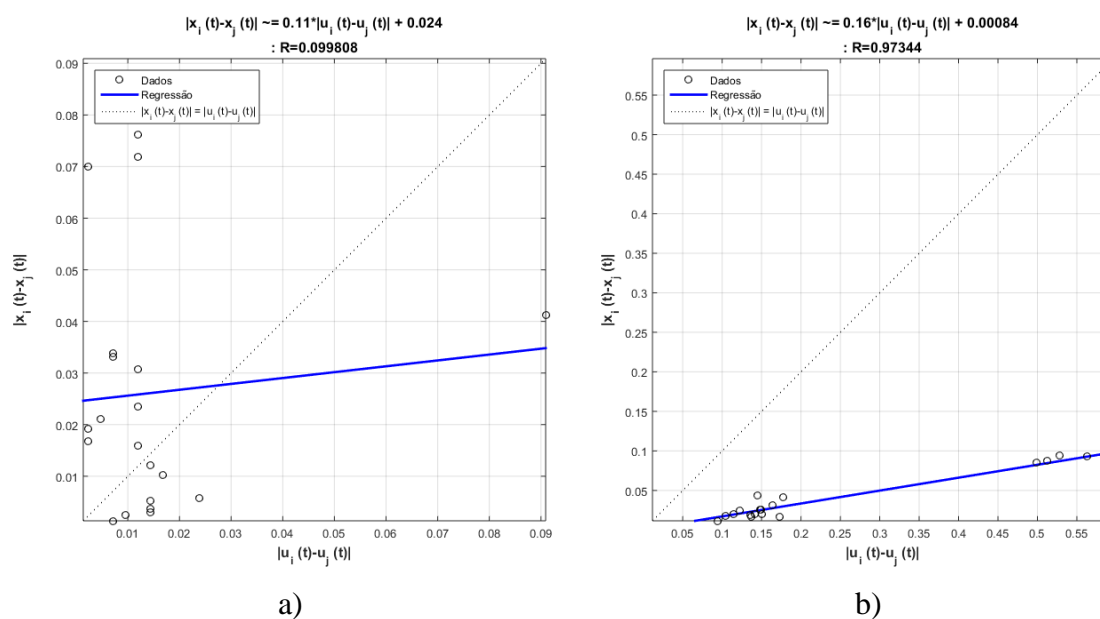


Figura 84. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 71. Parâmetros finais de configuração para a série M3-0351.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	301
Janela	4
Porcentagem de conexão	22%
Raio Espectral	0.4
Fator de escala das conexões de entrada	0.3
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.5

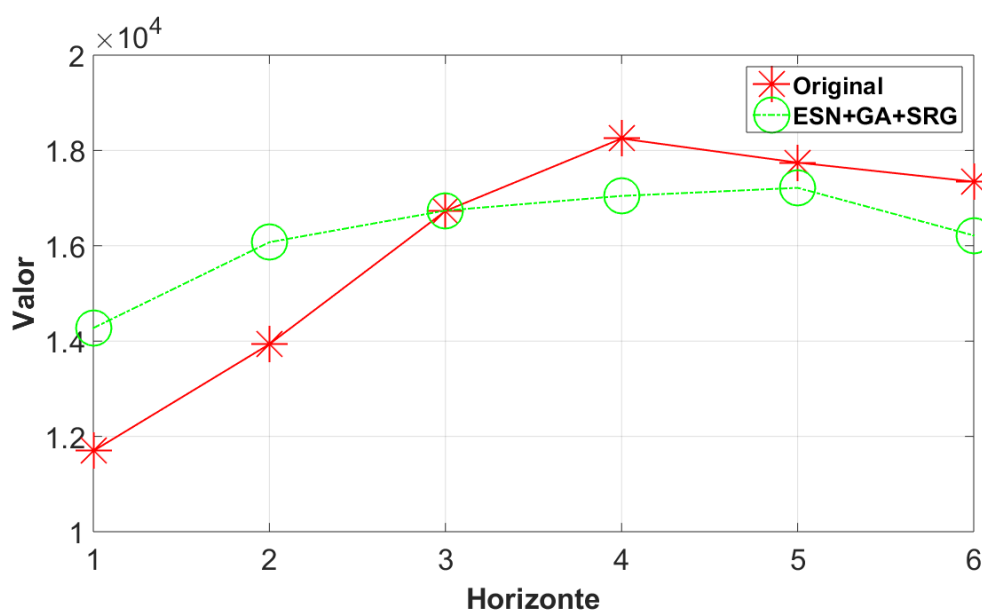


Figura 85. Resultados Série M3-0351 fora da amostra.

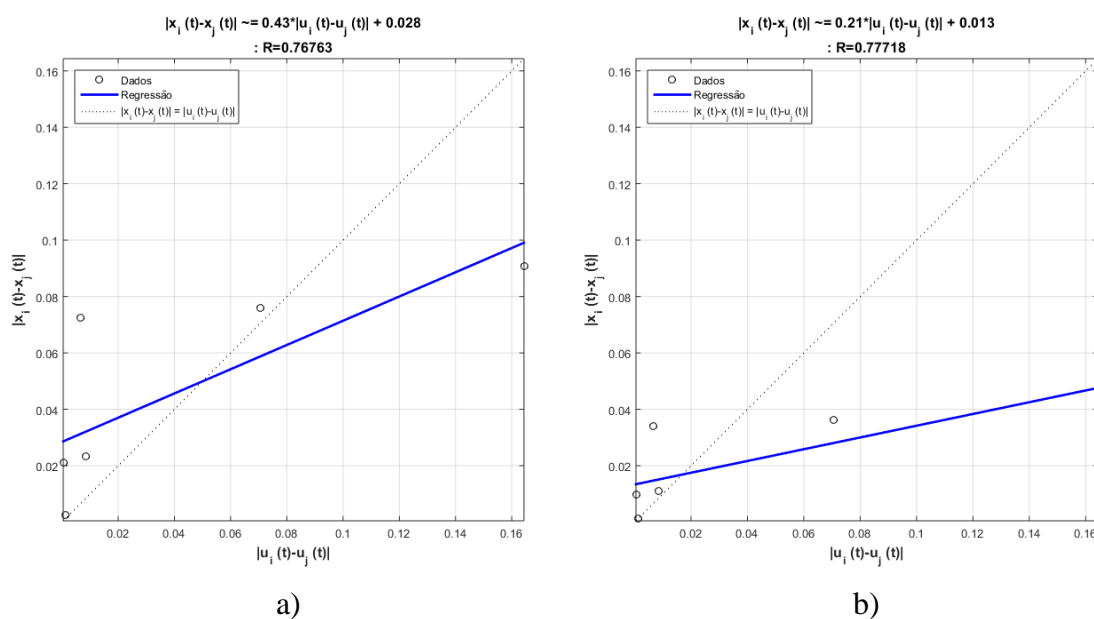


Figura 86. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 72. Parâmetros finais de configuração para a série M3-0564.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	195
Janela	3
Porcentagem de conexão	40%
Raio Espectral	0.2
Fator de escala das conexões de entrada	0.2
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.5

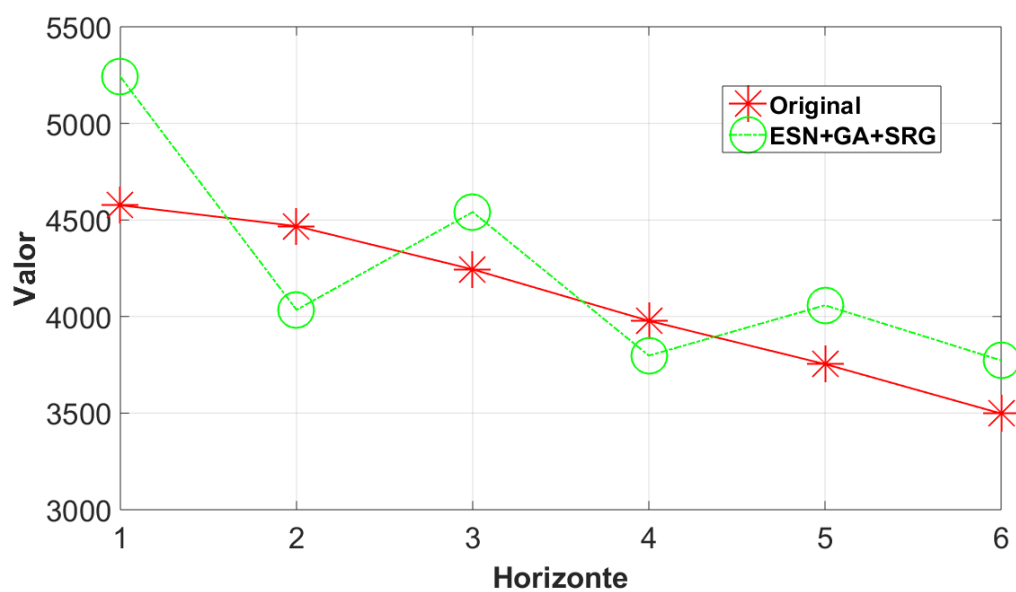


Figura 87. Resultados Série M3-0564 fora da amostra.

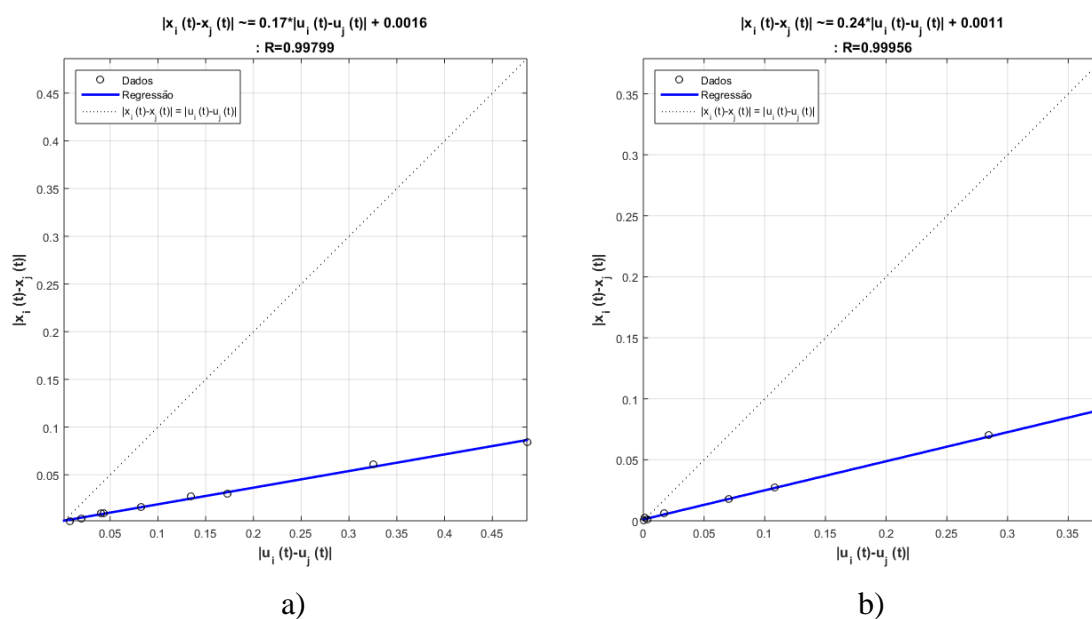


Figura 88. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 73. Parâmetros finais de configuração para a série M3-0599.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	173
Janela	5
Porcentagem de conexão	10%
Raio Espectral	0.6
Fator de escala das conexões de entrada	0.5
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.6

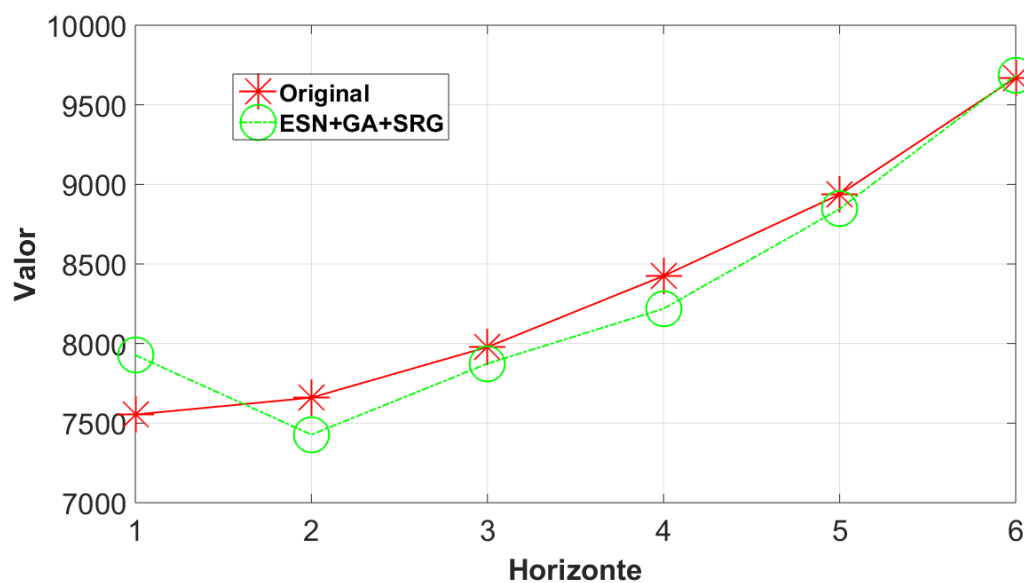


Figura 89. Resultados Série M3-0599 fora da amostra.

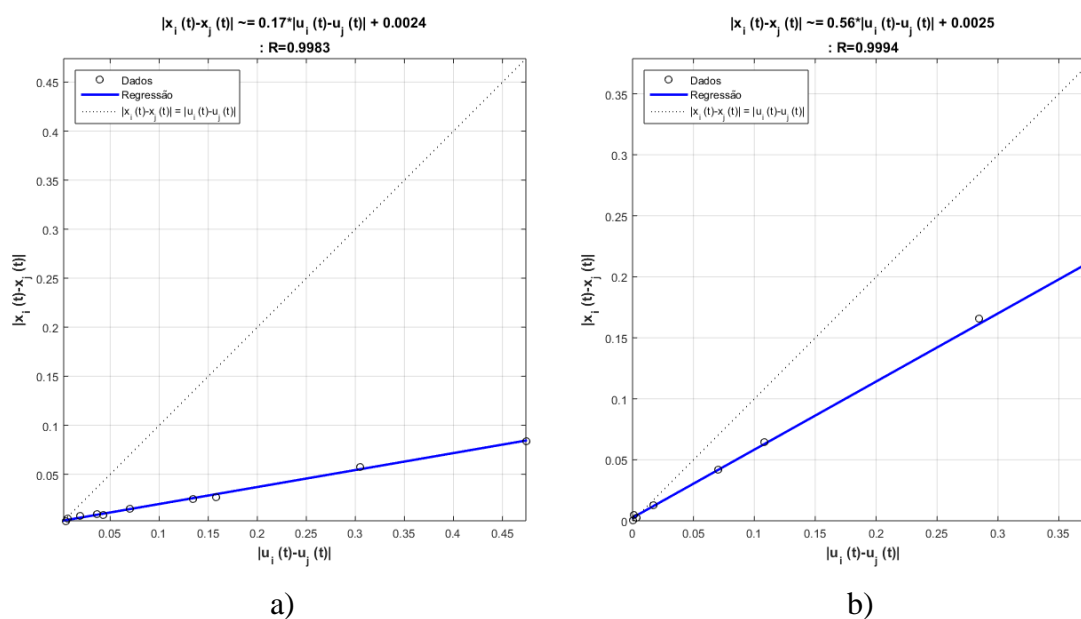


Figura 90. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 74. Parâmetros finais de configuração para a série M3-0642.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	342
Janela	5
Porcentagem de conexão	80%
Raio Espectral	0.6
Fator de escala das conexões de entrada	0.7
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.4

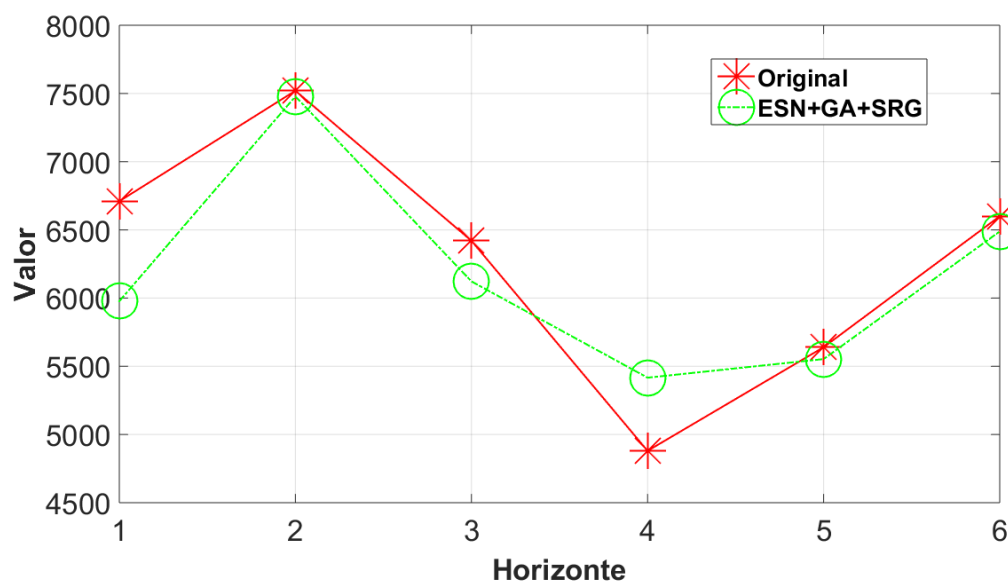


Figura 91. Resultados Série M3-0642 fora da amostra.

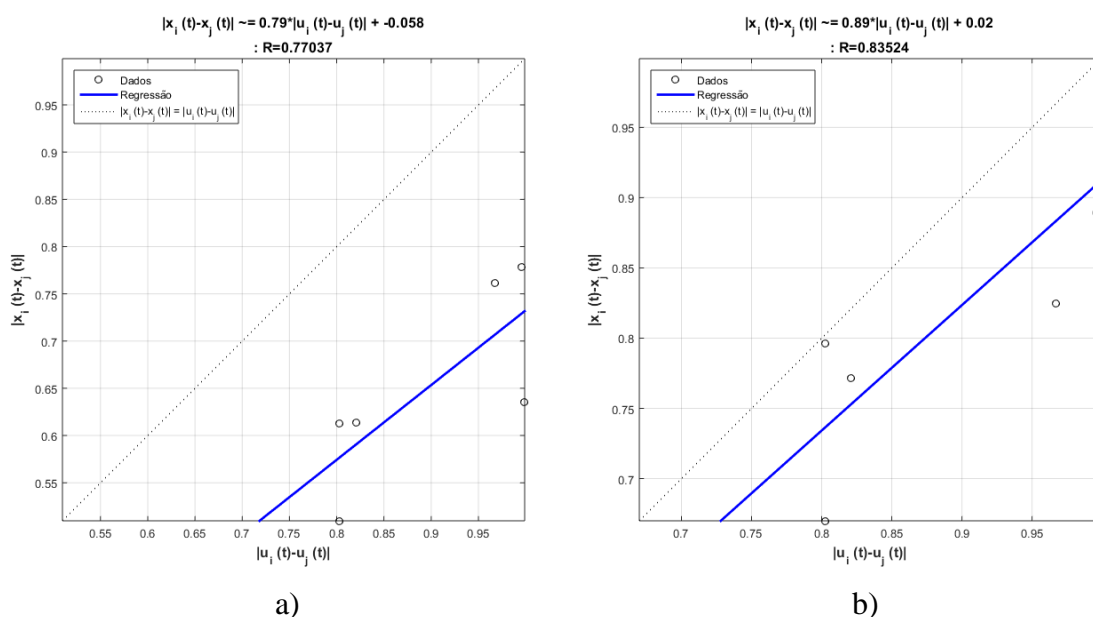


Figura 92. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Na figura 93 é apresentada a evolução do sMAPE para os 6 modelos com melhor desempenho. É possível observar que o modelo proposto teve o melhor desempenho ao longo do horizonte de previsão.

De forma geral se pode afirmar que, para as séries selecionadas no grupo de amostragem anual, o modelo proposto teve um desempenho comparativo melhor do que os modelos selecionados.

Uma das características que foi constatada nas previsões das séries NN3 também foi observada nas previsões do grupo de amostragem anual, que é a melhora de desempenho na previsão ao longo da previsão. Muitos fatores podem ser os responsáveis por este tipo de comportamento, porém, como também foi constatado na análise de sensibilidade (Capítulo 3), a propriedade de gerar históricos através da apresentação dos padrões de forma recorrente produz um conjunto de estados que permite aproveitar o atrelamento temporal das séries selecionadas.

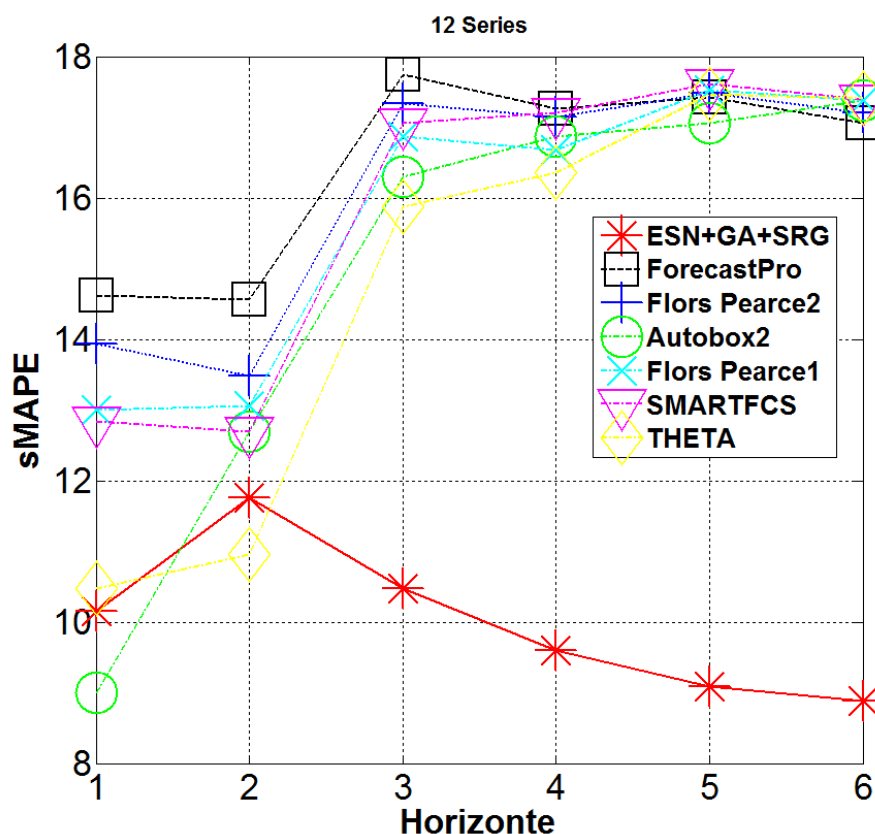


Figura 93. Evolução do sMAPE para cada modelo.

5.3.2.

Resultados para as séries do grupo trimestral da competição M3

Para cada uma das séries, as tabelas 75 a 86, apresentam os parâmetros de configuração finais do modelo *ESN-GA-SRG* para as séries do grupo trimestral e as figuras 94 a 117 apresentam os resultados da previsão do conjunto fora da amostra (teste) para o modelo proposto em cada uma das séries selecionadas e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG.

Tabela 75. Parâmetros finais de configuração para a série M3-0671.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	251
Janela	5
Porcentagem de conexão	26%
Raio Espectral	0.8
Fator de escala das conexões de entrada	0.8
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.1

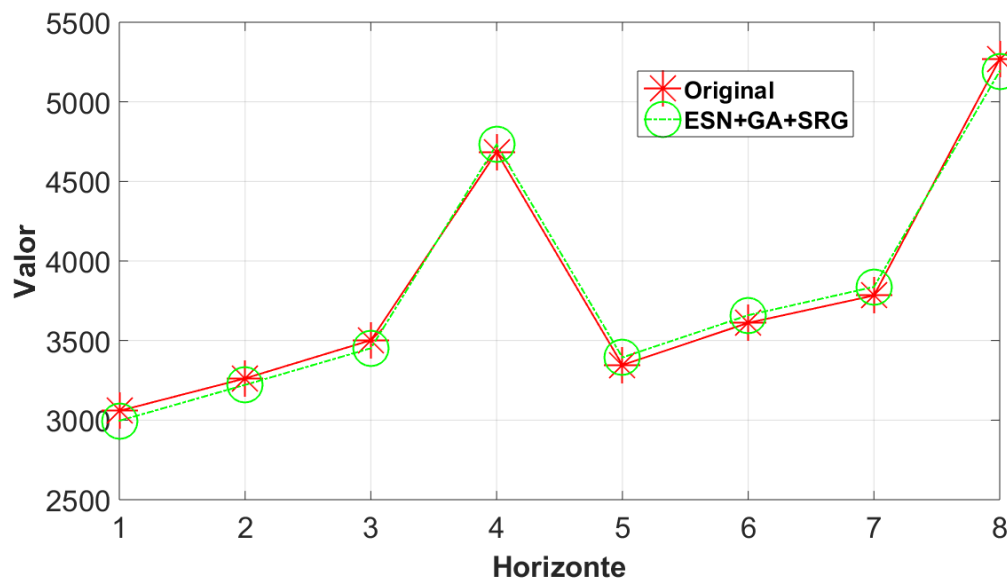


Figura 94. Resultados Série M3-0671 fora da amostra.

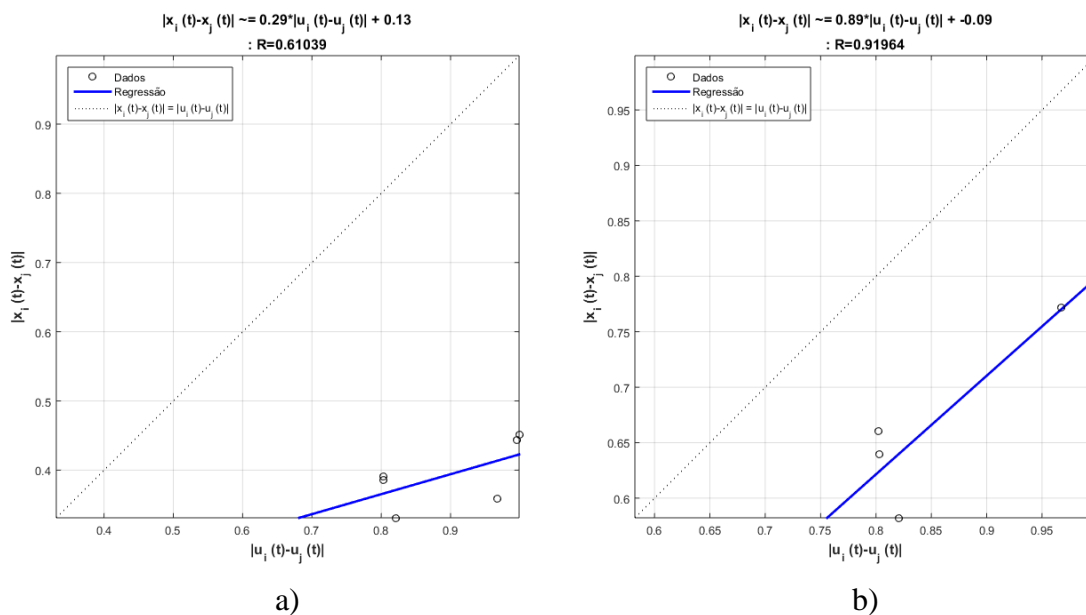


Figura 95. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 76. Parâmetros finais de configuração para a série M3-0677.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	358
Janela	4
Porcentagem de conexão	34%
Raio Espectral	0.7
Fator de escala das conexões de entrada	0.6
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.4

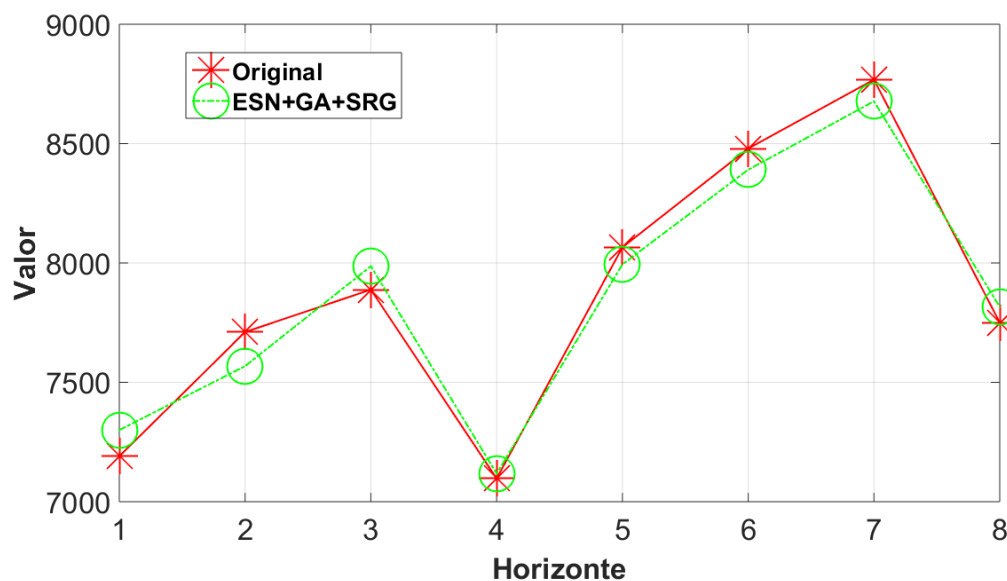


Figura 96. Resultados Série M3-0677 fora da amostra.

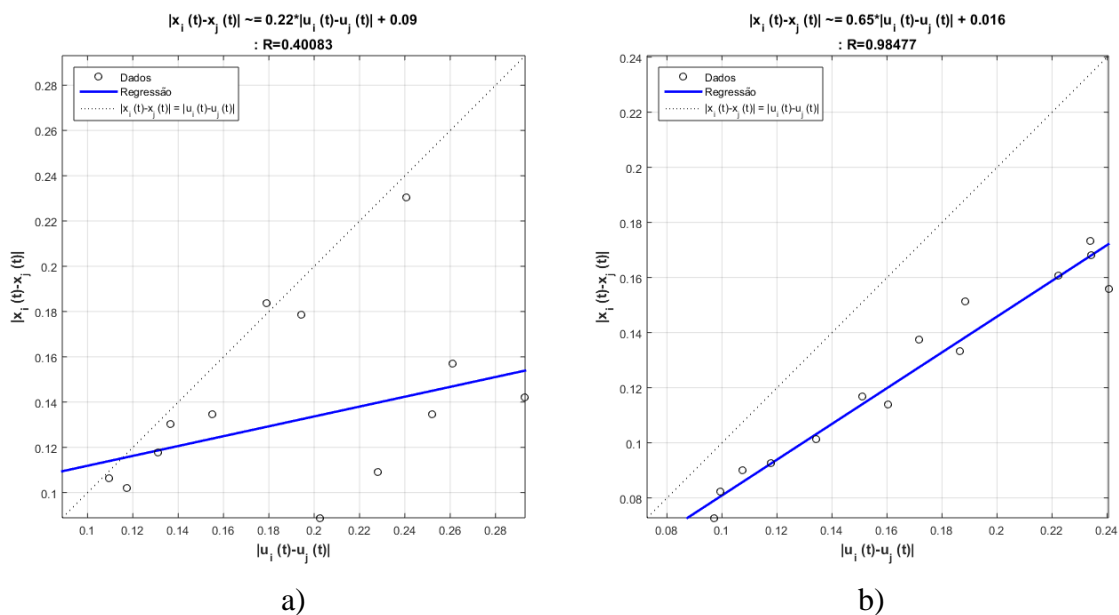


Figura 97. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 77. Parâmetros finais de configuração para a série M3-0692.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	259
Janela	5
Percentagem de conexão	33%
Raio Espectral	0.5
Fator de escala das conexões de entrada	0.6
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.6

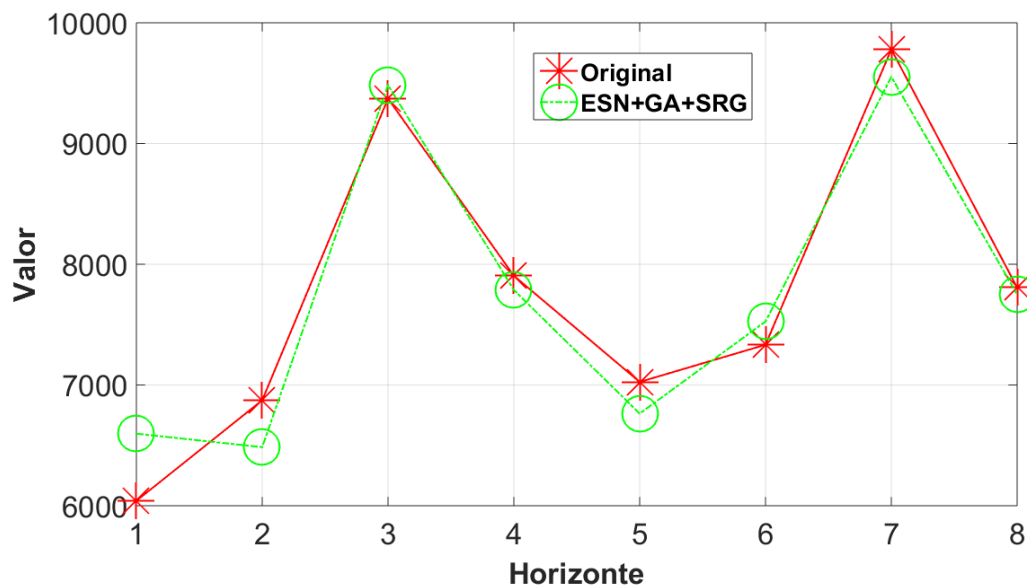


Figura 98. Resultados Série M3-0692 fora da amostra.

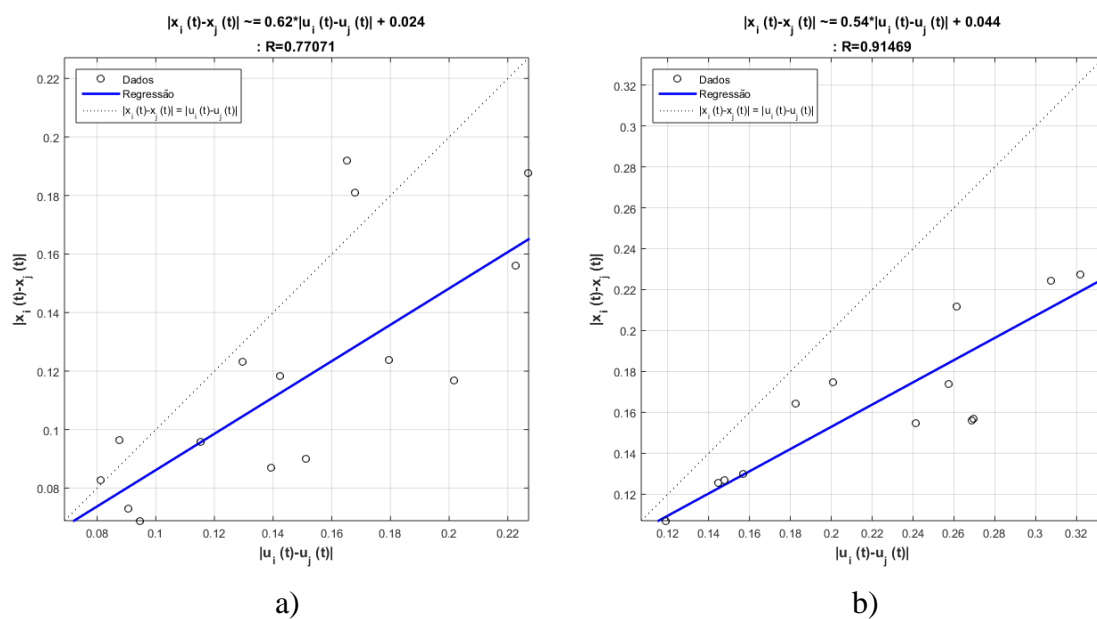


Figura 99. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 78. Parâmetros finais de configuração para a série M3-0713.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	430
Janela	4
Porcentagem de conexão	36%
Raio Espectral	1
Fator de escala das conexões de entrada	0.2
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.8

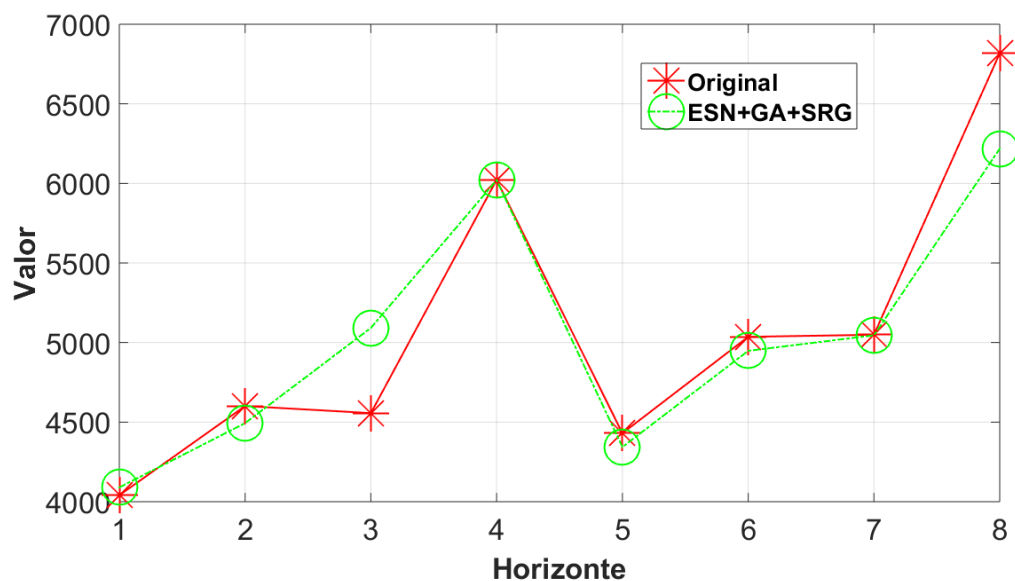


Figura 100. Resultados Série M3-0713 fora da amostra.

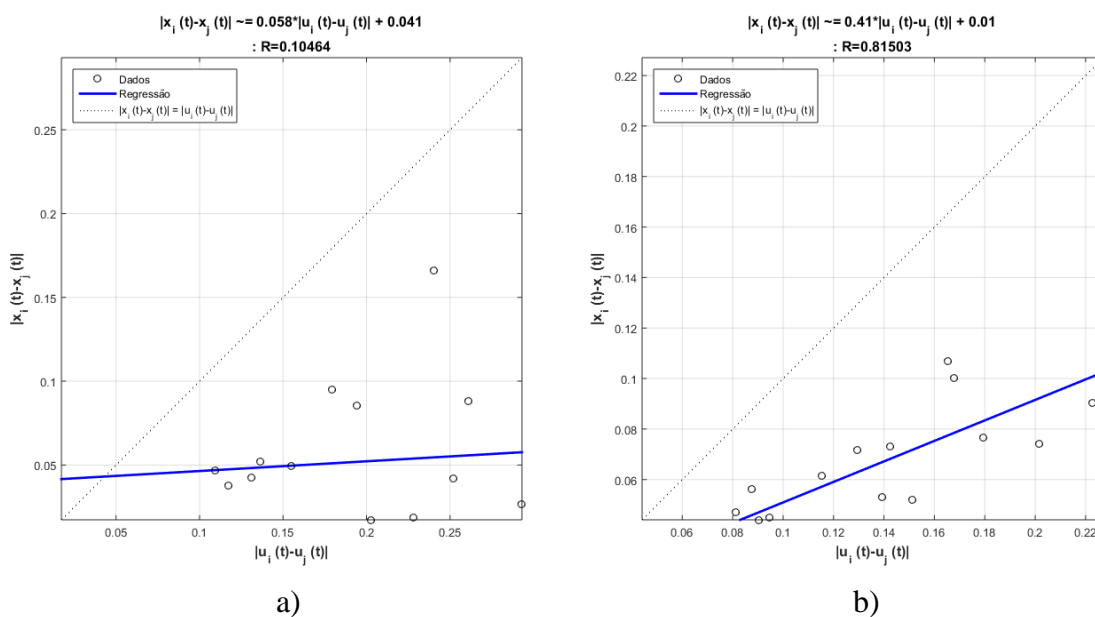


Figura 101. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 79. Parâmetros finais de configuração para a série M3-0730.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	327
Janela	7
Porcentagem de conexão	21%
Raio Espectral	0.4
Fator de escala das conexões de entrada	0.5
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.5

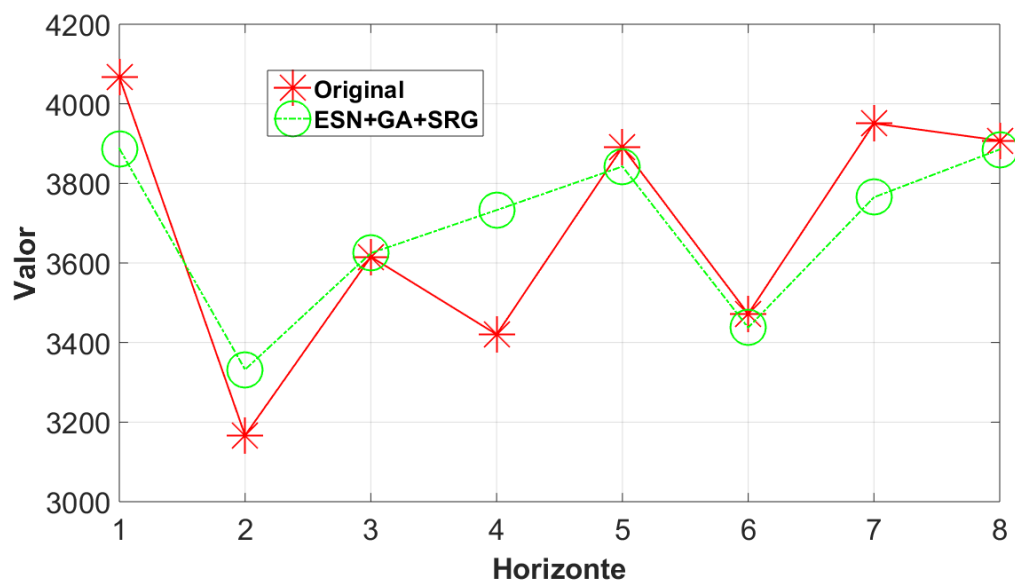


Figura 102. Resultados Série M3-0730 fora da amostra.

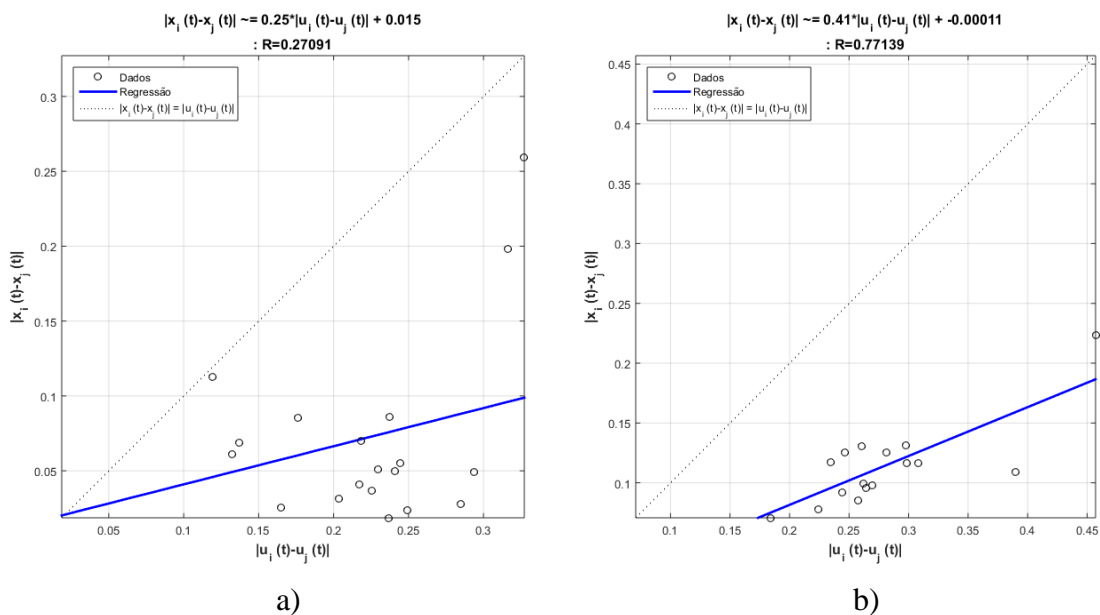


Figura 103. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 80. Parâmetros finais de configuração para a série M3-0758.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	198
Janela	8
Porcentagem de conexão	30%
Raio Espectral	0.9
Fator de escala das conexões de entrada	0.9
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.3

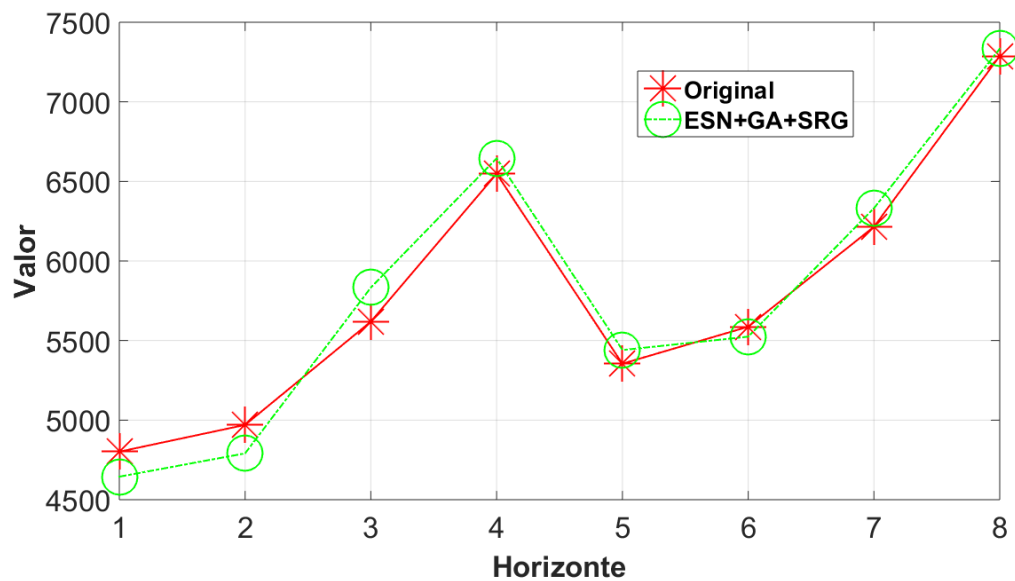


Figura 104. Resultados Série M3-0758 fora da amostra.

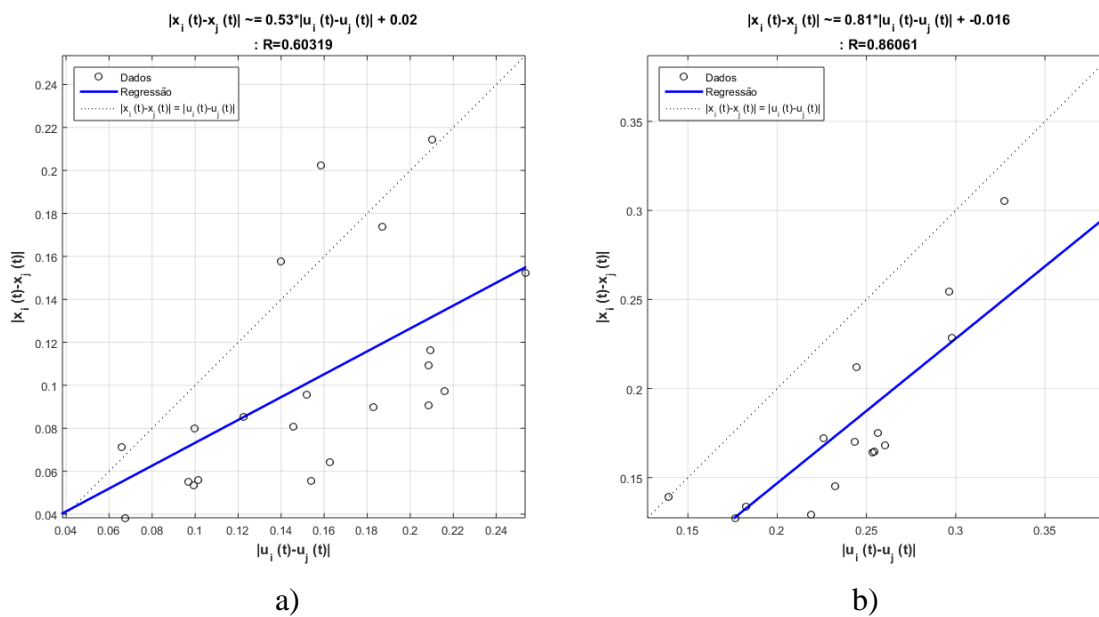


Figura 105. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 81. Parâmetros finais de configuração para a série M3-0773.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	429
Janela	3
Porcentagem de conexão	50%
Raio Espectral	0.4
Fator de escala das conexões de entrada	0.9
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.1

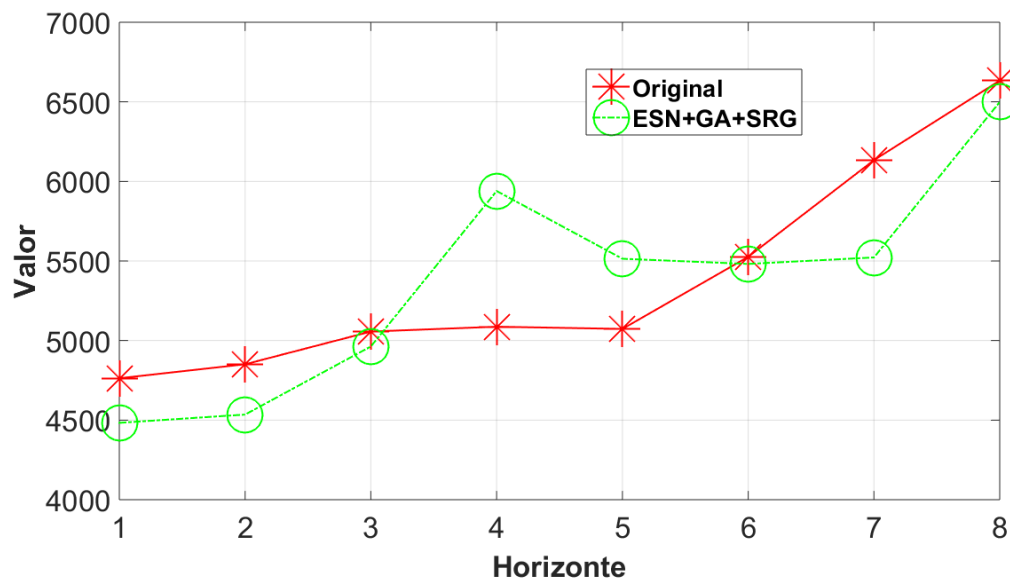


Figura 106. Resultados Série M3-0773 fora da amostra.

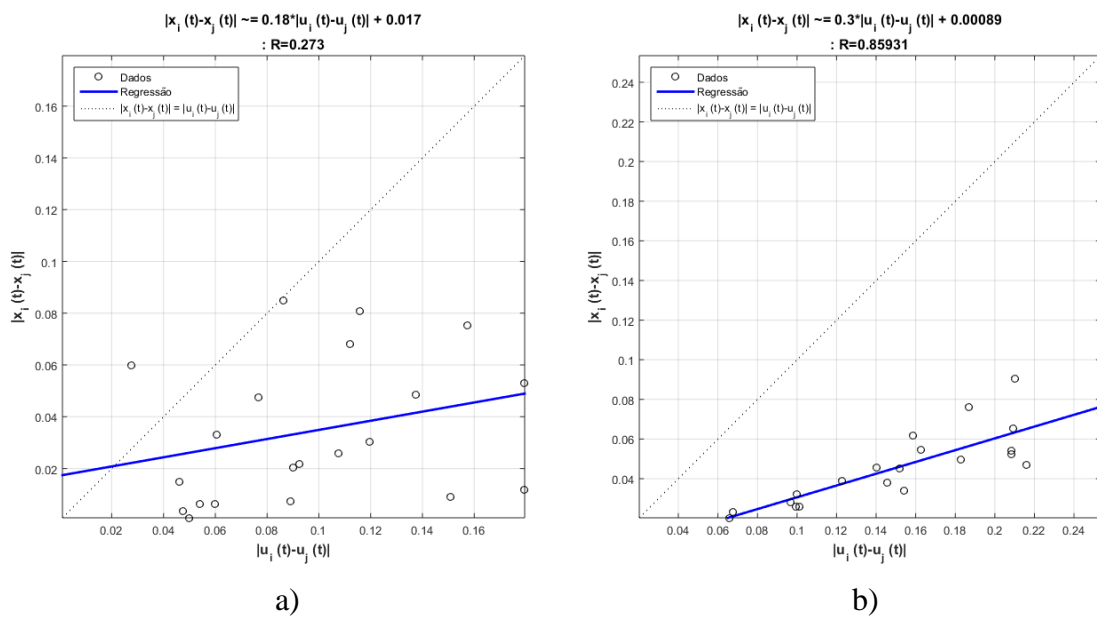


Figura 107. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 82. Parâmetros finais de configuração para a série M3-0829.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	346
Janela	9
Porcentagem de conexão	43%
Raio Espectral	0.9
Fator de escala das conexões de entrada	0.6
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.5

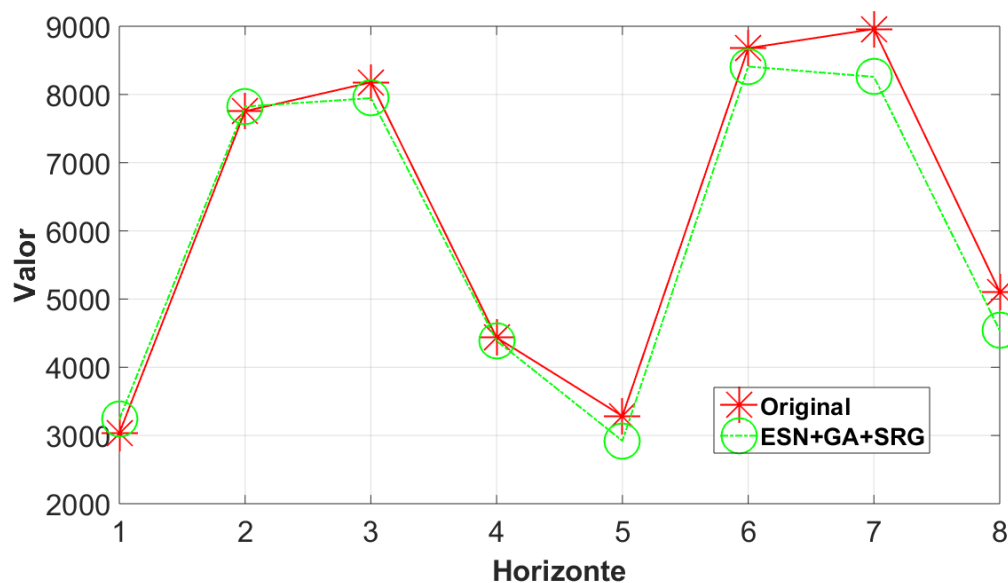


Figura 108. Resultados Série M3-0829 fora da amostra.

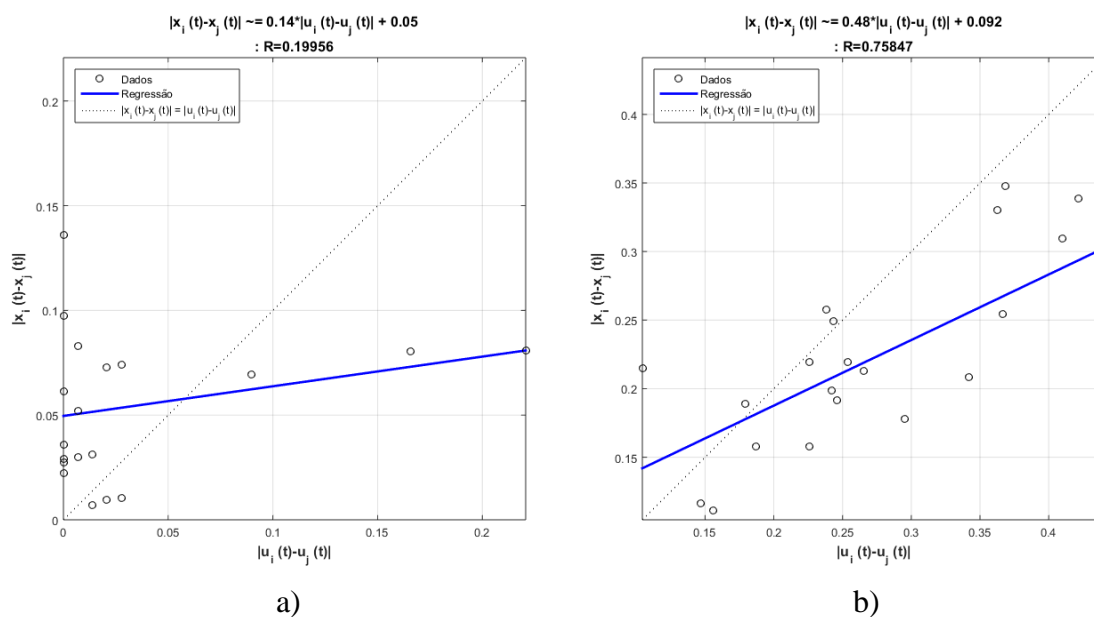


Figura 109. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 83. Parâmetros finais de configuração para a série M3-0847.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	216
Janela	4
Porcentagem de conexão	20%
Raio Espectral	0.3
Fator de escala das conexões de entrada	0.1
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.8

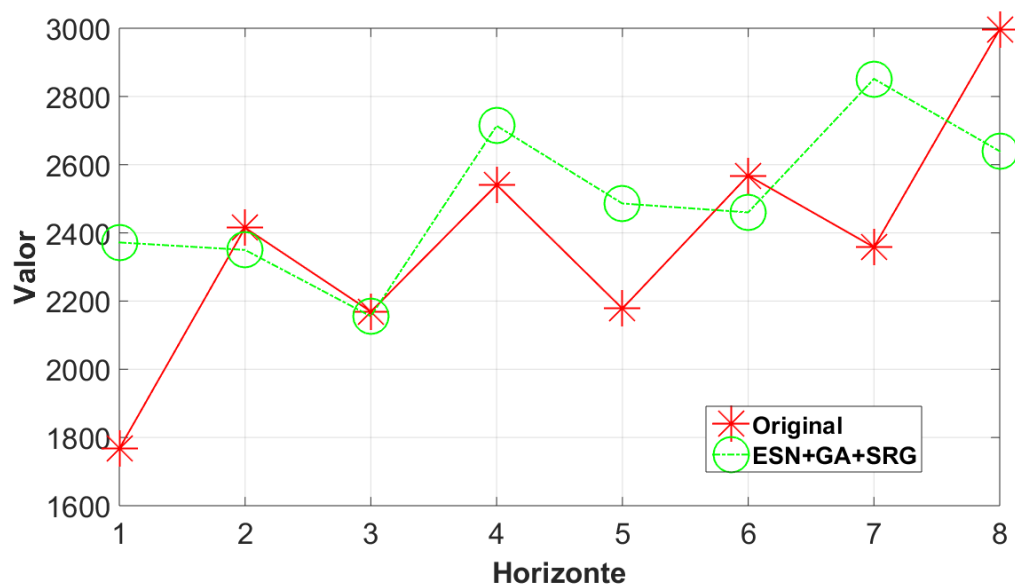


Figura 110. Resultados Série M3-0847 fora da amostra.

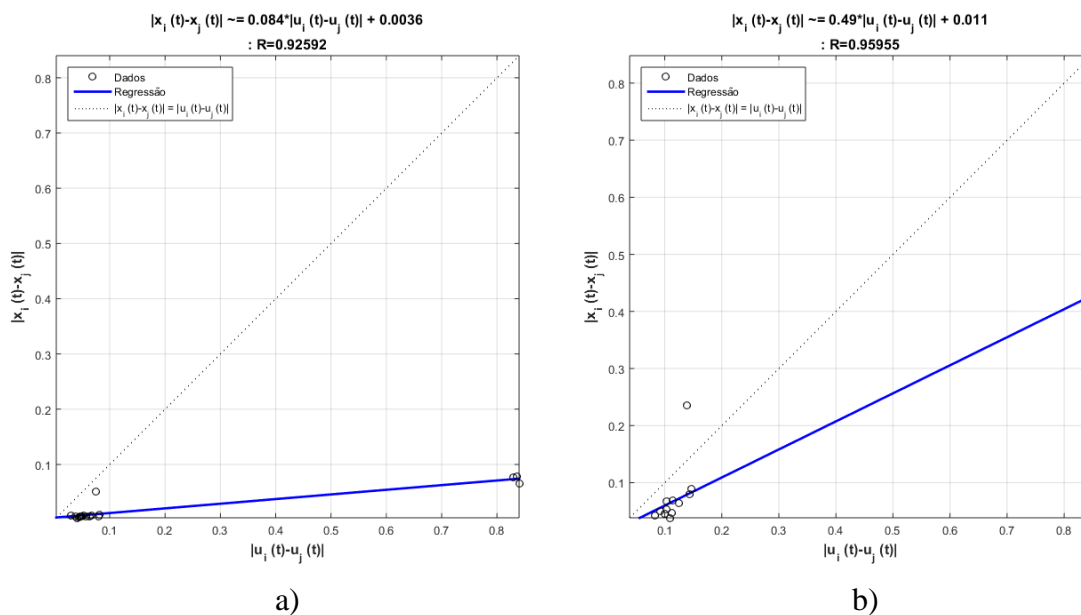


Figura 111. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 84. Parâmetros finais de configuração para a série M3-0854.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	336
Janela	5
Porcentagem de conexão	20%
Raio Espectral	1
Fator de escala das conexões de entrada	0.9
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.3

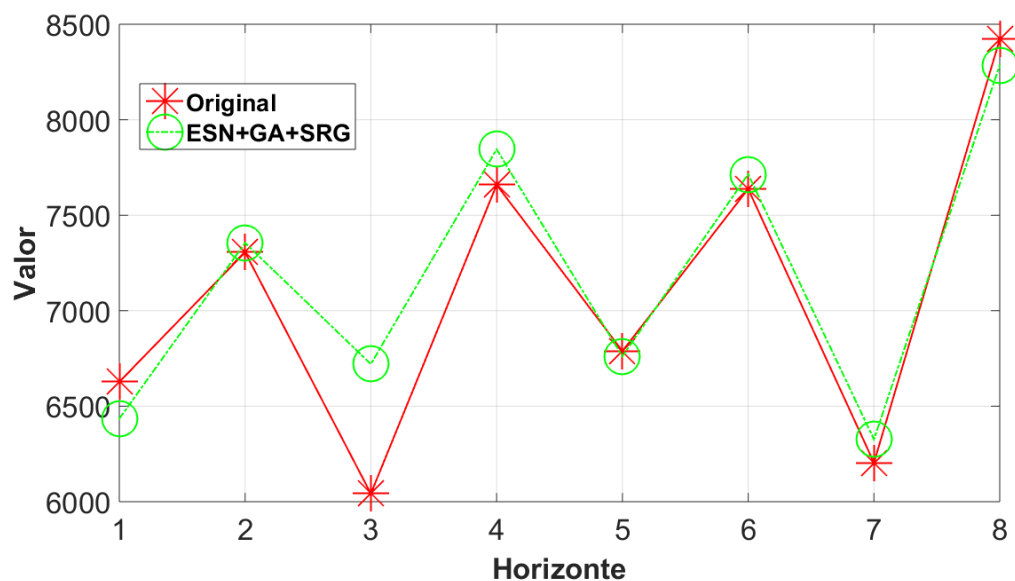


Figura 112. Resultados Série M3-0854 fora da amostra.

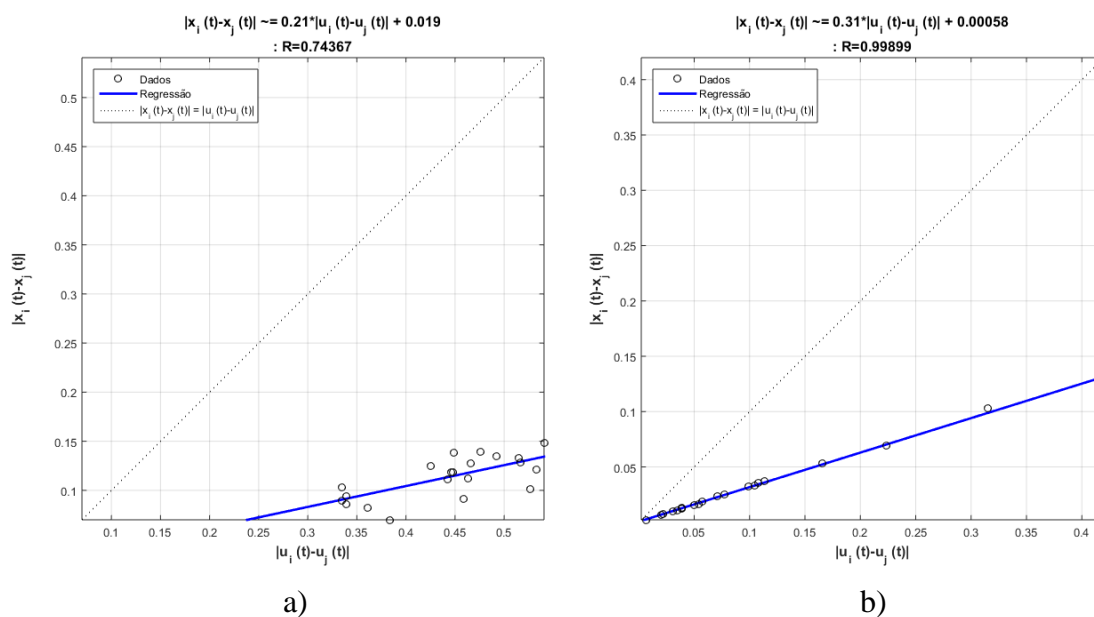


Figura 113. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 85. Parâmetros finais de configuração para a série M3-0865.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	264
Janela	9
Porcentagem de conexão	36%
Raio Espectral	0.8
Fator de escala das conexões de entrada	0.9
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.2

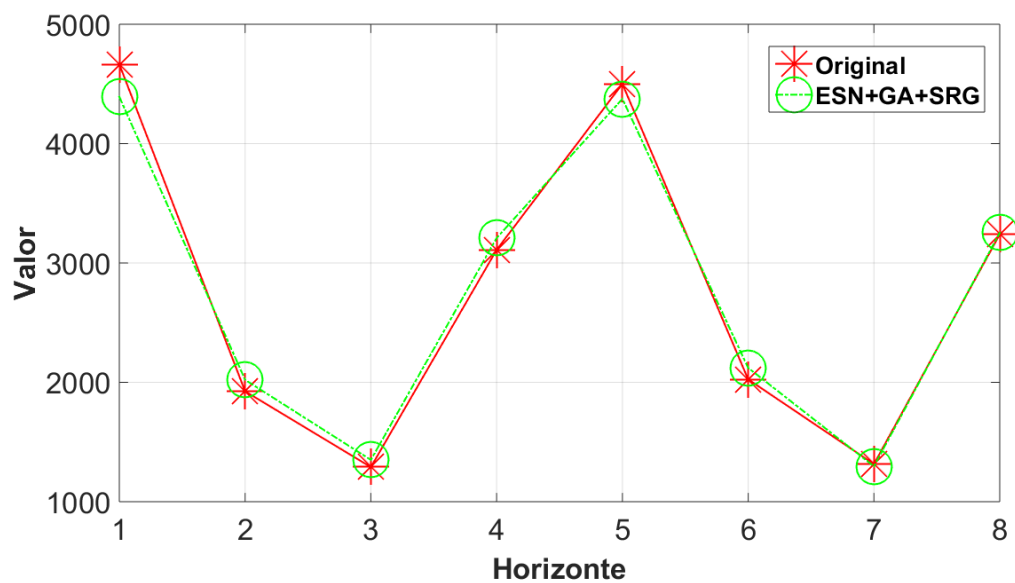


Figura 114. Resultados Série M3-0865 fora da amostra.

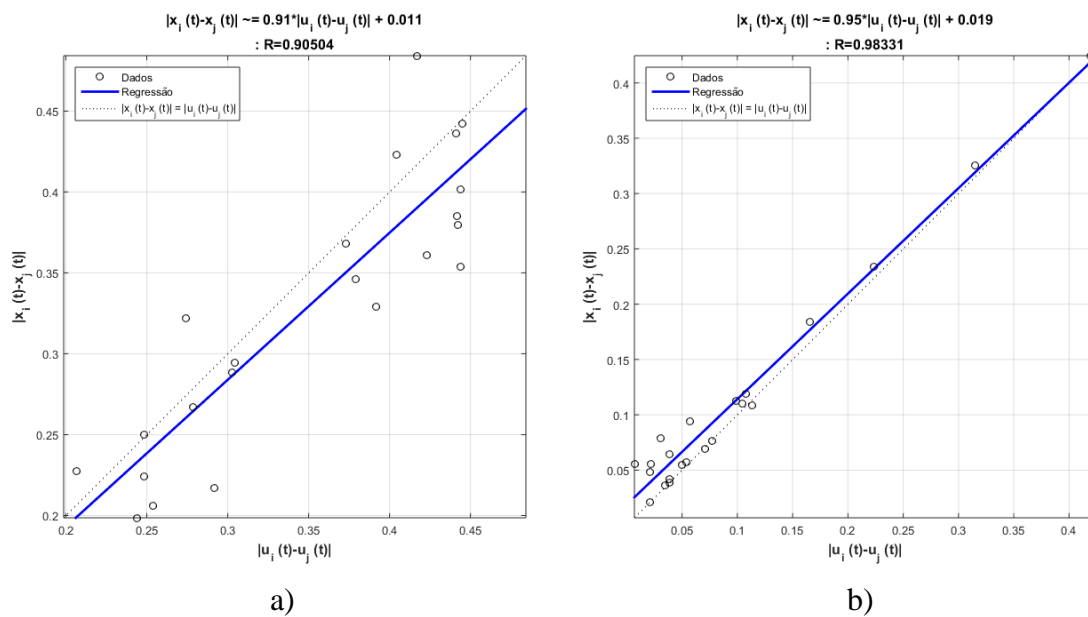


Figura 115. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 86. Parâmetros finais de configuração para a série M3-0896.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	256
Janela	11
Porcentagem de conexão	24%
Raio Espectral	0.7
Fator de escala das conexões de entrada	0.3
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.5

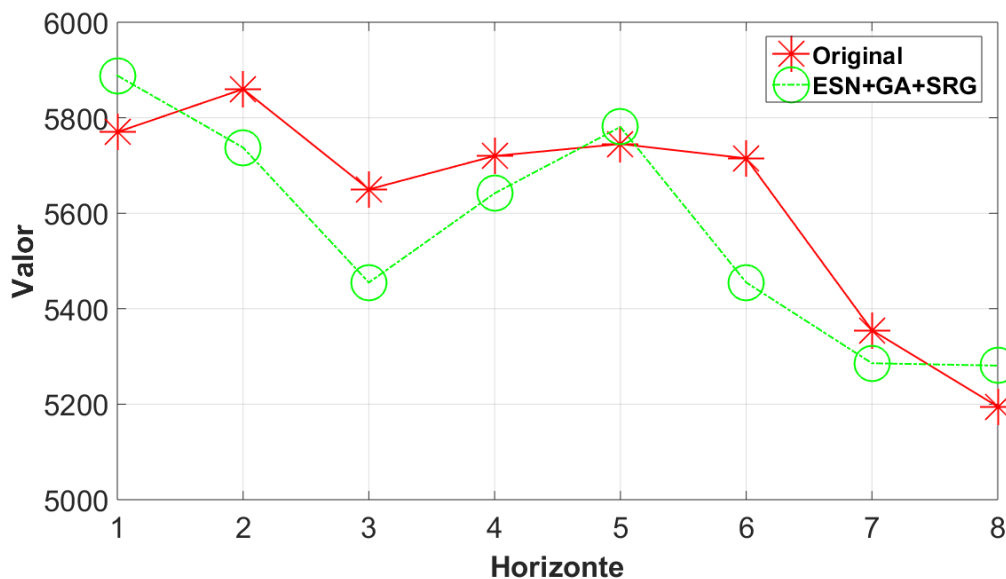


Figura 116. Resultados Série M3-0896 fora da amostra.

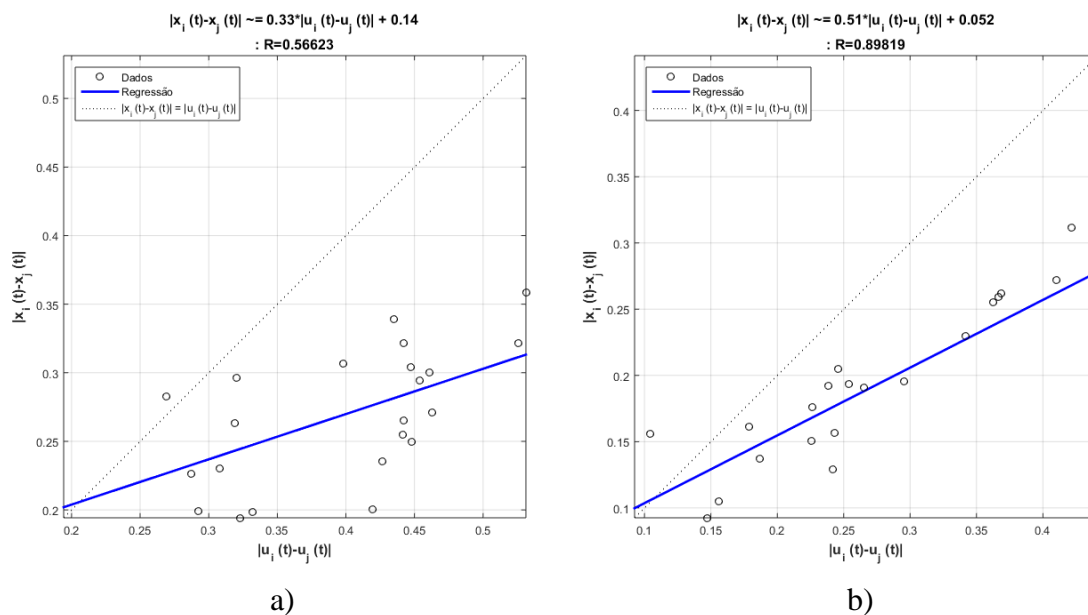


Figura 117. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Na figura 118 é apresentada a evolução do sMAPE para o grupo de amostragem trimestral. É possível observar que o modelo proposto manteve o tipo de comportamento observado no grupo anual, porém, desta vez os modelos selecionados para comparação tiveram resultados de sMAPE mais próximos ao desempenho do modelo proposto ESN-GA-SRG.

As respostas apresentadas em cada uma das previsões individuais pelo modelo proposto permitem concluir que a previsão obtida consegue acompanhar o comportamento de todas as séries pertencentes ao grupo de amostragem trimestral.

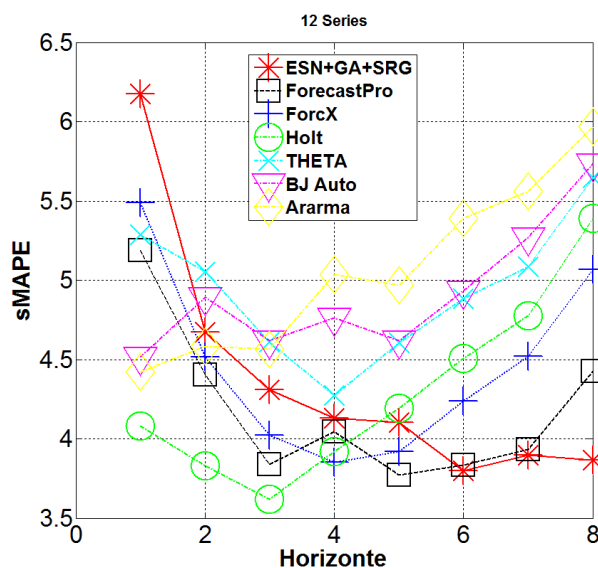


Figura 118. Evolução do SMAPE para cada modelo.

5.3.3.

Resultados para as séries do grupo mensal da competição M3

Para cada uma das séries, as tabelas 87 a 98, apresentam os parâmetros de configuração finais do modelo *ESN-GA-SRG* para as séries do grupo mensal e as figuras 119 a 142 apresentam os resultados da previsão do conjunto fora da amostra (teste) para o modelo proposto em cada uma das séries selecionadas e a regressão linear obtida para o *Reservoir* antes e depois de utilizar o método SRG

Tabela 87. Parâmetros finais de configuração para a série M3-1430.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	227
Janela	4
Percentagem de conexão	60%
Raio Espectral	0.1
Fator de escala das conexões de entrada	0.1
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.8

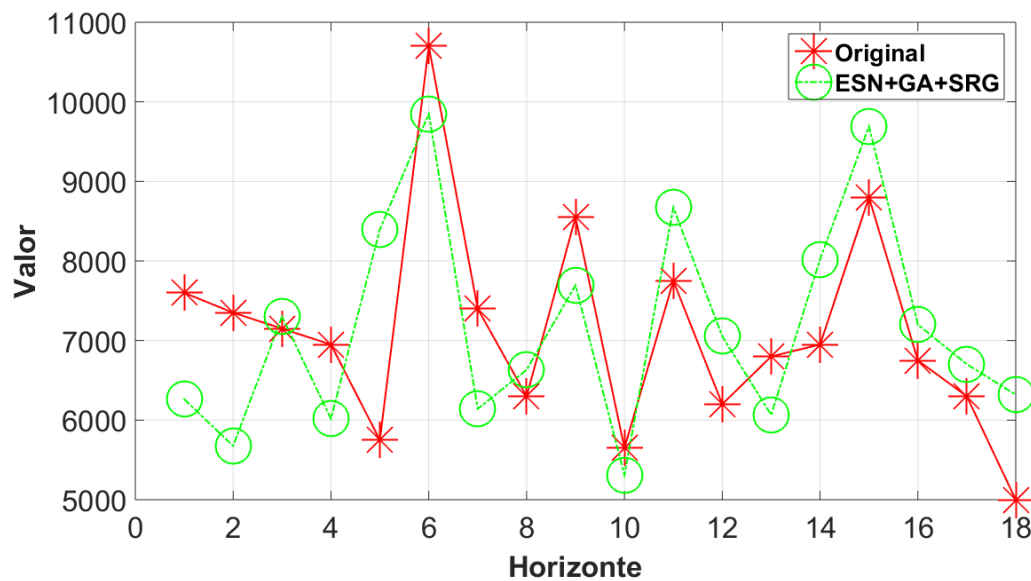


Figura 119. Resultados Série M3-1430 fora da amostra.

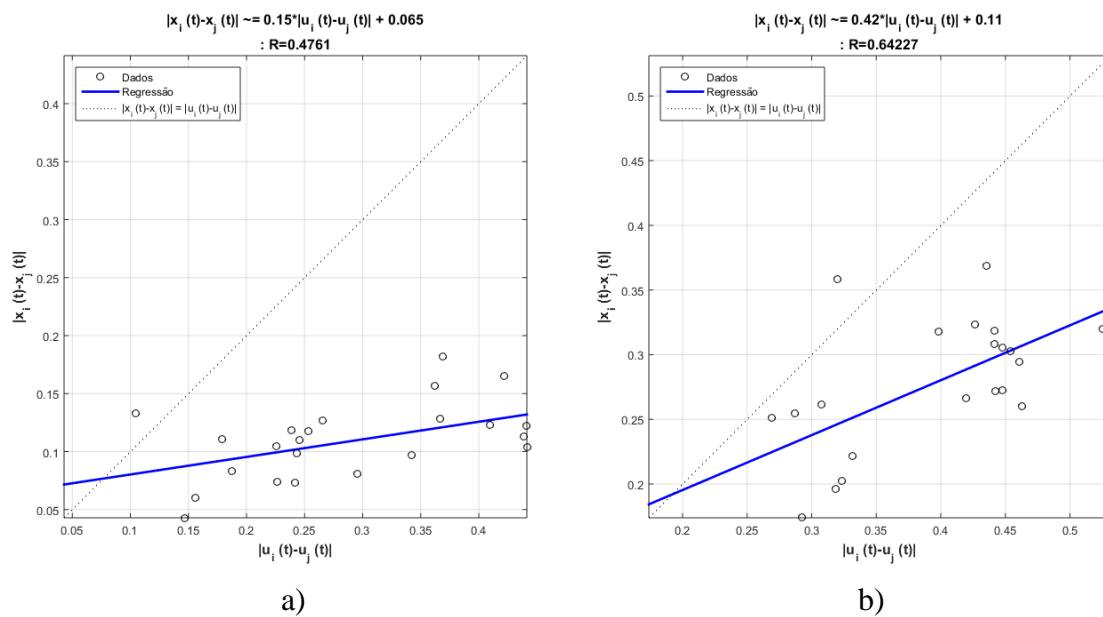


Figura 120. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 88. Parâmetros finais de configuração para a série M3-1460.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	362
Janela	3
Porcentagem de conexão	26%
Raio Espectral	0.9
Fator de escala das conexões de entrada	0.5
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.7

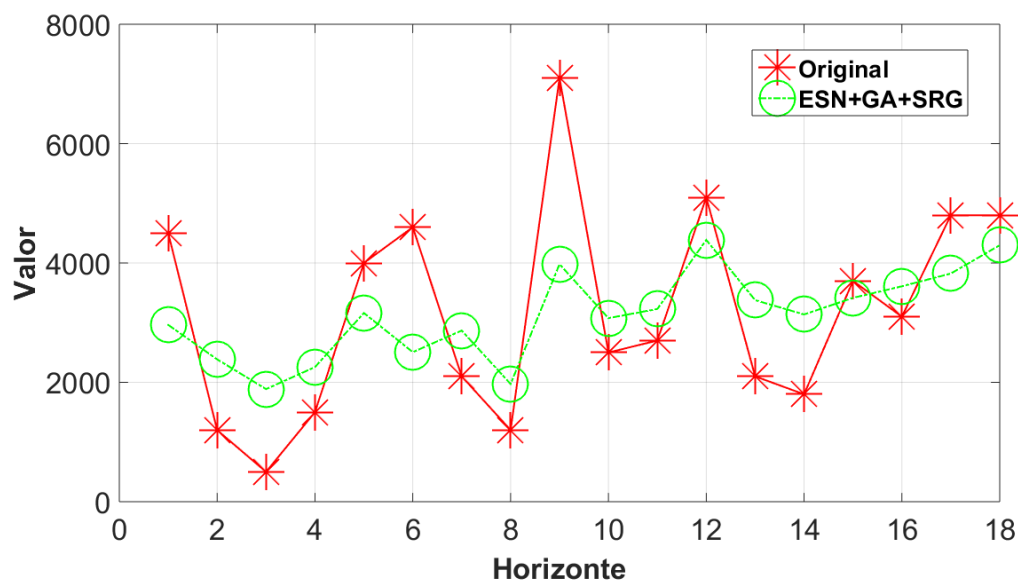


Figura 121. Resultados Série M3-1460 fora da amostra.

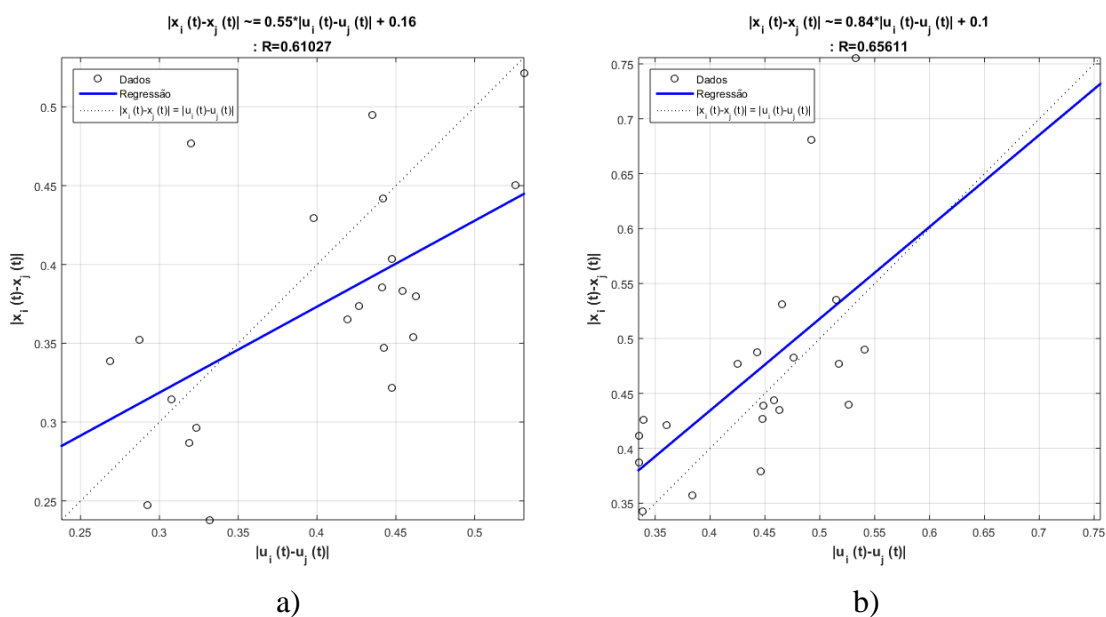


Figura 122. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 89. Parâmetros finais de configuração para a série M3-1725.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	462
Janela	12
Porcentagem de conexão	20%
Raio Espectral	0.4
Fator de escala das conexões de entrada	0.8
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.1

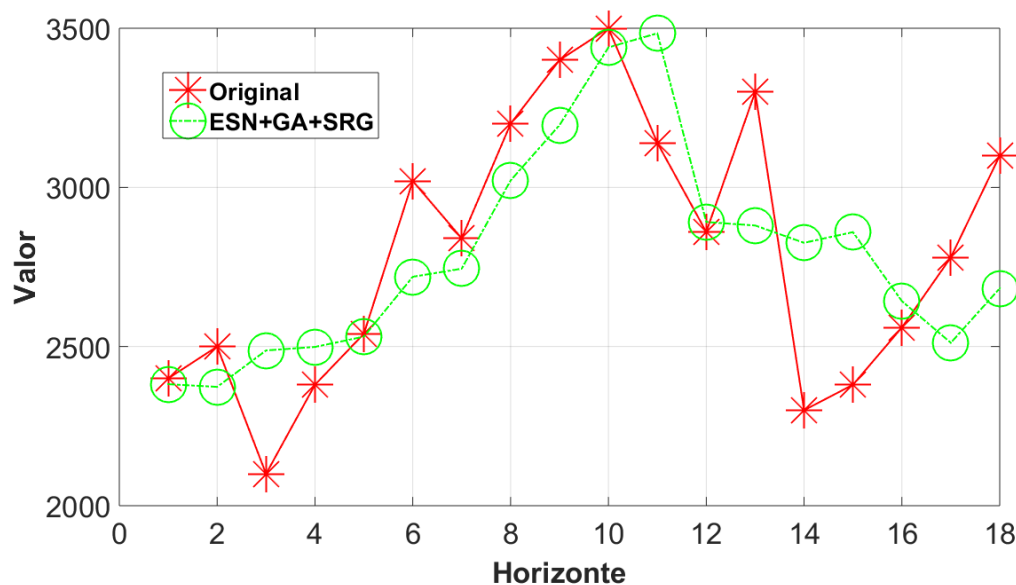


Figura 123. Resultados Série M3-1725 fora da amostra.

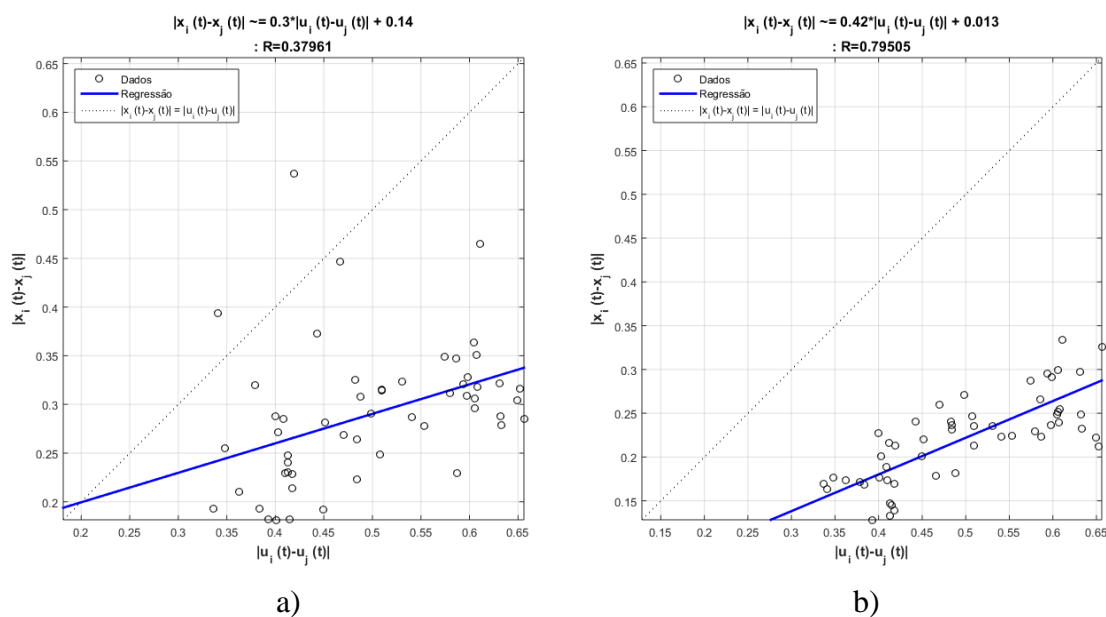


Figura 124. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 90. Parâmetros finais de configuração para a série M3-1900.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	294
Janela	10
Porcentagem de conexão	22%
Raio Espectral	0.7
Fator de escala das conexões de entrada	0.4
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.5

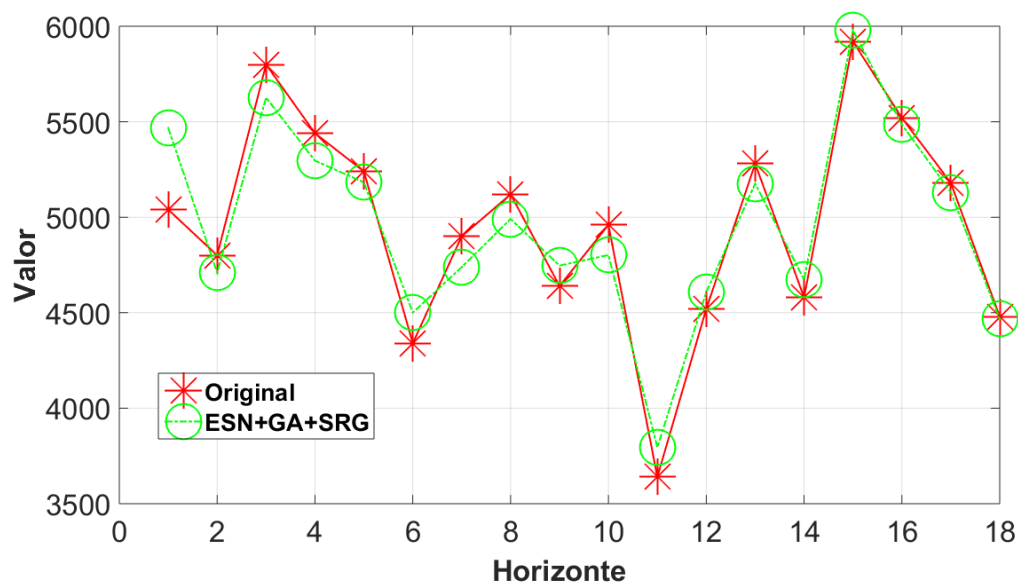


Figura 125. Resultados Série M3-1900 fora da amostra.

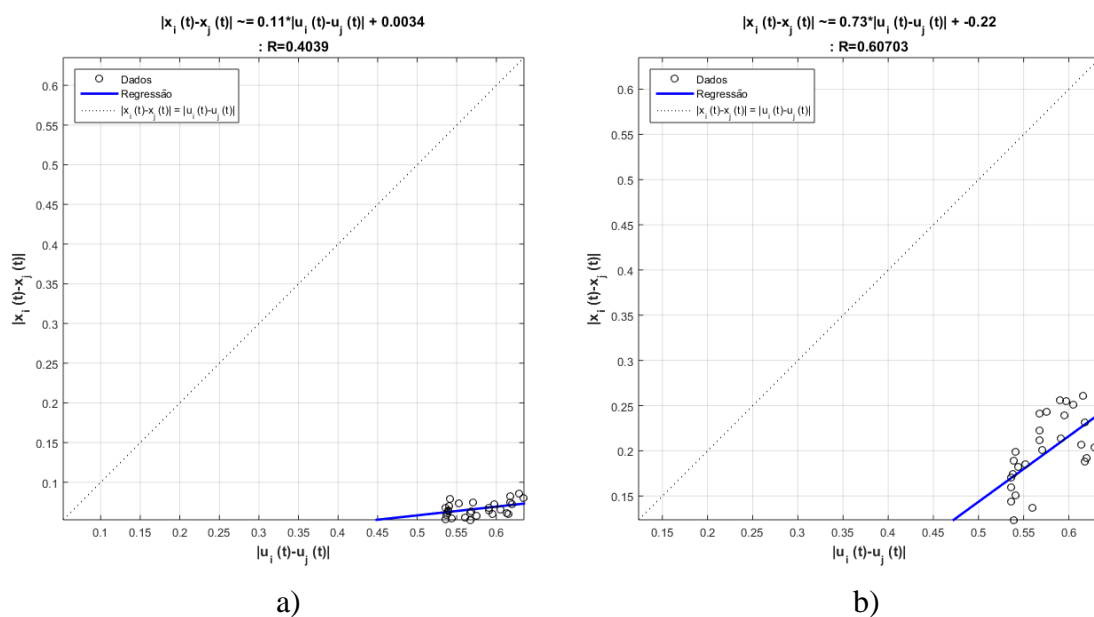


Figura 126. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 91. Parâmetros finais de configuração para a série M3-1906.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	468
Janela	10
Porcentagem de conexão	34%
Raio Espectral	0.6
Fator de escala das conexões de entrada	0.6
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.1

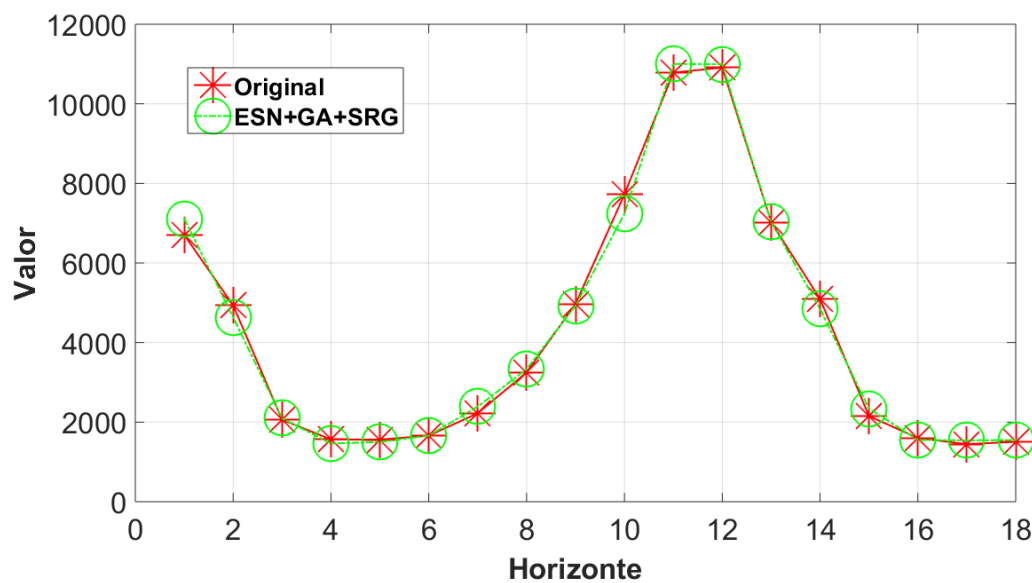


Figura 127. Resultados Série M3-1906 fora da amostra.

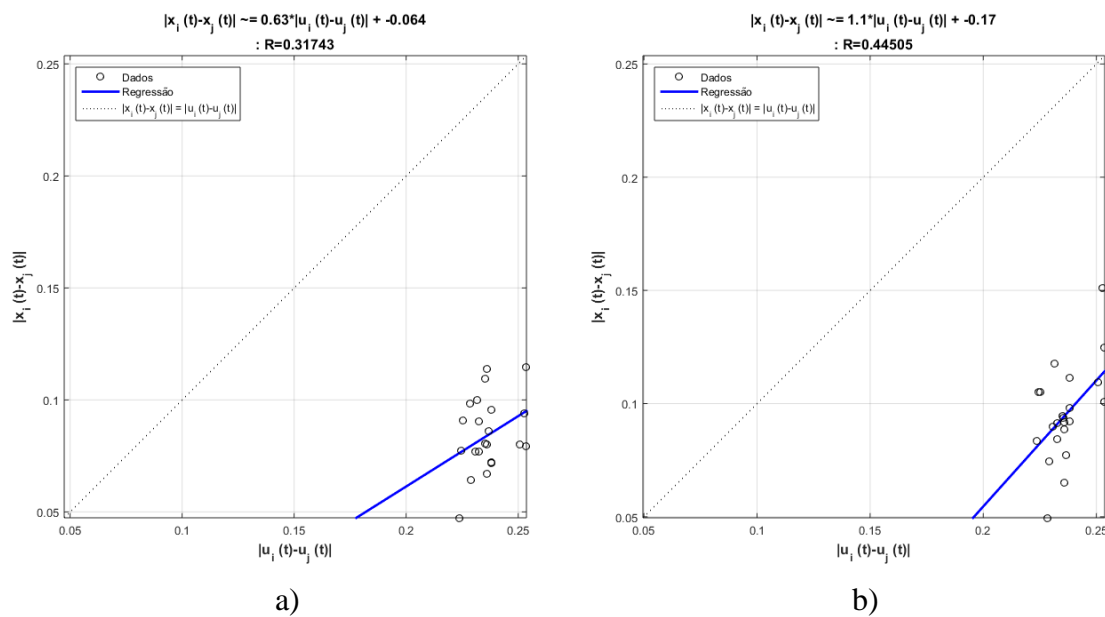


Figura 128. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 92. Parâmetros finais de configuração para a série M3-1912.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	338
Janela	5
Porcentagem de conexão	27%
Raio Espectral	0.6
Fator de escala das conexões de entrada	0.6
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.3

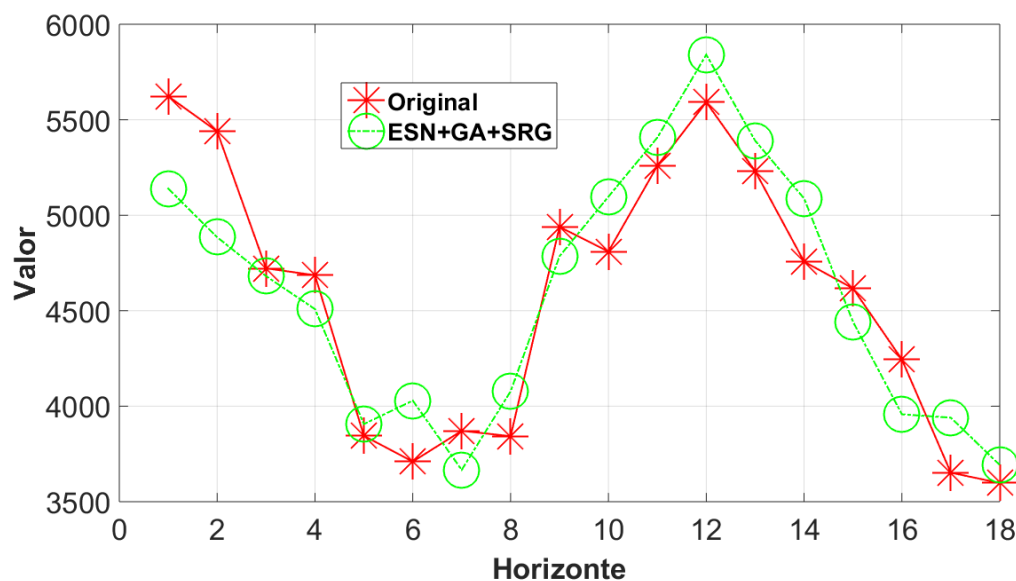


Figura 129. Resultados Série M3-1912 fora da amostra.

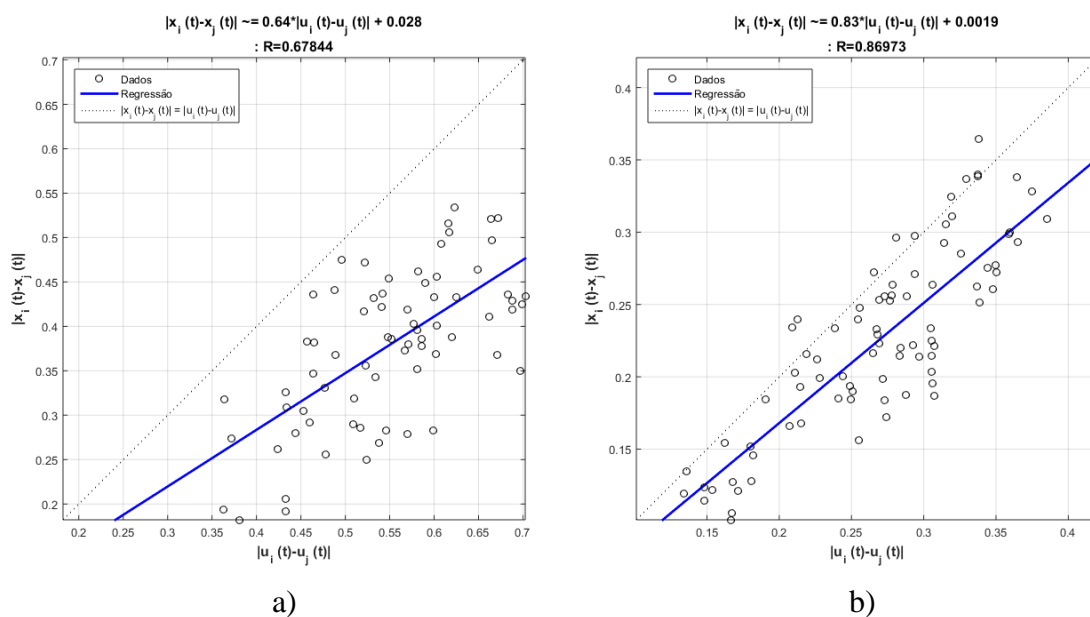


Figura 130. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 93. Parâmetros finais de configuração para a série M3-1996.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	299
Janela	12
Porcentagem de conexão	42%
Raio Espectral	0.3
Fator de escala das conexões de entrada	0.9
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0

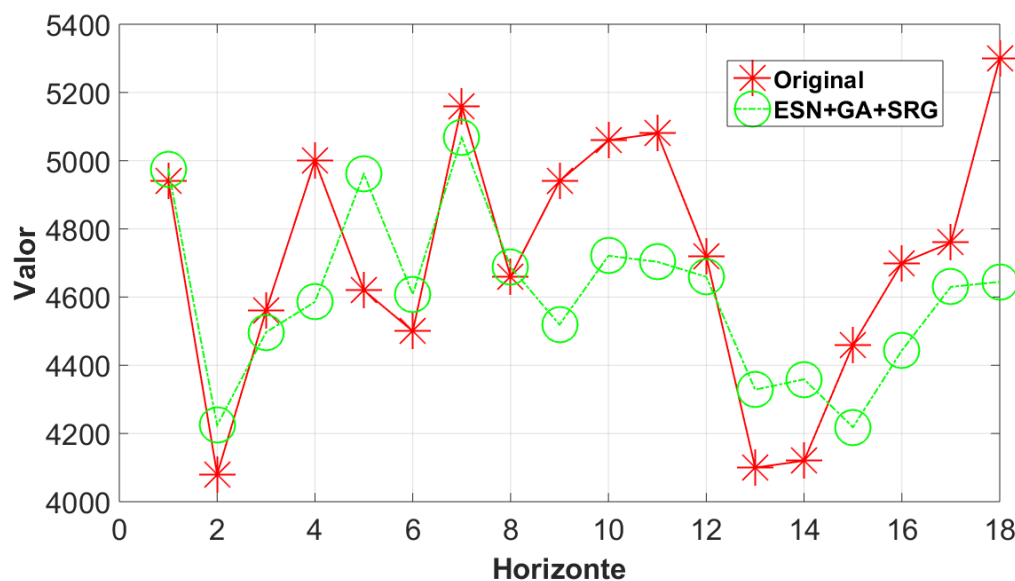


Figura 131. Resultados Série M3-1996 fora da amostra.

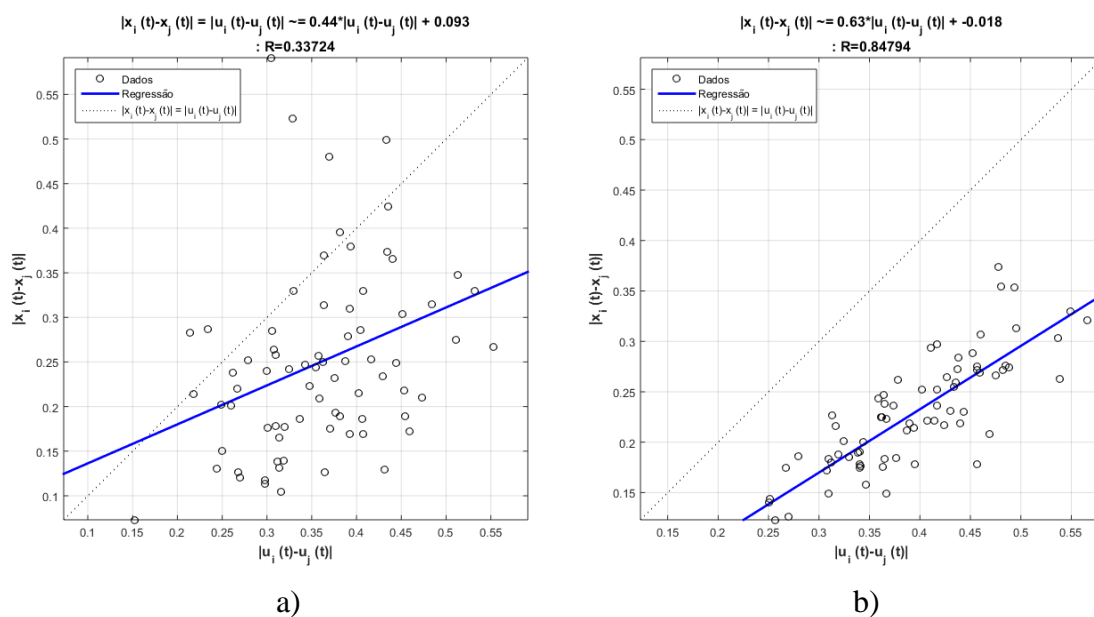


Figura 132. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 94. Parâmetros finais de configuração para a série M3-2001.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	477
Janela	7
Porcentagem de conexão	20%
Raio Espectral	1
Fator de escala das conexões de entrada	0.7
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.4

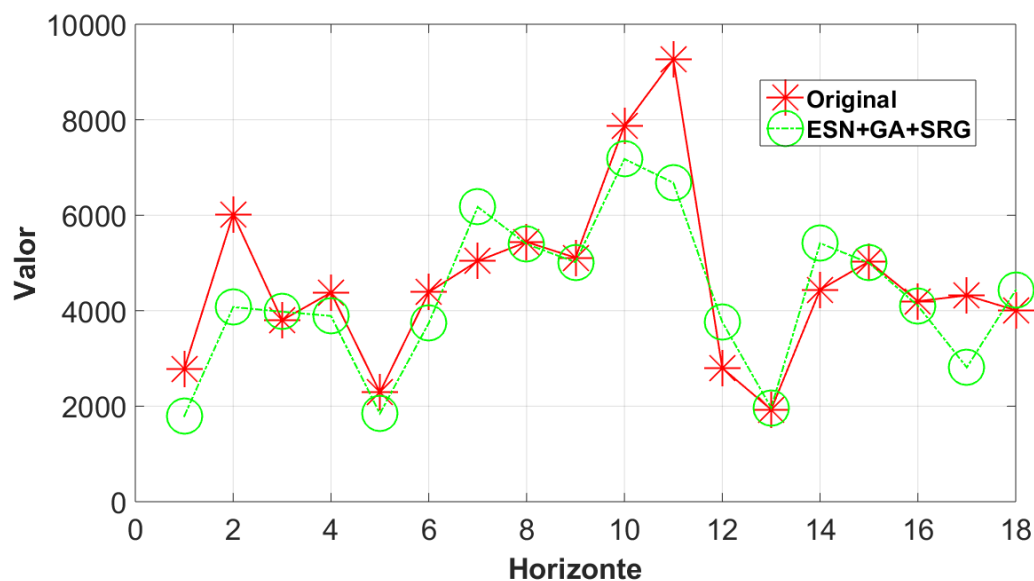


Figura 133. Resultados Série M3-2001 fora da amostra.

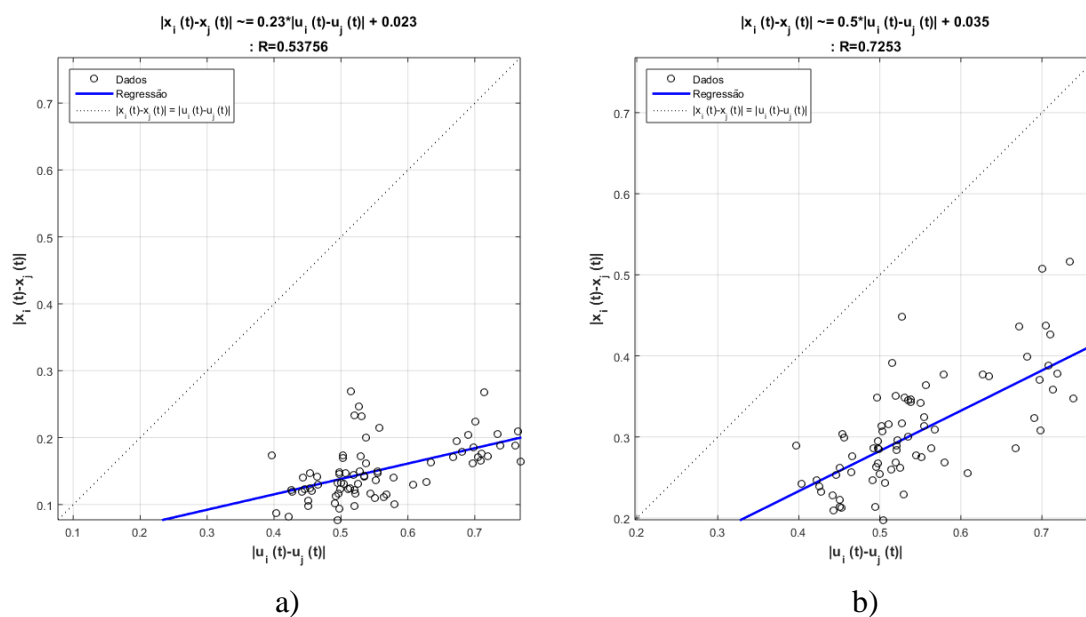


Figura 134. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 95. Parâmetros finais de configuração para a série M3-2013.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	451
Janela	8
Porcentagem de conexão	70%
Raio Espectral	8
Fator de escala das conexões de entrada	0.8
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.3

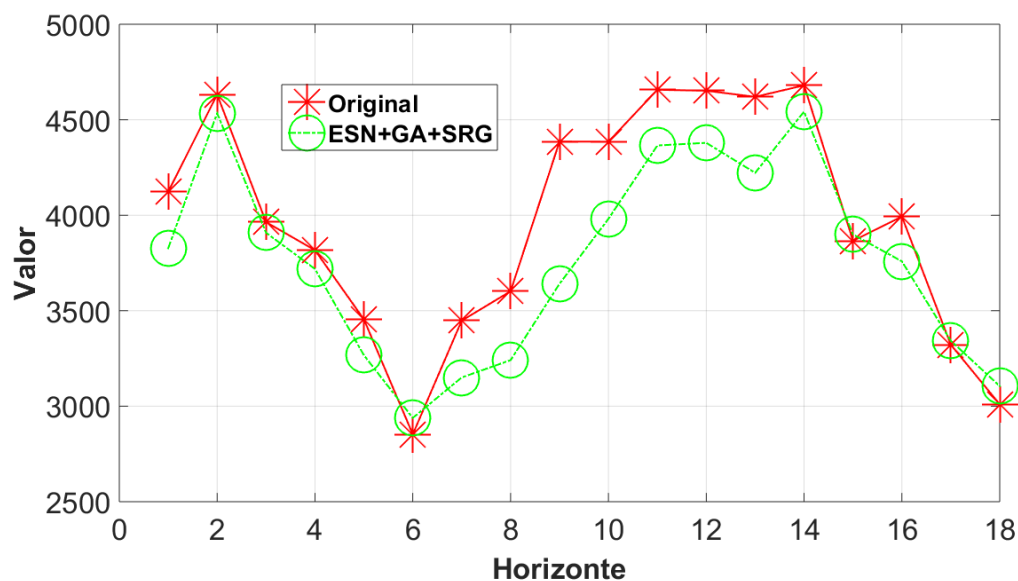


Figura 135. Resultados Série *M3-2013* fora da amostra.

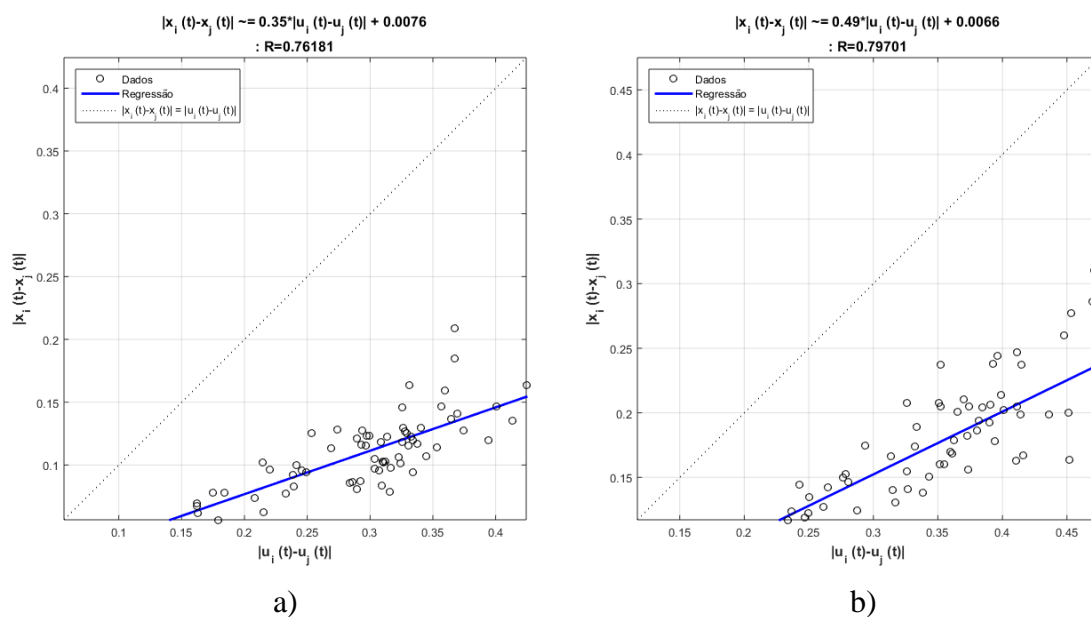


Figura 136. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 96. Parâmetros finais de configuração para a série *M3-2016*.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	356
Janela	8
Porcentagem de conexão	64%
Raio Espectral	1
Fator de escala das conexões de entrada	0.2
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.7

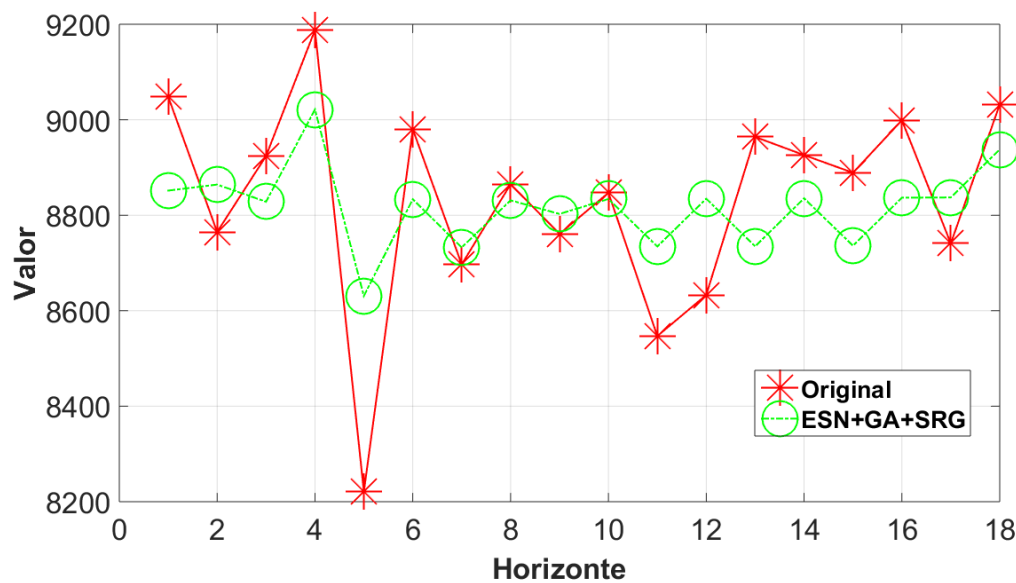


Figura 137. Resultados Série M3-2016 fora da amostra.

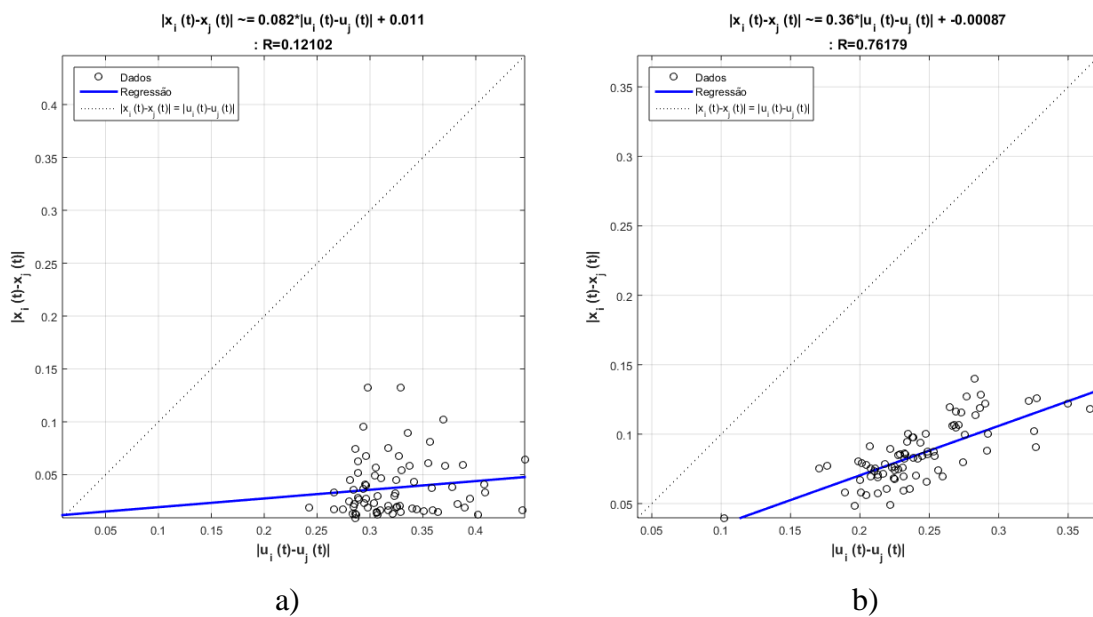


Figura 138. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 97. Parâmetros finais de configuração para a série M3-2089.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	187
Janela	7
Porcentagem de conexão	44%
Raio Espectral	0.4
Fator de escala das conexões de entrada	0.8
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.2

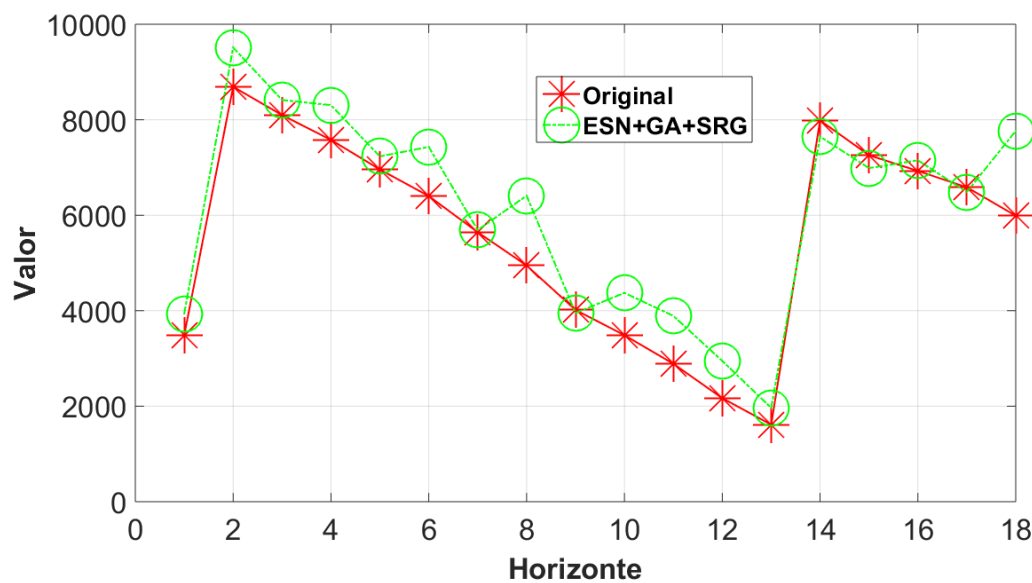


Figura 139. Resultados Série M3-2089 fora da amostra.

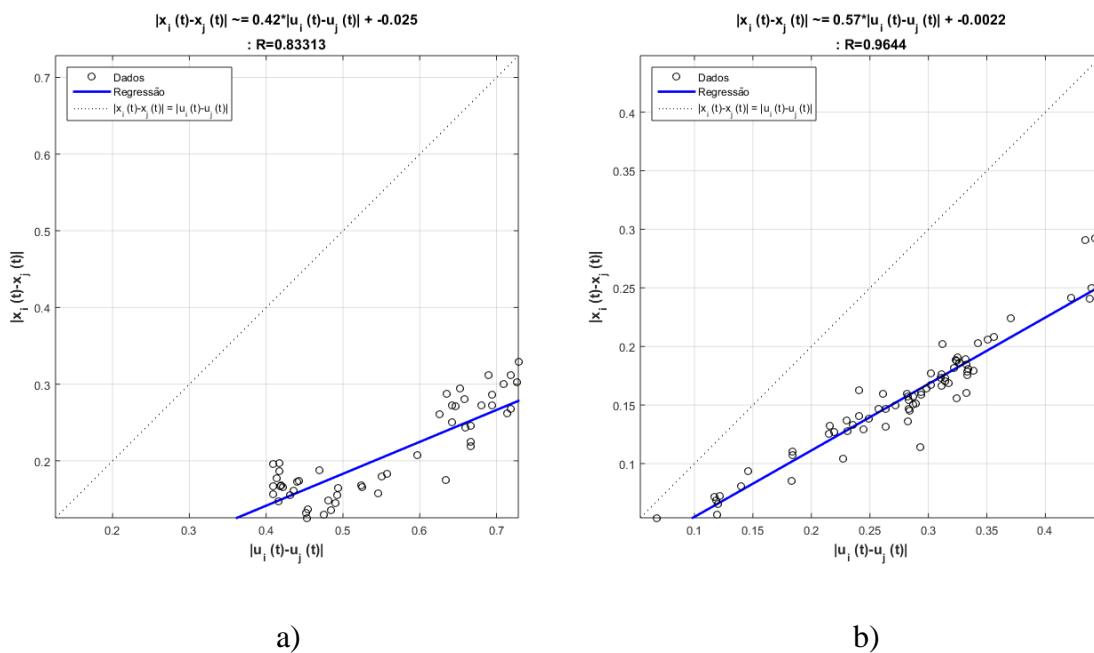


Figura 140. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Tabela 98. Parâmetros finais de configuração para a série M3-2579.

Parâmetro	Valor
Número de unidades no <i>Reservoir</i>	279
Janela	11
Porcentagem de conexão	20%
Raio Espectral	0.3
Fator de escala das conexões de entrada	0.4
Função de ativação dos neurônios do <i>Reservoir</i>	"Tanh"
Função de ativação do neurônio de Saída	"Linear"
Fator de escala das conexões recorrentes	0.5

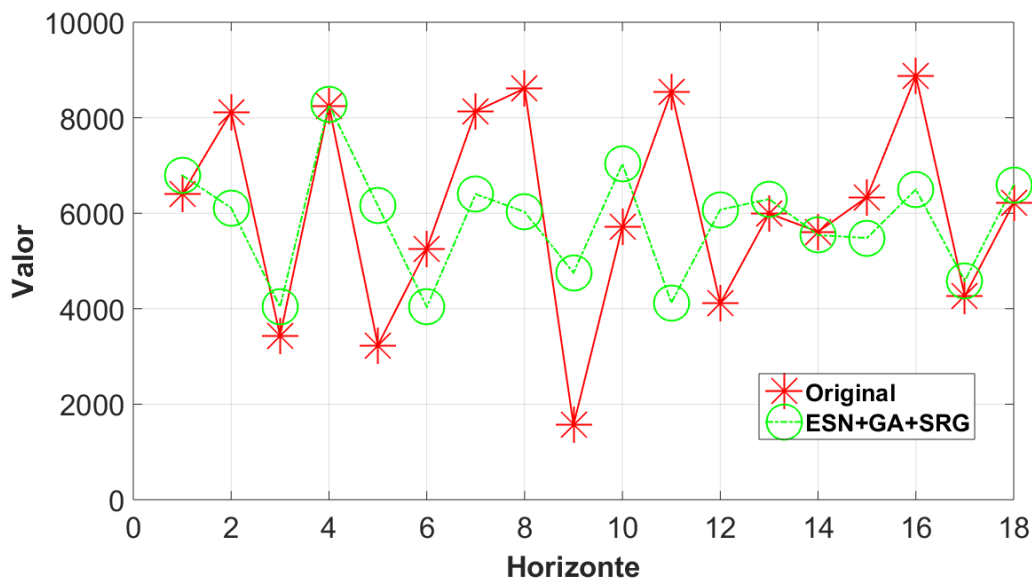


Figura 141. Resultados Série M3-2579 fora da amostra.

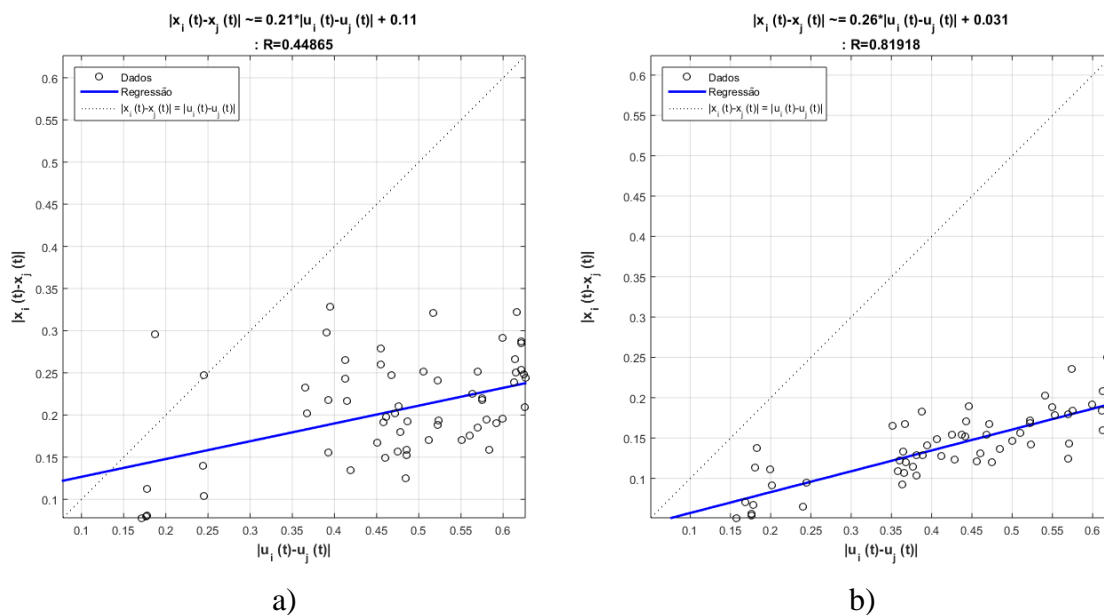


Figura 142. a) Regressão obtida antes do SRG. b) regressão obtida após do SRG.

Na figura 143 é apresentada a evolução do sMAPE para o grupo de amostragem mensal. É possível observar que para este grupo o modelo proposto teve um desempenho melhor do que os outros modelos selecionados. Neste cenário a métrica utilizada constatou que os erros obtidos conseguem melhorar em até 3 unidades a diferença das previsões.

De forma geral é possível constatar que o modelo proposto conseguiu acompanhar as previsões de cada uma das séries do grupo de amostragem mensal.

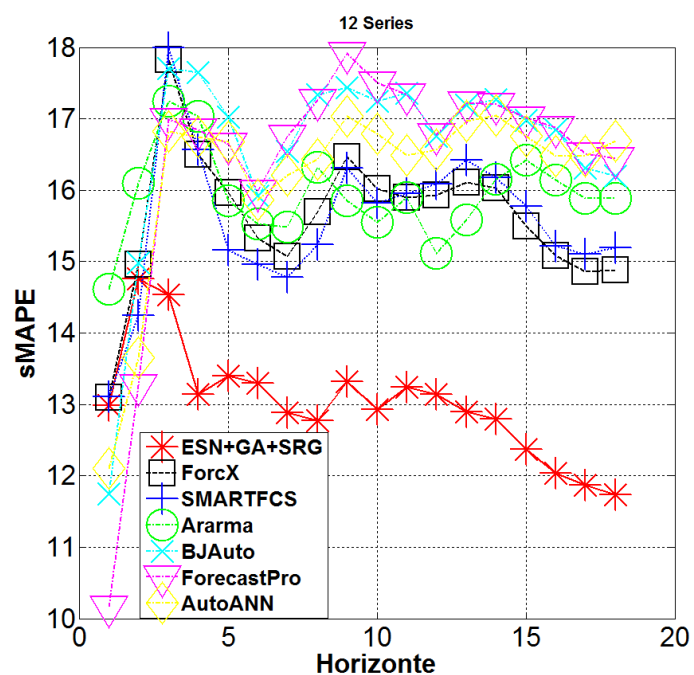


Figura 143. Evolução do sMAPE para cada modelo.

Nas figuras 144 a 146 são apresentados os resultados do sMAPE para cada série, permitindo comparar o desempenho obtido nas 12 séries de cada grupo de amostragem.

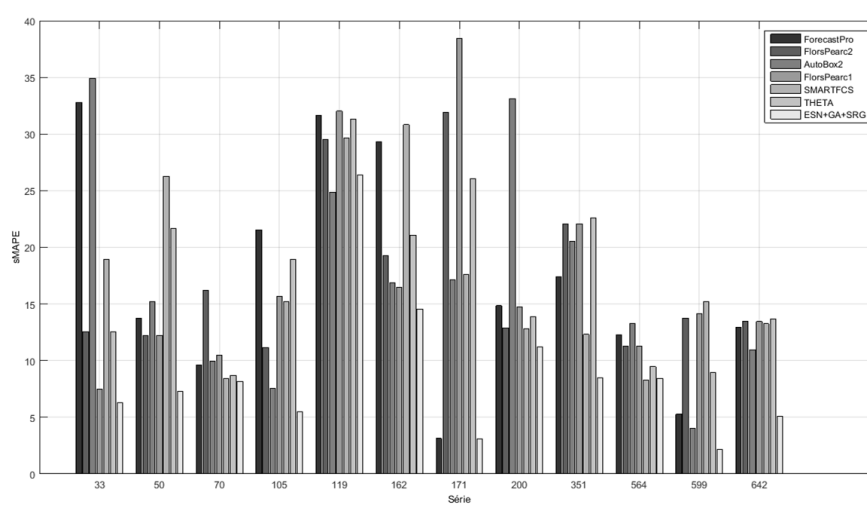


Figura 144. Diagrama de barras comparativas para as séries anuais.

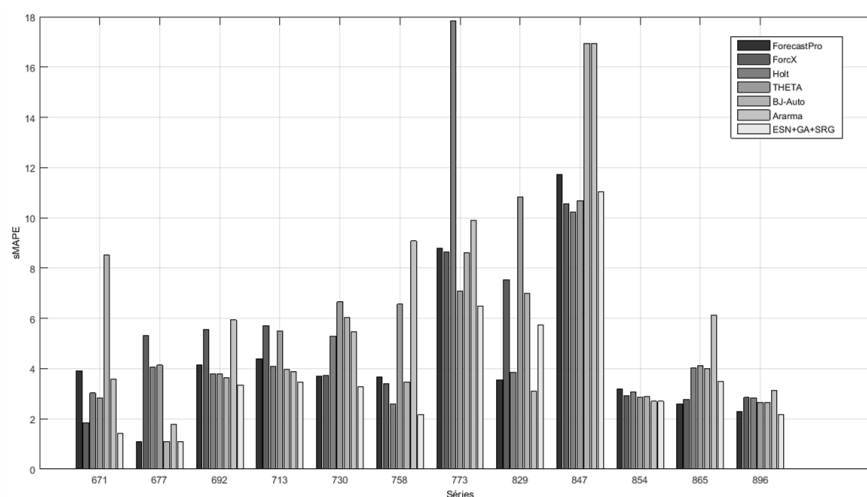


Figura 145. Diagrama de barras comparativas para as séries trimestrais.

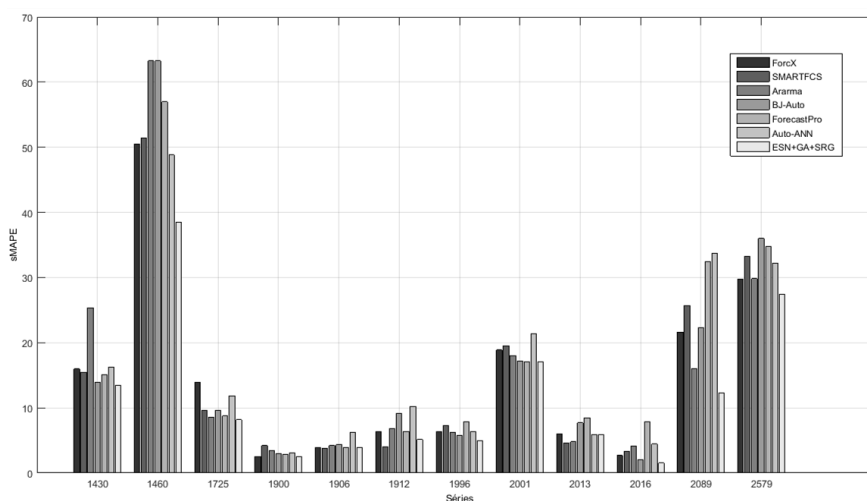


Figura 146. Diagrama de barras comparativas para as séries mensais.

5.3.4.

Resultados testes de significância estatística para as séries *M3*

Com o intuito de avaliar de forma estatística os resultados obtidos nas previsões *multi-step* para cada uma das 12 séries utilizadas, foram realizados os mesmos testes mencionados na seção 5.2.1:

- T-student
- Sync
- Wilcoxon
- Friedman
- Post-Hoc

Da mesma forma que no caso da competição NN3, o valor estabelecido para nível de significância, identificado como *p-value*, foi o 5%, e a comparação dos diferentes modelos é realizada utilizando o sMAPE acumulado para cada passo do horizonte de previsão.

Nas tabelas 99 a 101 são apresentados os resultados para os testes *t-student*, *Sync* e *Wilcoxon* para o grupo de amostragem anual, onde é possível constatar que o modelo proposto conseguiu ter diferenças estatisticamente significativas na maioria dos casos, quando a comparação é realizada par a par.

Tabela 99. Resultados do teste *t-student* para as séries M3 com amostragem anual.

Teste t				
Modelo	<i>p-valor</i>	Lim. Inferior	Lim. Superior	Conclusão
ForecastPro	9,58e-04	-8,8440	-4,0593	ESN+GA+SRG é melhor
Flors_Pearc2	0,0028	-8,9778	-3,2292	ESN+GA+SRG é melhor
Autobox2	0,0312	-9,1205	-0,6573	ESN+GA+SRG é melhor
Flors_Pearc1	0,0054	-8,9208	-2,6009	ESN+GA+SRG é melhor
SMARTFCS	0,0070	-9,1987	-2,4160	ESN+GA+SRG é melhor
THETA	0,0349	-9,0218	-0,5013	ESN+GA+SRG é melhor

Tabela 100. Resultados do teste *sync* para as séries M3 com amostragem anual

Teste sinal				
Modelo	<i>p-valor</i>	Lim. Inferior	Lim. Superior	Conclusão
ForecastPro	0,0313	-8,3458	-2,7967	ESN+GA+SRG é melhor
Flors_Pearc2	0,0313	-8,4055	-1,7118	ESN+GA+SRG é melhor
Autobox2	0,2188	-8,4942	1,1614	Equivalente
Flors_Pearc1	0,0313	-8,4968	-1,2894	ESN+GA+SRG é melhor
SMARTFCS	0,0313	-8,5349	-0,9250	ESN+GA+SRG é melhor
THETA	0,2188	-8,5150	0,8146	Equivalente

Tabela 101. Resultados do teste *Wilcoxon* para as séries M3 com amostragem anual

Teste Wilcoxon				
Modelo	<i>p-valor</i>	Lim. Inferior	Lim. Superior	Conclusão
ForecastPro	0,0313	-8,3458	-2,7967	ESN+GA+SRG é melhor
Flors_Pearc2	0,0313	-8,4055	-1,7118	ESN+GA+SRG é melhor
Autobox2	0,0938	-8,4942	1,1614	Equivalente
Flors_Pearc1	0,0313	-8,4968	-1,2894	ESN+GA+SRG é melhor
SMARTFCS	0,0313	-8,5349	-0,9250	ESN+GA+SRG é melhor
THETA	0,0938	-8,5150	0,8146	Equivalente

Na tabela 102 é apresentado o resultado obtido para o teste de Friedman, onde são comparados de forma simultânea os 6 modelos selecionados. É possível observar que o modelo proposto conseguiu obter a primeira posição na comparação com os outros seis modelos estatísticos.

Tabela 102. Resultados para o teste Friedman.

Teste Friedman							
Modelo	ESN+GA+SRG	ForecastPro	Flors_Pearc2	AutoBox2	Flors_Pearc1	SMARTFCS	THETA
h=1	2	7	6	1	5	4	3
h=2	2	7	6	4	5	3	1
h=3	1	7	6	3	4	5	2
h=4	1	7	5	4	3	6	2
h=5	1	3	5	2	6	7	4
h=6	1	2	3	4	5	6	7
Soma	8	33	31	18	28	31	19
Média	1,3333	5,5000	5,1667	3,0000	4,6667	5,1667	3,1667
Posição	1	7	5	2	4	5	3

Na tabela 103 é apresentado o resultado obtido para o teste de Post-Hoc, onde novamente são comparados todos os modelos de forma simultânea, mas, desta vez, é obtido o nível de significância para todos os modelos selecionados. Neste teste é possível observar que somente com a metade dos modelos comparados conseguiu obter diferenças estatisticamente significativas.

Tabela 103. Resultados para o teste Post-Hoc.

Teste Pos-Hoc						
Modelo	Modelo	Lim. Inferior	Dif. Médias	Lim. Superior	p-Valor	Conclusão
ESN+GA+SRG	ForecastPro	-7,8439	-4,1667	-0,4895	0,0147	ESN+GA+SRG é melhor
ESN+GA+SRG	Flors_Pearc2	-7,5105	-3,8333	-0,1561	0,0345	ESN+GA+SRG é melhor
ESN+GA+SRG	Autobox2	-5,3439	-1,6667	2,0105	0,8349	Equivalente
ESN+GA+SRG	Flors_Pearc1	-7,0105	-3,3333	0,3439	0,1052	Equivalente
ESN+GA+SRG	SMARTFCS	-7,5105	-3,8333	-0,1561	0,0345	ESN+GA+SRG é melhor
ESN+GA+SRG	THETA	-5,5105	-1,8333	1,8439	0,7628	Equivalente
ForecastPro	Flors_Pearc2	-3,343868	0,333333	4,010534	0,999970	Equivalente
ForecastPro	Autobox2	-1,177201	2,500000	6,177201	0,411687	Equivalente
ForecastPro	Flors_Pearc1	-2,843868	0,833333	4,510534	0,994278	Equivalente
ForecastPro	SMARTFCS	-3,343868	0,333333	4,010534	0,999970	Equivalente
ForecastPro	THETA	-1,343868	2,333333	6,010534	0,499857	Equivalente
Flors_Pearc2	Autobox2	-1,510534	2,166667	5,843868	0,590839	Equivalente
Flors_Pearc2	Flors_Pearc1	-3,177201	0,500000	4,177201	0,999680	Equivalente
Flors_Pearc2	SMARTFCS	-3,677201	0,000000	3,677201	1,000000	Equivalente
Flors_Pearc2	THETA	-1,677201	2,000000	5,677201	0,680078	Equivalente
Autobox2	Flors_Pearc1	-5,34387	-1,66667	2,01053	0,83492	Equivalente
Autobox2	SMARTFCS	-5,84387	-2,16667	1,51053	0,59084	Equivalente
Autobox2	THETA	-3,84387	-0,16667	3,51053	1,00000	Equivalente
Flors_Pearc1	SMARTFCS	-4,1772	-0,5000	3,1772	0,9997	Equivalente
Flors_Pearc1	THETA	-2,1772	1,5000	5,1772	0,8934	Equivalente
SMARTFCS	THETA	-1,6772	2,0000	5,6772	0,6801	Equivalente

Os resultados obtidos para os testes *t-student*, *Sync* e *Wilcoxon* são apresentados nas tabelas 104 a 106 para o grupo de amostragem trimestral, onde é

possível verificar que o modelo proposto se manteve equivalente a todos os modelos comparados.

Tabela 104. Resultados do teste *t-student* para as séries M3 com amostragem trimestral.

Teste t				
Modelo	<i>p-valor</i>	Lim. Inferior	Lim. Superior	Conclusão
ForecastPro	0,2736	-0,1876	0,5664	Equivalente
ForcX	0,7135	-0,5988	0,4321	Equivalente
Holt	0,8479	-0,8768	1,0379	Equivalente
THETA	0,0893	-1,2339	0,1119	Equivalente
B_J_Auto	0,1861	-1,4325	0,3361	Equivalente
Ararma	0,1561	-1,7219	0,3379	Equivalente

Tabela 105. Resultados do teste *Sync* para as séries M3 com amostragem trimestral.

Teste sinal				
Modelo	<i>p-valor</i>	Lim. Inferior	Lim. Superior	Conclusão
ForecastPro	0,7266	-0,5633	0,9890	Equivalente
ForcX	0,7266	-1,2028	0,6872	Equivalente
Holt	1	-1,5242	2,0962	Equivalente
THETA	0,0893	-1,2339	0,1119	Equivalente
B_J_Auto	0,1861	-1,4325	0,3361	Equivalente
Ararma	0,2891	-2,1050	1,7559	Equivalente

Tabela 106. Resultados do teste *Wilcoxon* para as séries M3 com amostragem trimestral.

Teste Wilcoxon				
Modelo	<i>p-valor</i>	Lim. Inferior	Lim. Superior	Conclusão
ForecastPro	0,3125	-0,2388	0,6298	Equivalente
ForcX	0,9453	-0,6214	0,4356	Equivalente
Holt	1	-0,8786	1,1535	Equivalente
THETA	0,0781	-1,1852	0,2532	Equivalente
B_J_Auto	0,1484	-1,3691	0,5735	Equivalente
Ararma	0,1953	-1,6631	0,4446	Equivalente

Na tabela 107 é apresentado o resultado obtido para o teste de Friedman, onde são comparados, de forma simultânea os 6 modelos selecionados. É possível observar que a colocação do modelo proposto nos primeiros passos prejudicou a posição final.

Tabela 107. Resultados para o teste Friedman.

Teste Friedman							
Modelo	ESN+GA+SRG	ForecastPro	ForcX	Holt	THETA	B_J_auto	Ararma
h=1	7	4	6	1	5	3	2
h=2	5	2	3	1	7	6	4
h=3	4	2	3	1	6	7	5
h=4	4	3	1	2	5	6	7
h=5	3	1	2	4	5	6	7
h=6	1	2	3	4	5	6	7
h=7	1	2	3	4	5	6	7
h=8	1	2	3	4	5	6	7
Soma	26	18	24	21	43	46	46
Média	3,2500	2,2500	3,0000	2,6250	5,3750	5,7500	5,7500
Posição	4	1	3	2	5	6	6

Os resultados para o teste de Post-Hoc são apresentados na tabela 108, onde são comparados todos os modelos de forma simultânea, para o modelo proposto os valores de significância obtidos foram maiores do que o limiar estabelecido (5%) por tanto não existe uma diferença estática significativa com os 6 modelos comparados.

Tabela 108. Resultados para o teste Post-Hoc.

Teste Pos-Hoc						
Modelo	Modelo	Lim. Inferior	Dif. Médias	Lim. Superior	<i>p</i> -Valor	Conclusão
ESN+GA+SRG	ForecastPro	-2,1845	1,0000	4,1845	0,9685	Equivalente
ESN+GA+SRG	ForcX	-2,9345	0,2500	3,4345	1,0000	Equivalente
ESN+GA+SRG	Holt	-2,5595	0,6250	3,8095	0,9974	Equivalente
ESN+GA+SRG	THETA	-5,3095	-2,1250	1,0595	0,4356	Equivalente
ESN+GA+SRG	B_J_Auto	-5,6845	-2,5000	0,6845	0,2368	Equivalente
ESN+GA+SRG	Arama	-5,6845	-2,5000	0,6845	0,2368	Equivalente
ForecastPro	ForcX	-3,9345	-0,7500	2,4345	0,9929	Equivalente
ForecastPro	Holt	-3,5595	-0,3750	2,8095	0,9999	Equivalente
ForecastPro	THETA	-6,3095	-3,1250	0,0595	0,0585	Equivalente
ForecastPro	B_J_Auto	-6,6845	-3,5000	-0,3155	0,0204	ForecastPro é melhor
ForecastPro	Arama	-6,6845	-3,5000	-0,3155	0,0204	ForecastPro é melhor
ForcX	Holt	-2,8095	0,3750	3,5595	0,9999	Equivalente
ForcX	THETA	-5,5595	-2,3750	0,8095	0,2960	Equivalente
ForcX	B_J_Auto	-5,9345	-2,7500	0,4345	0,1431	Equivalente
ForcX	Arama	-5,9345	-2,7500	0,4345	0,1431	Equivalente
Holt	THETA	-5,9345	-2,7500	0,4345	0,1431	Equivalente
Holt	B_J_Auto	-6,3095	-3,1250	0,0595	0,0585	Equivalente
Holt	Arama	-6,3095	-3,1250	0,0595	0,0585	Equivalente
THETA	B_J_Auto	-3,5595	-0,3750	2,8095	0,9999	Equivalente
THETA	Arama	-3,5595	-0,3750	2,8095	0,9999	Equivalente
B_J_Auto	Arama	-3,1845	0,0000	3,1845	1,0000	Equivalente

Para as séries mensais os resultados dos testes *t-student*, *Sync* e *Wilcoxon* são apresentados nas tabelas 109 a 111, onde o modelo proposto conseguiu confirmar que é significativamente diferente dos modelos comparados.

Tabela 109. Resultados do teste *t-student* para as séries M3 com amostragem mensal.

Teste t				
Modelo	<i>p</i> -valor	Lim. Inferior	Lim. Superior	Conclusão
ForcX	1,92e-09	-3,0984	-2,1377	ESN+GA+SRG é melhor
SMARTFCS	4,57e-08	-3,1364	-1,9744	ESN+GA+SRG é melhor
Arama	8,38e-11	-3,3591	-2,4837	ESN+GA+SRG é melhor
B_J_Auto	3,57e-08	-4,3768	-2,7779	ESN+GA+SRG é melhor
ForecastPro	3,79e-06	-4,4304	-2,3061	ESN+GA+SRG é melhor
Auto_ANN	2,87e-03	-4,0375	-2,3756	ESN+GA+SRG é melhor

Tabela 110. Resultados do teste *Sync* para as séries M3 com amostragem mensal.

Teste sinal				
Modelo	<i>p</i> -valor	Lim. Inferior	Lim. Superior	Conclusão
ForcX	7,62e-06	-3,1499	-2,5713	ESN+GA+SRG é melhor
SMARTFCS	1,44e-04	-3,4051	-1,8988	ESN+GA+SRG é melhor
Ararma	7,62e-06	-3,8944	-2,4571	ESN+GA+SRG é melhor
B_J_Auto	1,44e-04	-4,4714	-3,6137	ESN+GA+SRG é melhor
ForecastPro	0,0013	-4,5996	-3,2518	ESN+GA+SRG é melhor
Auto_ANN	0,0013	-4,2527	-3,2471	ESN+GA+SRG é melhor

Tabela 111. Resultados do teste *Wilcoxon* para as séries M3 com amostragem mensal.

Teste Wilcoxon				
Modelo	<i>p</i> -valor	Lim. Inferior	Lim. Superior	Conclusão
ForcX	1,80e-04	-3,1230	-2,3463	ESN+GA+SRG é melhor
SMARTFCS	2,53e-04	-3,2027	-1,8308	ESN+GA+SRG é melhor
Ararma	1,80e-04	-3,3559	-2,4642	ESN+GA+SRG é melhor
B_J_Auto	2,53e-04	-4,3875	-3,1440	ESN+GA+SRG é melhor
ForecastPro	4,19e-04	-4,4642	-2,9733	ESN+GA+SRG é melhor
Auto_ANN	3,00e-04	-4,0240	-2,7882	ESN+GA+SRG é melhor

Na tabela 112 é apresentado o resultado obtido para o teste de Friedman, é possível observar que somente nos dois primeiros passos o modelo proposto no conseguiu a primeira colocação, mas ao longo do horizonte de previsão alcançou a melhor posição para terminar como o melhor colocado.

Tabela 112. Resultados para o teste Friedman.

Teste Friedman							
Modelo	ESN+GA+SRG	ForcX	SMARTFCS	Ararma	B_J_Auto	ForecastPro	Auto_ANN
h=1	4	5	6	7	2	1	3
h=2	4	5	3	7	6	1	2
h=3	1	6	7	4	5	3	2
h=4	1	2	3	6	7	5	4
h=5	1	4	2	3	7	5	6
h=6	1	3	2	4	6	7	5
h=7	1	3	2	4	6	7	5
h=8	1	3	2	4	7	6	5
h=9	1	4	3	2	6	7	5
h=10	1	4	3	2	6	7	5
h=11	1	3	4	2	7	6	5
h=12	1	3	4	2	6	7	5
h=13	1	3	4	2	6	7	5
h=14	1	2	4	3	7	6	5
h=15	1	2	3	4	6	7	5
h=16	1	2	3	4	6	7	5
h=17	1	2	3	4	5	7	6
h=18	1	2	3	4	5	6	7
Soma	24	58	61	68	106	102	85
Média	1,3333	3,2222	3,3889	3,7778	5,8889	5,6667	4,7222
Posição	1	2	3	4	7	6	5

Finalmente na tabela 113 é apresentado o resultado obtido para o teste de Post-Hoc, desta vez o modelo proposto conseguiu diferenças significativas com 4 dos 6 modelos comparados, confirmando o bom desempenho para previsão de series mensais.

Tabela 113. Resultados para o teste Post-Hoc.

Teste Pos-Hoc						
Modelo	Modelo	Lim. Inferior	Dif. Médias	Lim. Superior	p-Valor	Conclusão
ESN+GA+SRG	ForcX	-4,0119	-1,8889	0,2341	0,1190	Equivalente
ESN+GA+SRG	SMARTFCS	-4,1786	-2,0556	0,0675	0,0651	Equivalente
ESN+GA+SRG	Ararma	-4,5675	-2,4444	-0,3214	0,0122	ESN+GA+SRG é melhor
ESN+GA+SRG	B_J_Auto	-6,6786	-4,5556	-2,4325	0,0000	ESN+GA+SRG é melhor
ESN+GA+SRG	ForecastPro	-6,4564	-4,3333	-2,2103	0,0000	ESN+GA+SRG é melhor
ESN+GA+SRG	Auto_ANN	-5,5119	-3,3889	-1,2659	0,0001	ESN+GA+SRG é melhor
ForcX	SMARTFCS	-2,2897	-0,1667	1,9564	1,0000	Equivalente
ForcX	Ararma	-2,6786	-0,5556	1,5675	0,9876	Equivalente
ForcX	B_J_Auto	-4,7897	-2,6667	-0,5436	0,0040	ForcX é melhor
ForcX	ForecastPro	-4,5675	-2,4444	-0,3214	0,0122	ForcX é melhor
ForcX	Auto_ANN	-3,6230	-1,5000	0,6230	0,3627	Equivalente
SMARTFCS	Ararma	-2,5119	-0,3889	1,7341	0,9982	Equivalente
SMARTFCS	B_J_Auto	-4,6230	-2,5000	-0,3770	0,0093	SMARTFCS é melhor
SMARTFCS	ForecastPro	-4,4008	-2,2778	-0,1547	0,0262	SMARTFCS é melhor
SMARTFCS	Auto_ANN	-3,4564	-1,3333	0,7897	0,5128	Equivalente
Ararma	B_J_Auto	-4,2341	-2,1111	0,0119	0,0524	Equivalente
Ararma	ForecastPro	-4,0119	-1,8889	0,2341	0,1190	Equivalente
Ararma	Auto_ANN	-3,0675	-0,9444	1,1786	0,8468	Equivalente
B_J_Auto	ForecastPro	-1,9008	0,2222	2,3453	0,9999	Equivalente
B_J_Auto	Auto_ANN	-0,9564	1,1667	3,2897	0,6692	Equivalente
ForecastPro	Auto_ANN	-1,1786	0,9444	3,0675	0,8468	Equivalente

No próximo capítulo são apresentadas as conclusões gerais do trabalho desenvolvido e os trabalhos futuros a serem realizados para incrementar o entendimento e capacidade de modelos híbridos que utilizam as ESNs.

6.

Considerações finais

6.1.

Conclusões

Esta tese apresentou o modelo *ESN-GA-SRG* que considera a otimização de parâmetros globais para redes *ESNs* em tarefas de previsão *multi-step* de series temporais, onde a obtenção dos parâmetros foi realizada mediante a utilização de um Algoritmo Genético e a seleção da melhor topologia de Reservoir foi realizada com o método *SRG*.

Para a confirmação da seleção dos parâmetros a ser otimizados, foi realizado um estudo de sensibilidade para cada um dos 5 parâmetros considerados importantes na literatura recopilada, desta forma foi possível determinar para tarefas de previsão *multi-step* em particular quais os parâmetros considerados com maior ingerência na configuração das *ESNs*.

O objetivo principal foi gerar uma maior compressão do funcionamento das redes *ESN*, e assim poder aproveitar o potencial que oferece este tipo de modelos (Redes Recorrentes) para originar um modelo que tenha uma otimização conjunta de parâmetros globais e ao mesmo tempo uma interpretação do comportamento que contribua na pesquisa continua.

o modelo geral resulta em uma fácil compreensão, isto é, o grau de dificuldade para implementá-lo é baixo, mas, o grau de entendimento dos ecoes gerados pela rede durante o processo de mapeamento dos padrões é pouco, por tanto, esta tese procura gerar uma abordagem que permita diminuir dito grau.

Nesta tese foram apresentados os fundamentos das redes *ESN*, experimentos que permitiram realizar uma análise de sensibilidade em diferentes tipos de séries assim como experimentos empregando os *benchmarks* mais utilizado na previsão *multi-step* de séries temporais, adicionalmente são apresentadas as conclusões do modelo *ESN-GA-SRG*.

Em geral os resultados obtidos demonstraram que o modelo *ESN-GA-SRG* conseguiu melhorar as previsões realizadas por outros modelos e foi possível compreender o comportamento das topologias utilizadas, como por exemplo a sensibilidade na configuração dos parâmetros onde os resultados obtidos confirmaram que a simples mudança de um parâmetro faz com que os resultados mudem para valores de previsão não desejáveis.

Uma outra conclusão obtida tem a ver com as características próprias das séries e como o modelo proposto consegue previsões que podem variar em termos de desempenho. No caso do coeficiente de assimetria, foi observado que séries temporais

com valores altos geram maior dificuldade na previsão. No caso de características como sazonalidade e estacionaridade, o modelo proposto conseguiu previsões consideradas com erros baixos e para o caso da amostragem das séries, os resultados das previsões foram equivalentes para as séries com amostragem trimestral nas duas bases de dados utilizadas.

6.2.

Próximas etapas

Como próximos passos sugere-se realizar um tratamento das séries, com relação a propriedades como sazonalidade e tendência, antes da aplicação do modelo. Recomenda-se, também, analisar o comportamento quando utilizados neurônios do tipo *Integrate-and-Fire*. Esta nova abordagem permitirá incluir novos parâmetros e também explorar a dinâmica deste tipo de redes com padrões que contenham codificação temporal.

Quando criados os estados pelo *Reservoir* o estado inicial pode ser zero ou um valor randômico; em ambos os casos este valor fara parte do ponto de partida do próximo estado, em alguns apartes da teoria das *ESNs* é proposta a alternativa de desconsiderar os primeiros estados, isto é, esperar que o efeito produzido por este valor inicial desvaneça-se até o ponto de não ser considerado (Lazar, 2007).

Para próximos trabalhos considerar este tipo de parâmetro complementaria o estudo e a compreensão da propagação da informação quando gerado os estado no *Reservoir*.

Como apresentado no capítulo 2, existe outro modelo que faz uso de uma metodologia similar, as *Liquid State Machine* pelas suas siglas em inglês *LSM*, como trabalho de comprovação uma linha de pesquisa futura poderia ser a generalização do principio utilizado no modelo do presente trabalho para outros modelos de *Reservoir Computing*.

Referências bibliográficas

- Bian, Z.; Miller, D. J.; Wang, Y., "Nonlinear System Modeling with Random Matrices: Echo State Networks Revisited", **IEEE Trans. Neural Networks and Learning Systems**, p.175-182, 2012.
- Bian, X.; Mou, C., "Identification of non-linear dynamic model of UUV based on ESN neural network", **30th Chinese Control Conference (CCC)**, p.1432-1437, 2011.
- BOX, G. E. P.; JENKINS, G. M., "Time Series Analysis: Forecasting and Control". **Oakland: Holden-Day**, 1976.
- BREVE, F., "Echo State Network Sistemas Inteligentes". Universidade de São Paulo, São Carlos, Instituto de Ciências Matemáticas e de Computação, 2005.
- Cernansky, M.; Frank, S., "Generalization and Systematicity in Echo State Networks", **30th Annual Conference of the Cognitive Science Society**, p.733-738, 2008.
- Csercsik, D.; Szederkényi, G.; Hangos, K. M; Farkas, I, "Model synthesis identification a Hodgkin-Huxley-type neuron model," Control Conference (ECC), 2009 European, Budapest, p.3058-3063, 2009.
- Chen, H.; Hardle, W.; Jeong, S., "Nonparametric Risk Management with Generalized Hyperbolic Distributions", **Journal of the American Statistical Association**, p.910–923, 2008.
- Cherif, A.; Cardot, H.; Boné, R., "SOM time series clustering and prediction with recurrent neural networks", **Journal Neurocomputing**, Vol. 74, p.1936-1944, 2011.
- Conover, W. J., "Practical Nonparametric Statistics", 2 Ed, New York, John Wiley, 1980.
- de Aquino, R.; Souza, R.; Neto, O., Lira, M.; Carvalho, M.; Ferreira, A., "Echo state networks, artificial neural networks and fuzzy systems models for improve short-term wind speed forecasting," 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, pp. 1-8, 2015.
- Dominey, P. F.; Arbib, M. A.; Joseph, J. P., "A model of cortico-striatal plasticity for learning oculomotor associations and sequences", **Journal Cognitive Neuroscience**. p.311–336, 1995.
- Elman, J., "Finding Structure in Time", **Cognitive Science 14**, p.179-211, 1990.
- Feldbusch, F.; Kaiser, F., "Simulation of Spiking Neural Nets with INSPiRE ME," 2005 IEEE International Conference on Systems, Man and Cybernetics, Waikoloa, HI, p. 999-1004, 2005.
- Ferreira, A. T., "Comparing evolutionary methods for Reservoir computing pre-training", **International Joint Conference on Neural Networks 2011**, p.283-290, 2011.
- Gibbons, T., "Unifying Quality Metrics for Reservoir Networks", **International Joint Conference on Neural Networks**, p.1-7, 2010.
- Golestaneh, F.; Pinson, P.; Gooi, H. B., "Very Short-Term Nonparametric Probabilistic Forecasting of Renewable Energy Generation— With Application to Solar Energy," IEEE Transactions on Power Systems, vol. 31, no. 5, p.3850-3863, 2016.
- GOLDBERG, D., "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley Longman Publishing Co, 372p, 1999.
- Goldenholz, D., "Liquid computig: A real effect. Technical report", Boston University Department of Biomedical Engineering, 2002.

González, C.; Mira-McWilliams, J.; Juárez, I., "Important variable assessment and electricity price forecasting based on regression tree models: classification and regression trees, Bagging and Random Forests," *IET Generation, Transmission & Distribution*, vol. 9, no. 11, p1120-1128, 2015.

Goodman, E. and Ventura, D., "Spatiotemporal pattern recognition via liquid state machines," *Proceedings of the International Joint Conference on Neural Networks*, p3848–3853, 2006.

Grando, N., "Máquina de estado líquido para previsão de séries temporais contínuas: aplicação na demanda de energia elétrica", Tese de doutorado - UTFPR, 145p, 2010.

Graves, A.; Schmidhuber, J., "Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks", *Advances in Neural Information Processing Systems* **22**, p.545-552, 2009.

Graves, A.; Mohamed, A.; Hinton, G., "Speech Recognition with Deep Recurrent Neural Networks", *International Conference on Acoustics, Speech and Signal Processing*, 2013.

Hassanzadeh, M.; Evrenosoğlu, C. Y.; Mili, L., "A Short-Term Nodal Voltage Phasor Forecasting Method Using Temporal and Spatial Correlation," *IEEE Transactions on Power Systems*, vol. 31, no. 5, p.3881-3890, 2016.

Hopfield, J., "Neural networks and physical systems with emergent collective computationalabilities", *National academy of Science of United states*, Vol. 79, p.2554-2558, 1982.

Ilies, I.; Jaeger, H., "Stepping forward through echoes of the past: forecasting with Echo State Networks", *Artificial Neural Networks & Computational Intelligence Forecasting Competition*, 2007.

Ishii, K.; Van der Zant, T.; Becanovic, V.; Ploger, P., "Optimization of parameters of echo state network and its application to underwater robot," *SICE 2004 Annual Conference*, Sapporo, 2004, pp. 2800-2805 vol. 3.

Jaeger, H., "The "Echo State" Approach to Analysing and Training Recurrent Neural Networks – with an Erratum note", Technical Report Fraunhofer Institute for Autonomous Intelligent Systems, 2001.

Jaeger, H., "Adaptive nonlinear system identification with echo state networks", Paper *Advances in Neural Information Processing Systems*, 2002.

Jaeger, H.; Haas, H., "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication", *American Association for the Advancement of Science*, 2004.

Jaeger, H., "Reservoir riddles: Suggestions for echo state network research", *IEEE International Joint Neural Networks*, 2005.

Jaeger, H.; Lukosevicius, M.; Popovici, D.; & Siewert, U., "Optimization and applications of echo state networks with leaky- integrator neurons", *Neural Networks Journal*, p.335-352, 2007.

Jaeger, H., "Erratum note for the techreport: the "echo state" approach to analysing and training recurrent neural networks", Technical Report, 2010.

Jianhui, X. H., "Prediction of chaotic time series based on the recurrent predictor neural network", *IEEE Transactions on Signal Processing*, 2005.

Jordan, M. I., "Serial order: A parallel distributed approach", *ICS Report 8604*, 1986.

Kosko, B., "Bidirectional Associative memories", *IEEE Transactions on Systmes, Man, And Cybernetics*, p.49-60, 1988.

Lacerda, E. G. M., "Algoritmos Genéticos: Aspectos Práticos", 36p, 2008.

Lazar, A.; Pippa, G.; Triesch, J., "Fading memory and time series prediction in recurrent networks with different forms of plasticity", **Neural Networks Journal**, Vol. 20, p.312–322, 2007.

Lazo, J. G., "Determinação do Valor de Opções Reais por Simulação Monte Carlo com Aproximação por Números Fuzzy e Algoritmos Genéticos", Tese de Doutorado - PUC-Rio, 2004.

Legenstein, R.; Maass, W., "Edge of chaos and prediction of computational performance for Neural circuit models", **Neural Networks Journal**, Vol. 20, p.323-334, 2007.

Lin, Q.; Biao, Yang.; Jiao X., "The NGP-stability of Runge–Kutta methods for systems of neutral delay differential equations Numer", p.451–459, 1999.

Luque, B.; Sole, R. V., "Lyapunov exponents in random Boolean networks", p33–45, 2000.

Lukoševičius, M.; Jaeger, H., "Overview of *Reservoir* recipes", Technical Report University Bremen, 2007.

Lukoševičius, M.; Jaeger, H. "*Reservoir* computing approaches to recurrent neural network training", **Computer Science Review**, p.127-149, 2009.

Mackey, M. C.; Leon, G., "Oscillation and Chaos in Physiological Control Systems", **Science, New Series**, Vol. 197, p.287-289, 1977.

Michelewicz, Z., "Genetic Algorithms + Data Structures = Evolution Programs", **Springer series Artificial Intelligence**, 388p, 1996.

Mitul, A., "Classification of real time moving object using echo state network", **International Conference on Informatics, Electronics & Vision**, 2013.

Natschlaeger, T.; Maass, W., "Spiking neurons and the induction of finite state machines. Theoretical Computer Science", **Special Issue on Natural Computing**, p.251-265, 2002.

Natschlager, T.; Bertschinger, N.; Legenstein, R. (2005). "At the edge of chaos: Real- " time computations and self-organized criticality in recurrent neural networks", **Advances in neural information processing systems**, p.145–152, 2005.

Norton, D., "Improving Liquid State Machines Through Iterative Refinement of the Reservoir," Thesis, Computer Science, Brigham Young University, 2008.

NN3. NN3 Forecast Competition Home. Disponível em: < <http://www.neural-forecasting-competition.com/NN3/index.htm>>. Acesso em 10/06/2015.

Paine, R. E.; Tani, J., "Motor primitive and sequence self-organization in a hierarchical recurrent neural network", **Neural Networks Journal**, p.1291-1309, 2004.

Park, H.; Baldick, R.; Morton D. P., "A Stochastic Transmission Planning Model With Dependent Load and Wind Forecasts," IEEE Transactions on Power Systems, vol. 30, no. 6, p3003-3011, 2015.

Prokhorov, D., "Echo state networks: appeal and challenges," Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005., 2005, pp. 1463-1466 vol. 3.

Qiao, J.; Li, F.; Han, H.; Li, W., "Growing Echo-State Network With Multiple Subreservoirs," in IEEE Transactions on Neural Networks and Learning Systems , 2016, vol.PP, no.99, pp.1-14

Rosenstein, M.; Collins, J., "A practical method for calculating largest Lyapunov exponents from small data sets", 2012.

Salmen, M.; Ploger, P. G., "Echo State Networks used for Motor Control", *IEEE International Conference on Robotics and Automation*, p.1953-1958, 2005.

Santos, J.; Campo, A., "Biped Locomotion control with evolved adaptive center-crossing continuous time recurrent neural networks", *Neurocomputing*, 2012.

Schrauwen, B.; Wardermann, M.; Verstraeten, D.; Steil, J. J.; Stroobandt, D., "Improving Reservoirs using intrinsic plasticity", *Neurocomputing*, p.1159–1171, 2008.

Schrauwen, B.; et al., "On Computational Power and the Order-Chaos Phase Transition in Reservoir Computing," in *Advances in Neural Information Processing Systems* Vol. 21, 2008.

Shao, H.; Deng, X.; Cui, F., "Short-term wind speed forecasting using the wavelet decomposition and AdaBoost technique in wind farm of East China," *IET Generation, Transmission & Distribution*, vol. 10, no. 11, p2585-2592, 2016.

Steil, J., "Backpropagation-decorrelation: recurrent learning with $O(N)$ complexity", *IEEE International Joint Conference on Neural Networks*, Vol. 2, p843–848, 2004.

Steil, J. J., "Online stability of backpropagation-decorrelation recurrent learning", *Neurocomputing*, p.642–650, 2006.

Steil, J. J., "Online Reservoir adaptation by intrinsic plasticity for backpropagation", *Neural Networks Journal*, p353-364, 2007.

Tutschku, K., "Recurrent Multilayer Perceptrons for identification and control: the road to applications", Technical Report University of Wurzburg, 1995.

Valle, R., "Ponderação Neural de Experts", Tese de Doutorado – PUC-Rio, 181p, 2012.

Velasco, C., "ESNPredictor: Ferramenta de Previsão de Séries Temporais Baseada em Echo State Networks Otimizadas por Algoritmos Genéticos e Particle Swarm Optimization", Dissertação de Mestrado – PUC-Rio, 93p, 2014.

Verstraeten, D.; Schrauwen, B.; Stroobandt, D.; Campenhout, J. V., "Isolated word recognition with the liquid state machine: A case study", *Information Processing Letters*, p.521–528, 2005.

Verstraeten, D.; Schrauwen, B.; D'Haene, M.; Stroobandt, D., "An experimental unification of Reservoir computing methods", *Neural Networks Journal*, p.391-403, 2007.

Vogels, T. P.; Rajan, K.; Abbott, L., "Neural networks dynamics", *Annual Review of Neuroscience*, p.357–376, 2005.

Wang, Z.; Guo, L.; Adjouadi, M., "A biological plausible Generalized Leaky Integrate-and-Fire neuron model," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, p. 6810-6813, 2014.

Xu, D.; Lan, J.; Principe, J., "Direct adaptive control: an echo state network and genetic algorithm approach Neural Networks", *IEEE International Joint Neural Networks*, Vol. 3, p.1483-2486, 2005.

Yao, X.; Lun, S., "Parameters selection method of the leaky integrator echo state network for time series," *Mechatronics and Control (ICMC)*, 2014 International Conference on, Jinzhou, 2014, pp. 1288-1293.

Yuenyong, S., "Fast and effective tuning of Echo State Network reservoir parameters using evolutionary algorithms and template matrices," 2015 International Computer Science and Engineering Conference (ICSEC), Chiang Mai, 2015, pp. 1-6.

Zhao, J.; Deng, J.; Ye, W.; Lu, X., "Combined forecast method of HMM and LS-SVM about electronic equipment state based on MAGA," Journal of Systems Engineering and Electronics, vol. 27, no. 3, p730-738, 2016.

Zimmermann, H. G.; Grothmann, R.; Tietz, Ch., "Forecasting Market Prices with Causal-Retro-Causal Neural Networks", International Conference on Operations Research, 2012.

A

Versão reduzida NN3

A seguir são apresentadas nas figuras 147 a 157 as 11 séries originais que compõem a versão reduzida da competição NN3, com os conjuntos de treinamento, validação e teste.

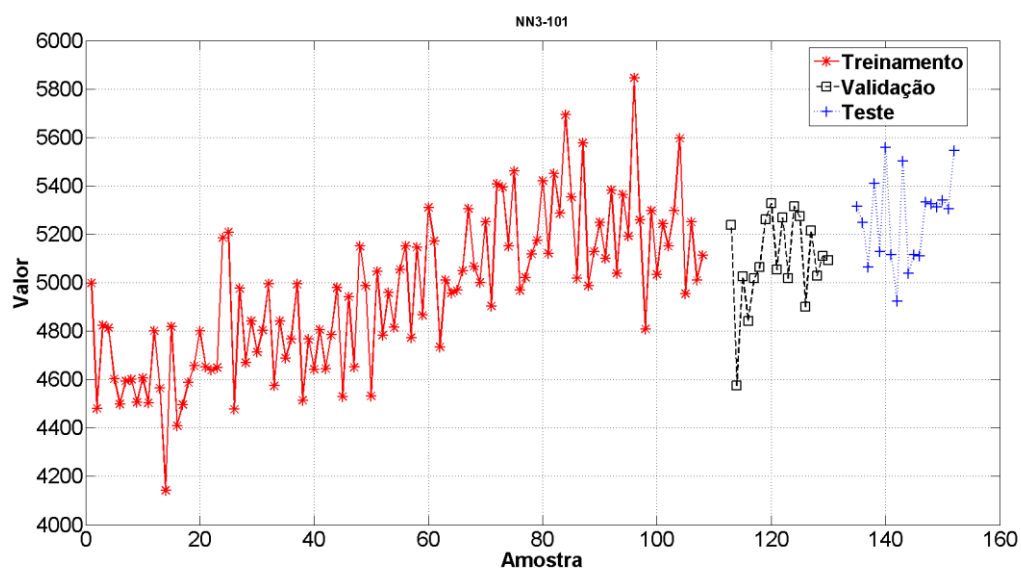


Figura 147. Conjuntos de Treinamento, Validação e Teste da Série NN3-101.

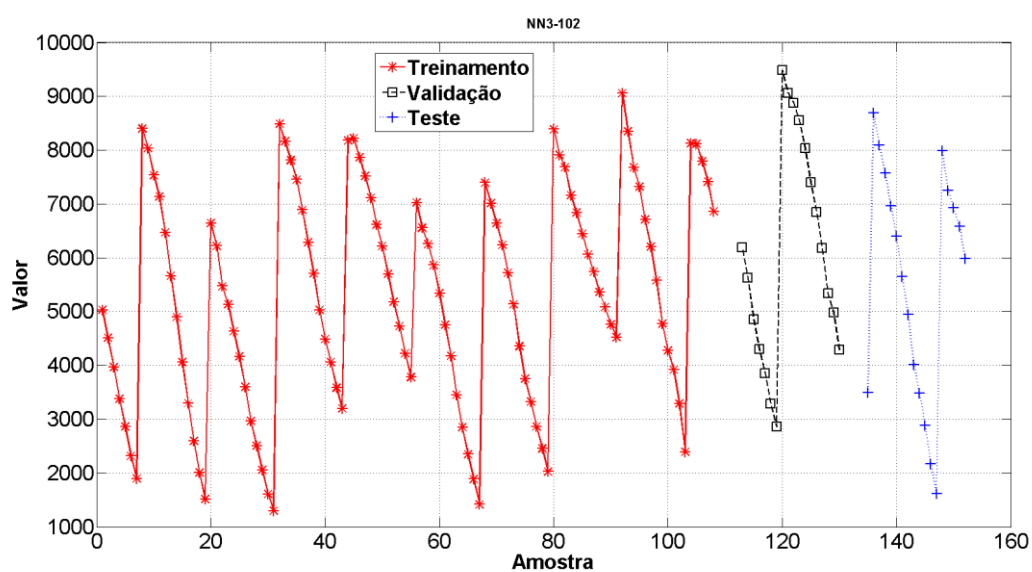


Figura 148. Conjuntos de Treinamento, Validação e Teste da Série NN3-102.

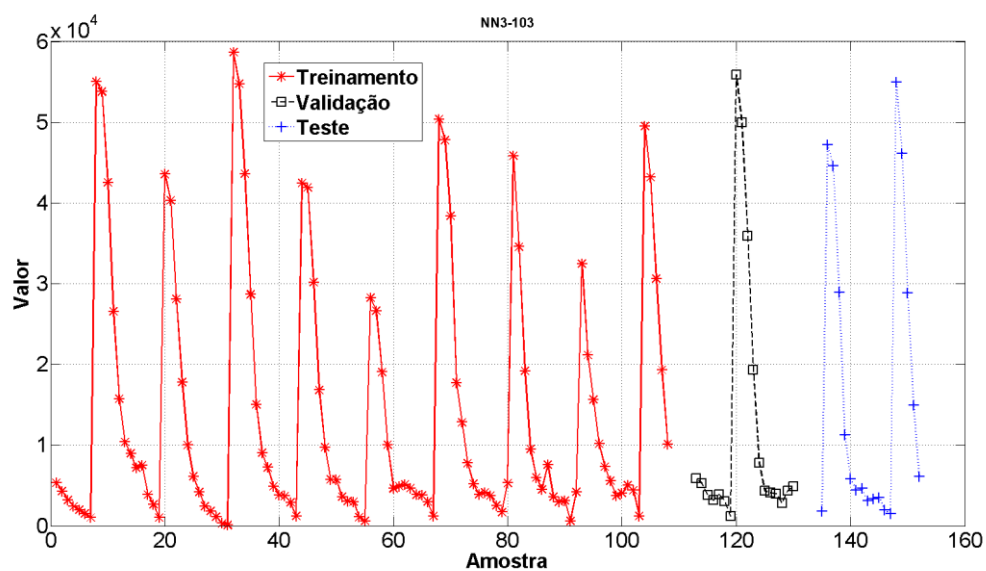


Figura 149. Conjuntos de Treinamento, Validação e Teste da Série NN3-103.

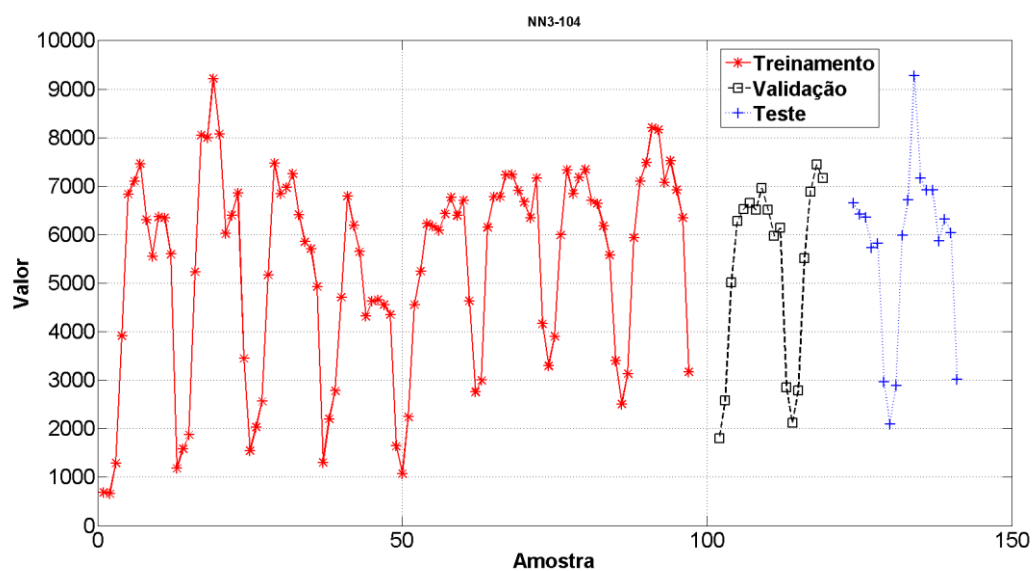


Figura 150. Conjuntos de Treinamento, Validação e Teste da Série NN3-104.

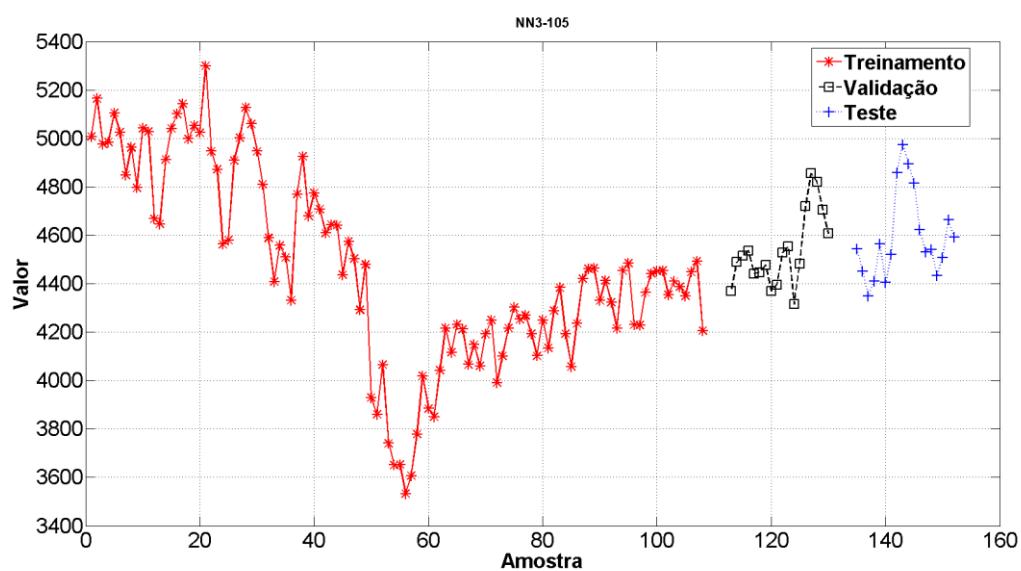


Figura 151. Conjuntos de Treinamento, Validação e Teste da Série NN3-105.

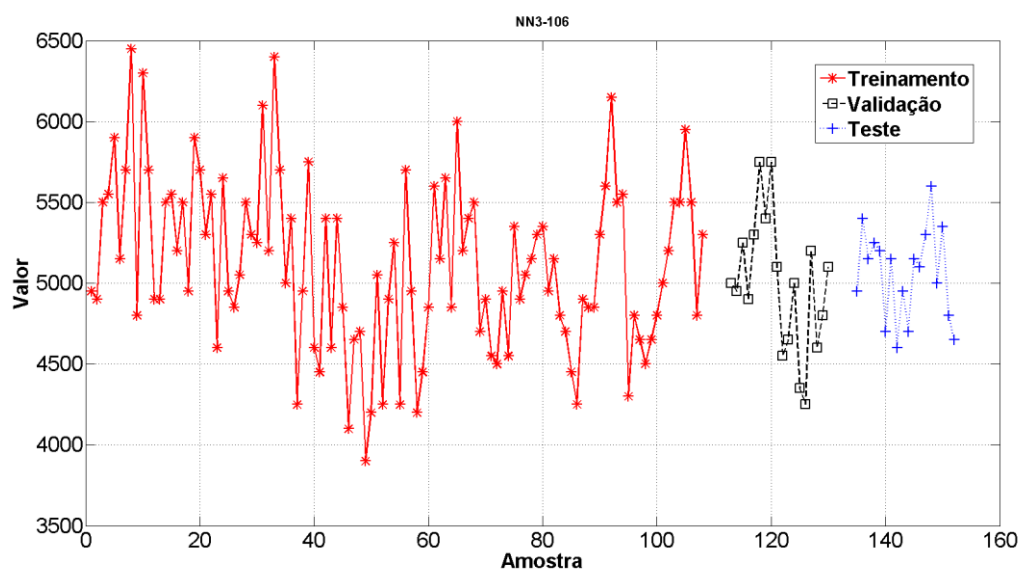


Figura 152. Conjuntos de Treinamento, Validação e Teste da Série NN3-106.

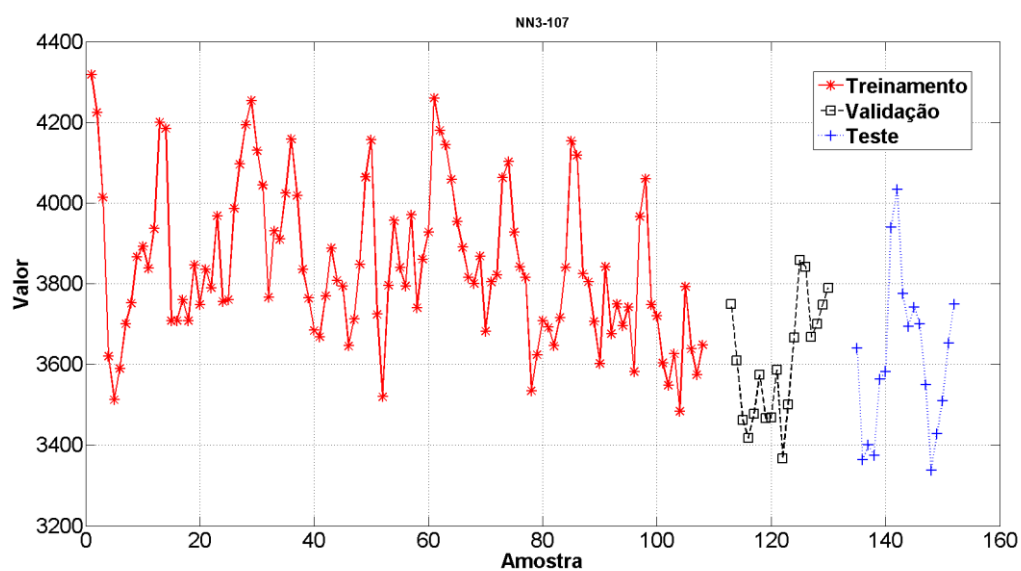


Figura 153. Conjuntos de Treinamento, Validação e Teste da Série NN3-107.

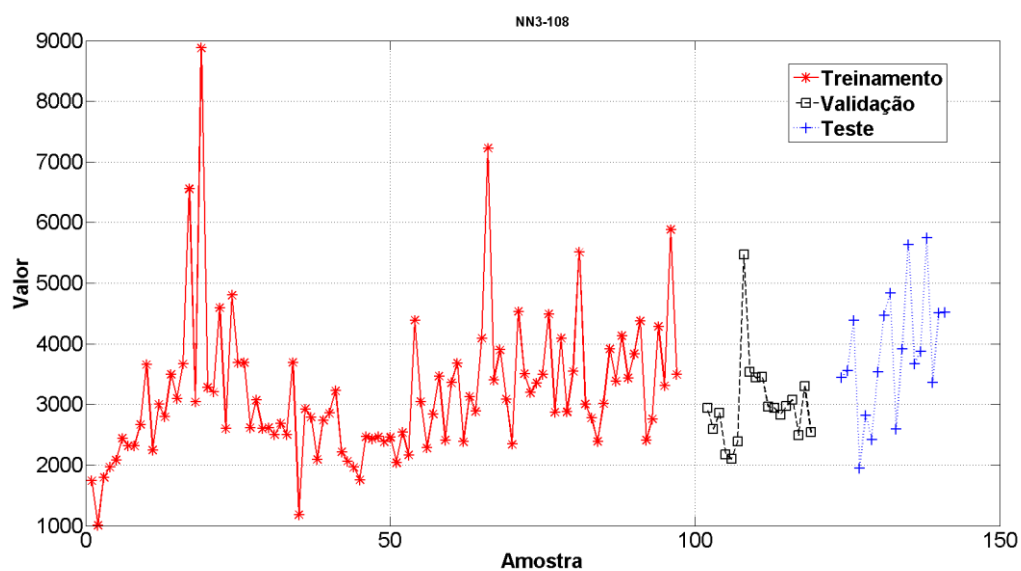


Figura 154. Conjuntos de Treinamento, Validação e Teste da Série NN3-108.

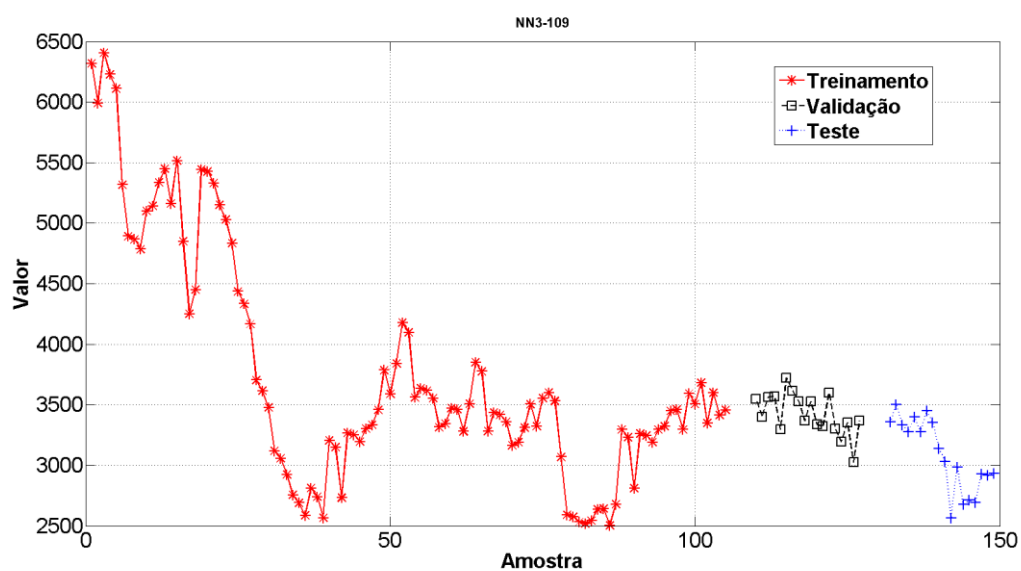


Figura 155. Conjuntos de Treinamento, Validação e Teste da Série NN3-109.

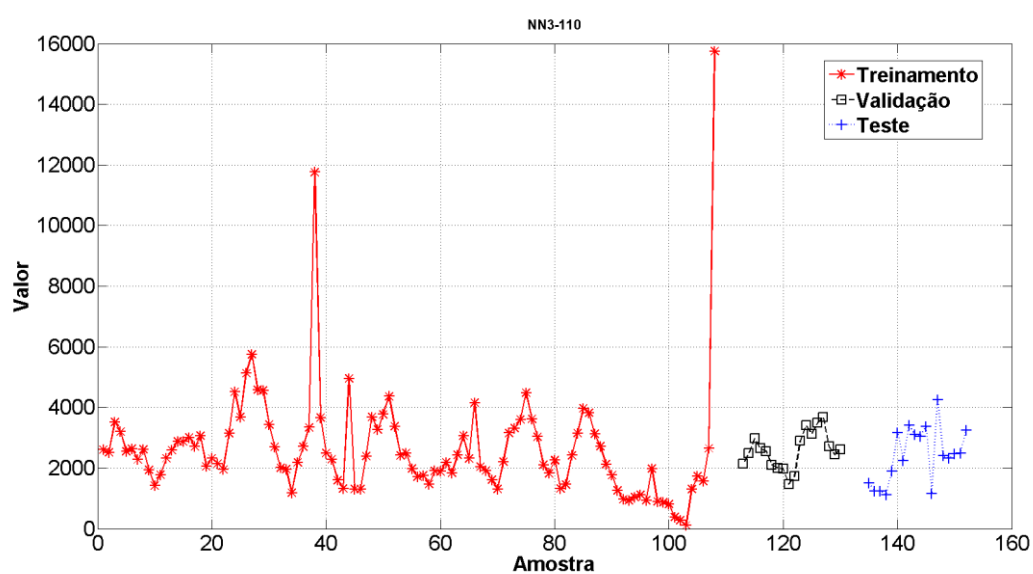


Figura 156. Conjuntos de Treinamento, Validação e Teste da Série NN3-110.

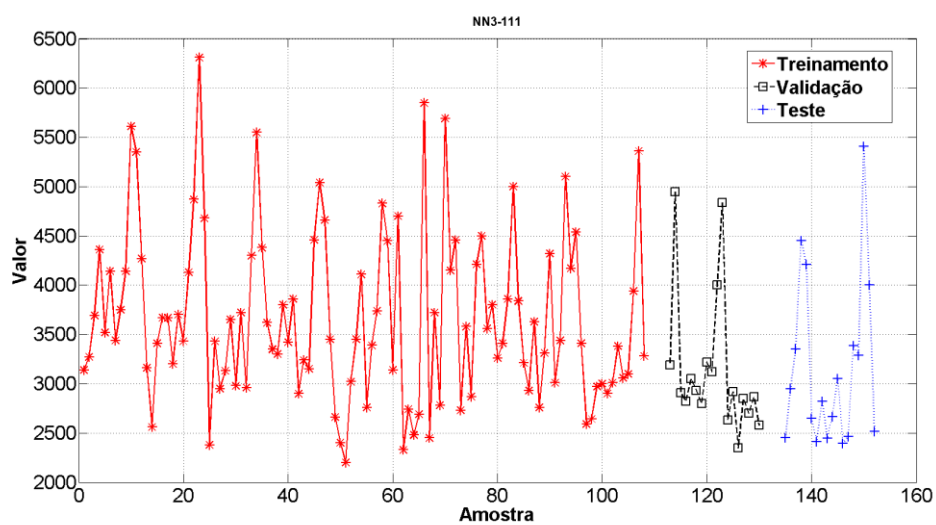


Figura 157. Conjuntos de Treinamento, Validação e Teste da Série NN3-111.

B

Séries selecionadas Base de Dados M3

A seguir são apresentadas nas figuras 158 a 194 as 36 séries que compõem a os 3 grupos de selecionados por amostragem.

- **Série M3_0033**

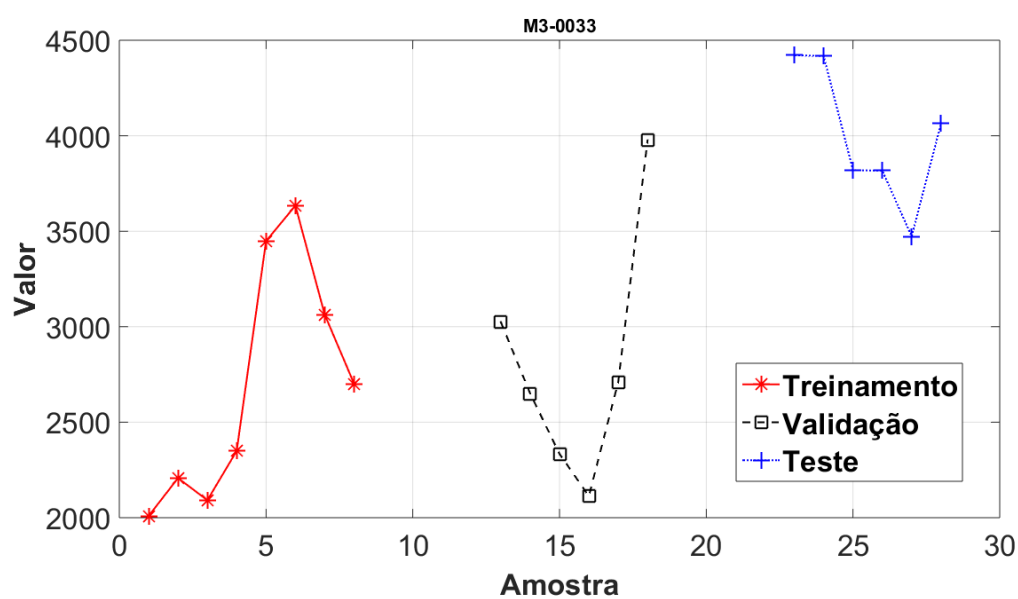


Figura 158. Conjuntos de Treinamento, Validação e Teste da Série M3-0033.

- **Série M3_0050**

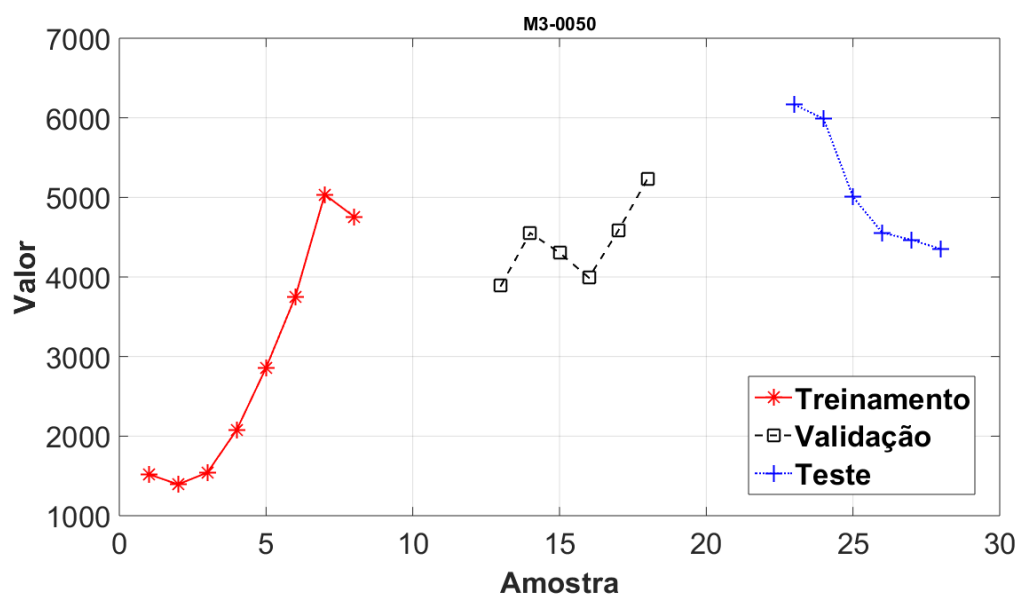


Figura 159. Conjuntos de Treinamento, Validação e Teste da Série M3-0050.

- **Série M3_0070**

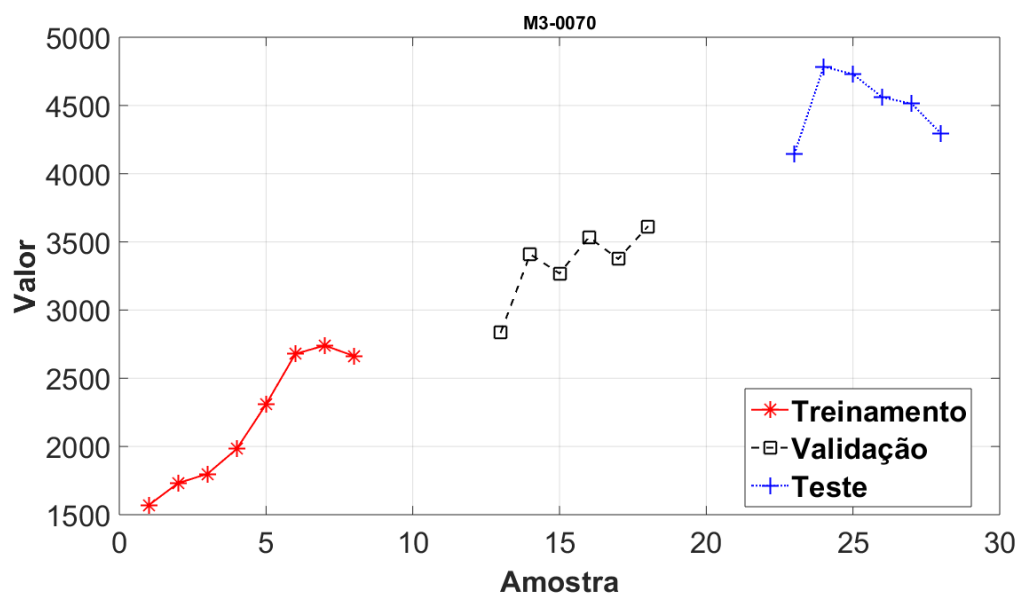


Figura 160. Conjuntos de Treinamento, Validação e Teste da Série M3-0070.

- **Série M3_0105**

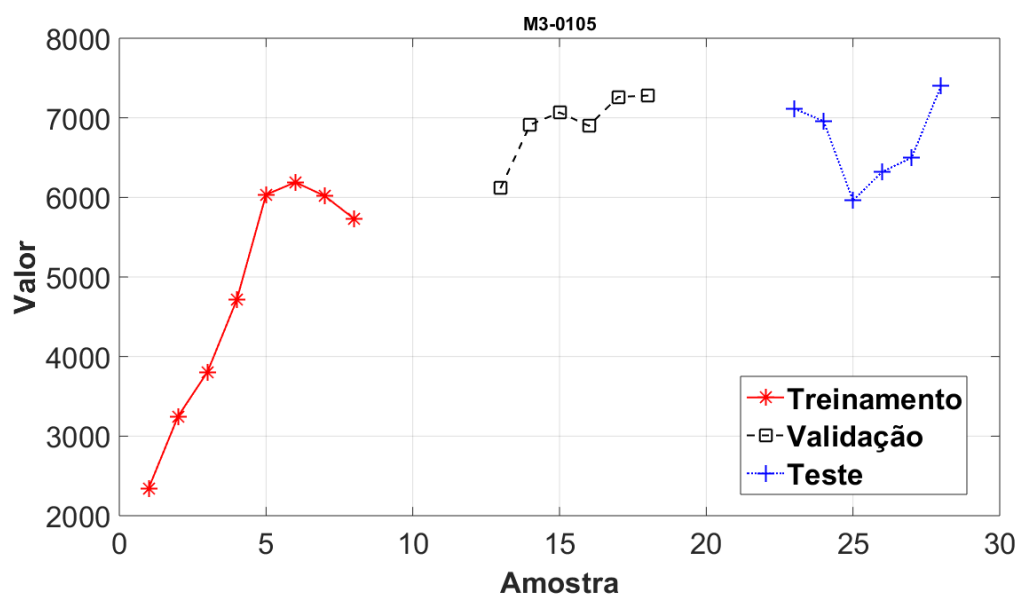


Figura 161. Conjuntos de Treinamento, Validação e Teste da Série M3-0105.

- **Série M3_0116**

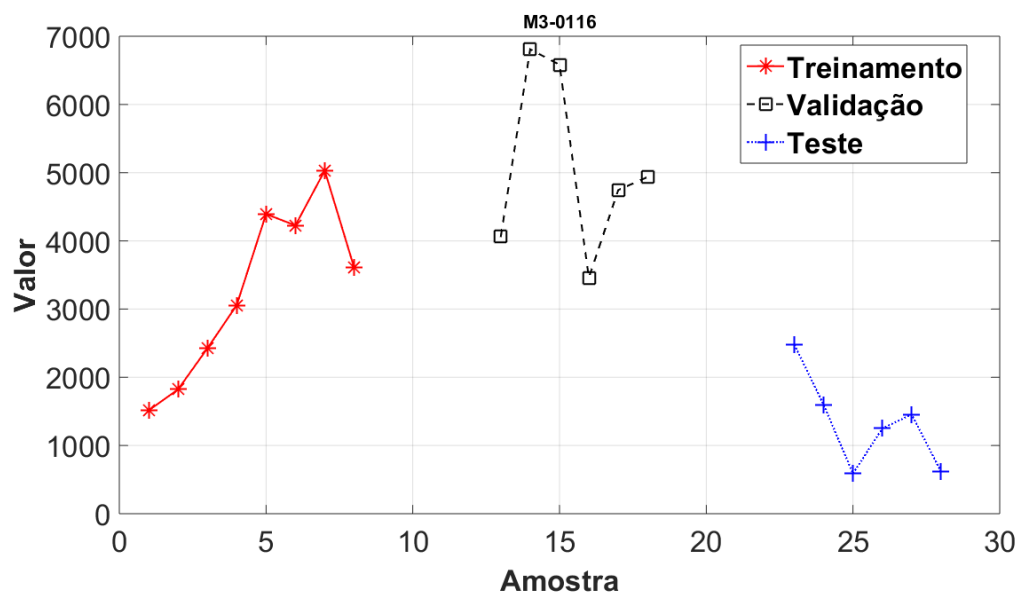


Figura 162. Conjuntos de Treinamento, Validação e Teste da Série M3-0116.

- **Série M3_0119**

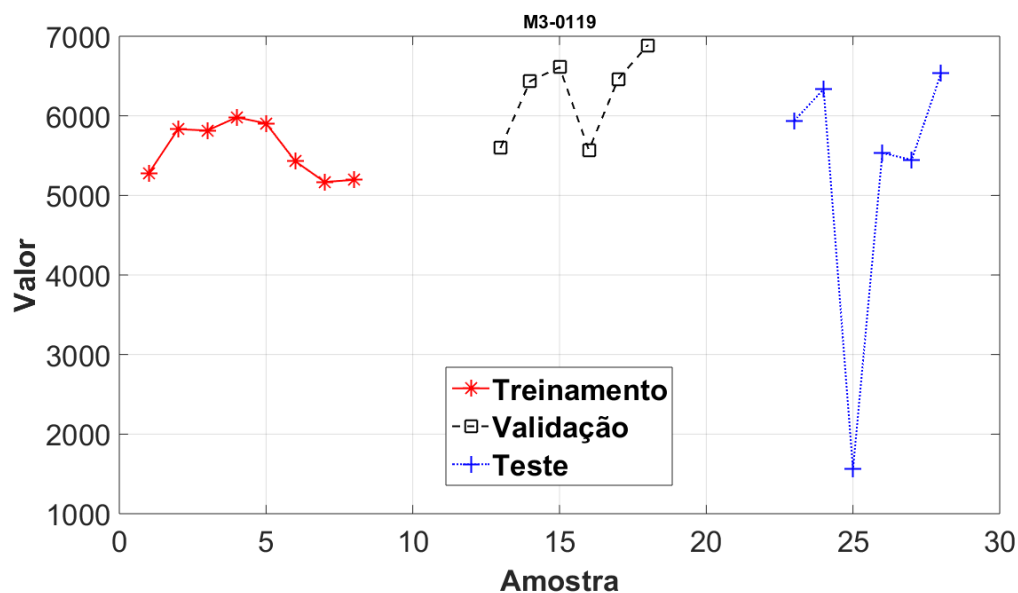


Figura 163. Conjuntos de Treinamento, Validação e Teste da Série M3-0119.

- **Série M3_0162**

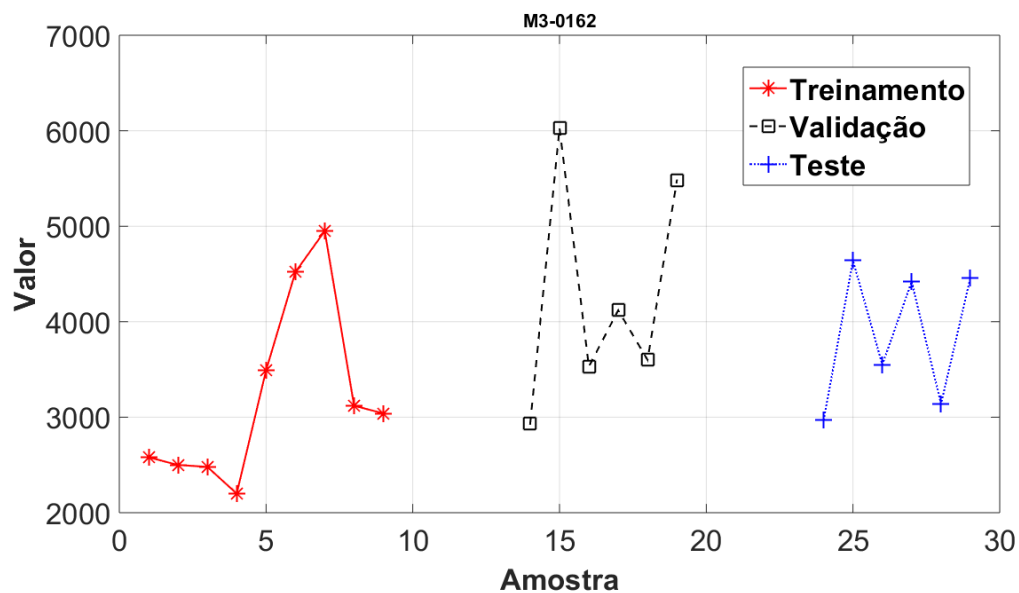


Figura 164. Conjuntos de Treinamento, Validação e Teste da Série M3-0162.

- **Série M3_0171**

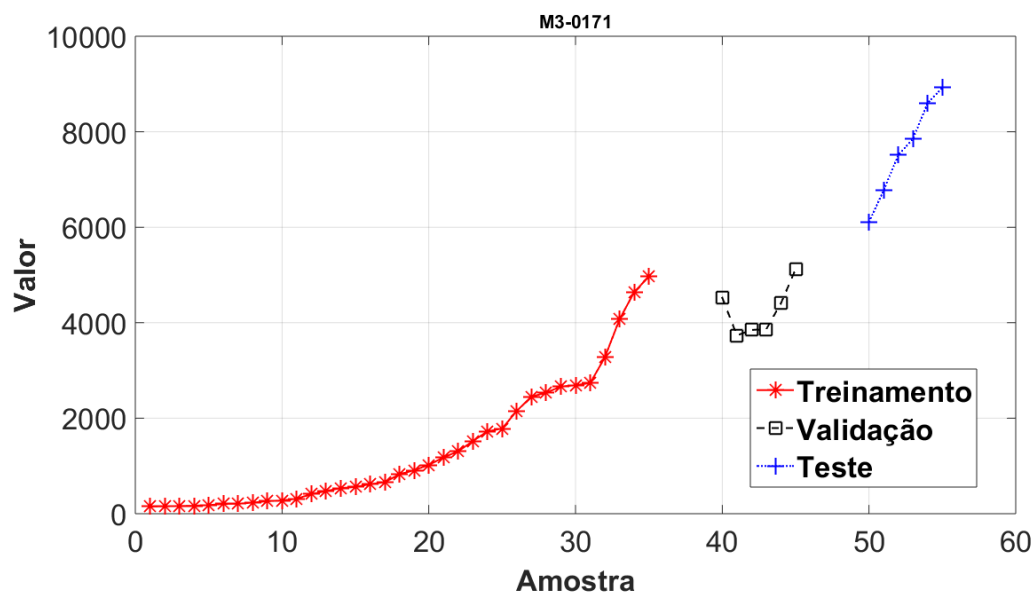


Figura 165. Conjuntos de Treinamento, Validação e Teste da Série M3-0171.

- **Série M3_0200**

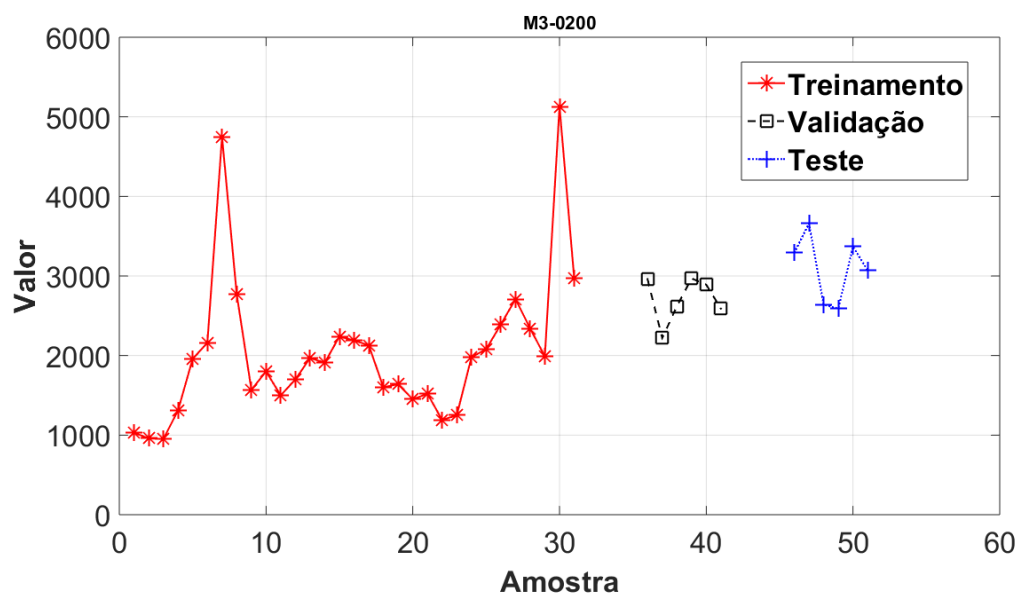


Figura 166. Conjuntos de Treinamento, Validação e Teste da Série M3-0200.

- **Série M3_0351**

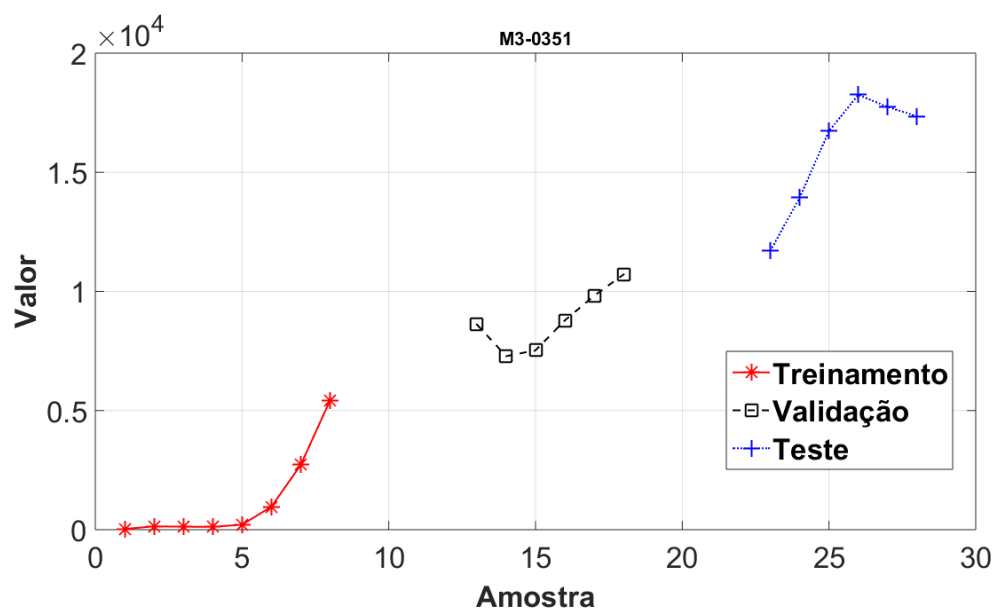


Figura 167. Conjuntos de Treinamento, Validação e Teste da Série M3-0351.

- **Série M3_0564**

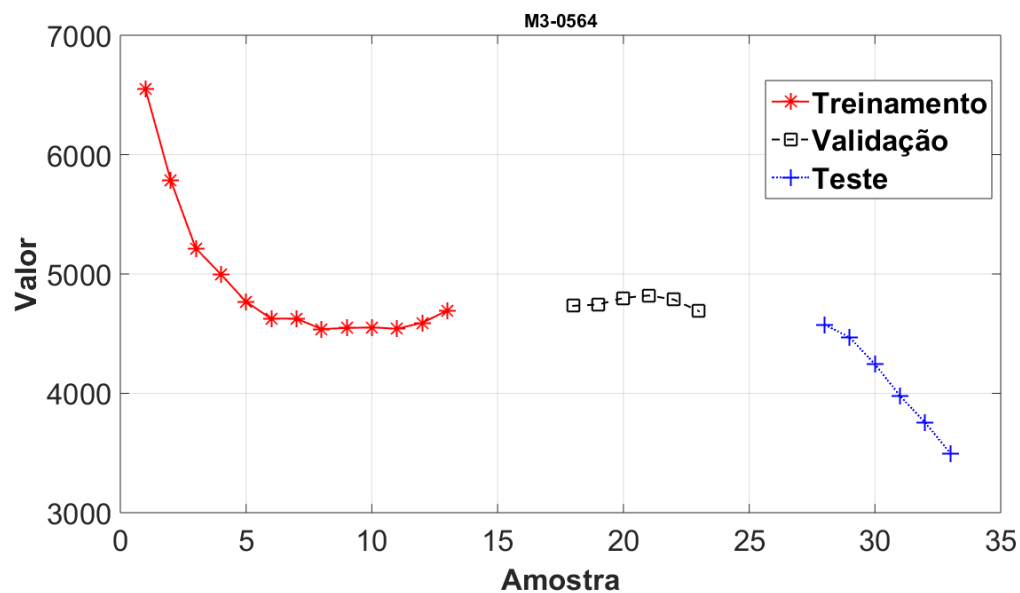


Figura 168. Conjuntos de Treinamento, Validação e Teste da Série M3-0564.

- **Série M3_0599**

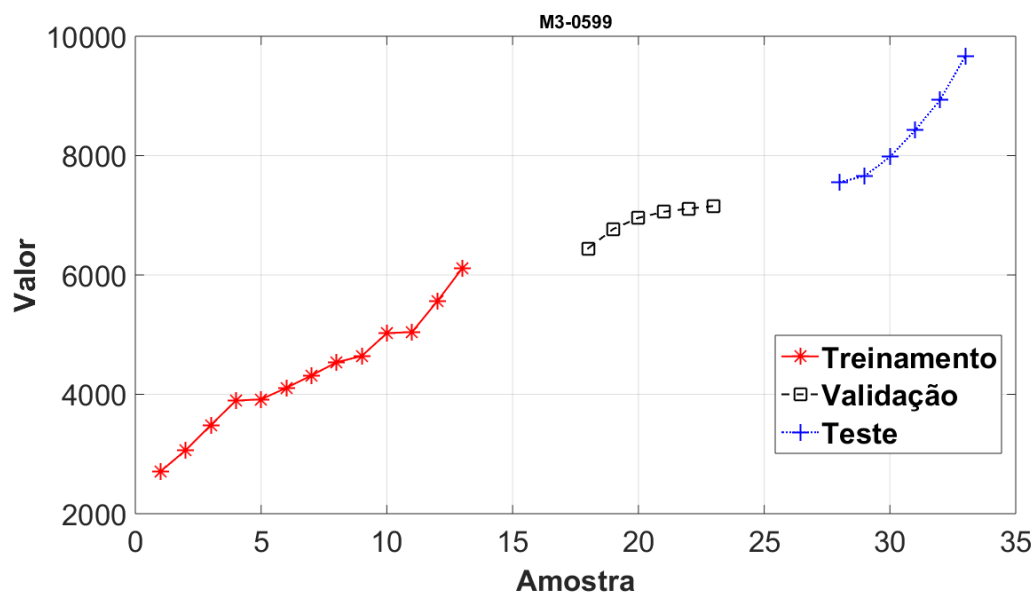


Figura 169. Conjuntos de Treinamento, Validação e Teste da Série M3-0599.

- **Série M3_0642**

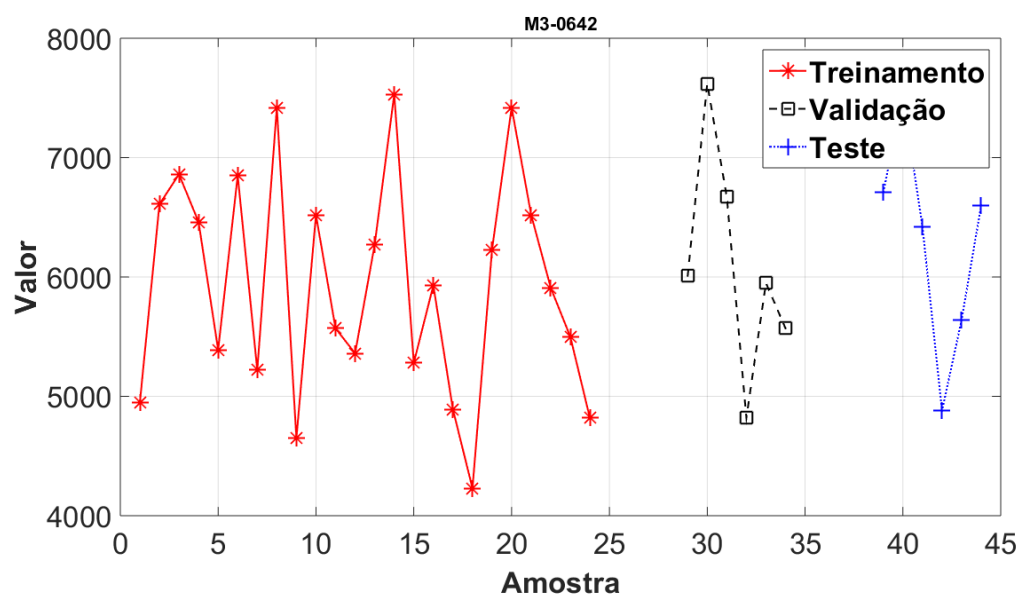


Figura 170. Conjuntos de Treinamento, Validação e Teste da Série M3-0642.

- **Série M3_0671**

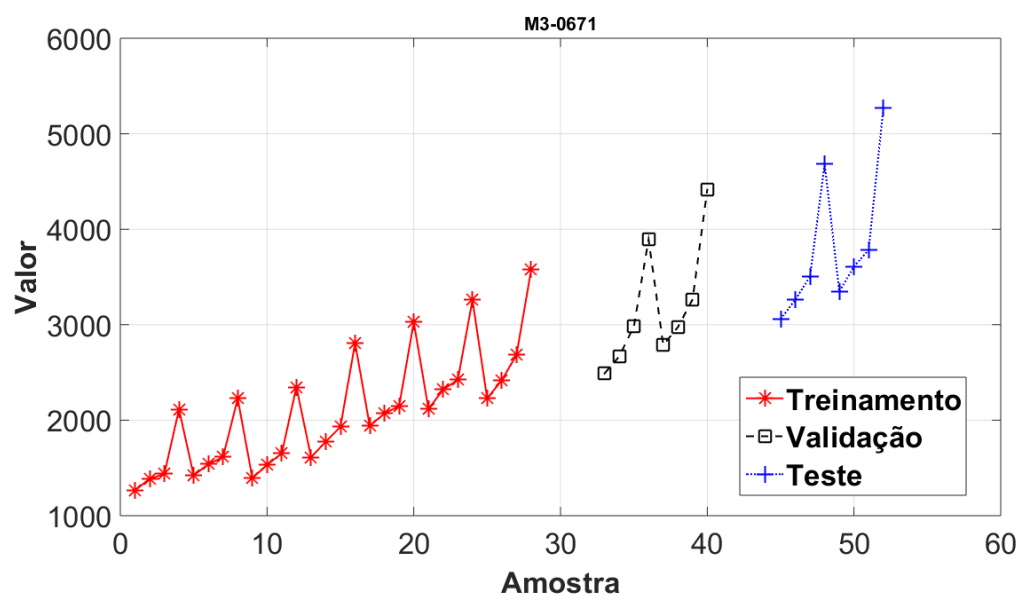


Figura 171. Conjuntos de Treinamento, Validação e Teste da Série M3-0671.

- **Série M3_0677**

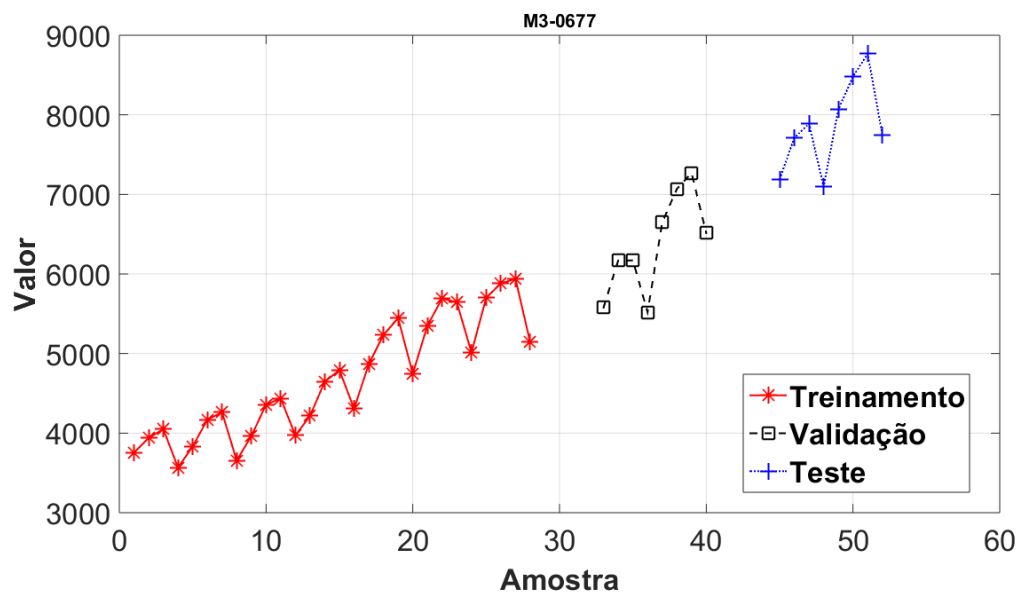


Figura 172. Conjuntos de Treinamento, Validação e Teste da Série M3-0671.

- **Série M3_0692**

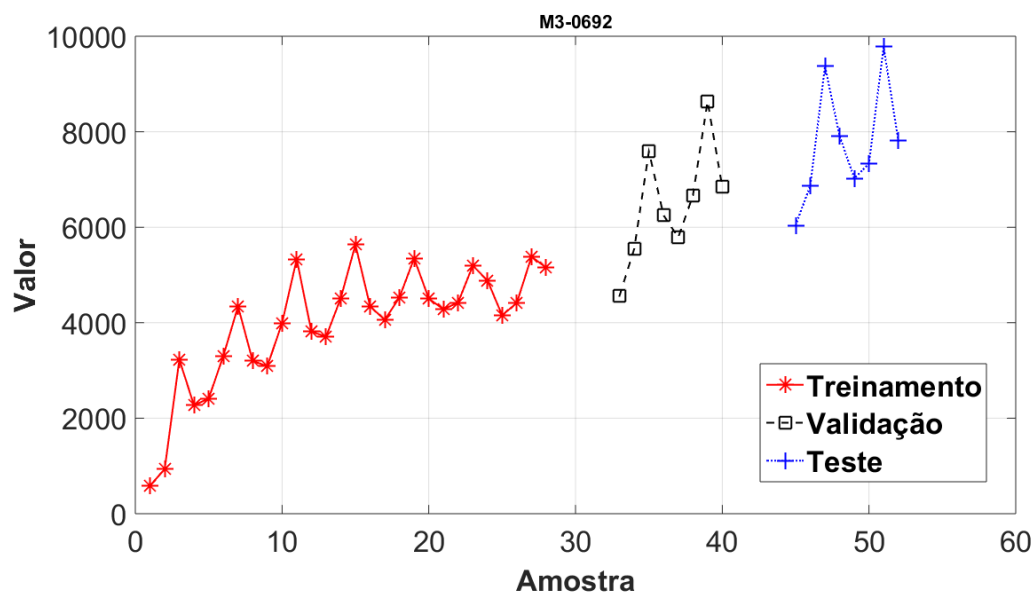


Figura 173. Conjuntos de Treinamento, Validação e Teste da Série M3-0692.

- **Série M3_0713**

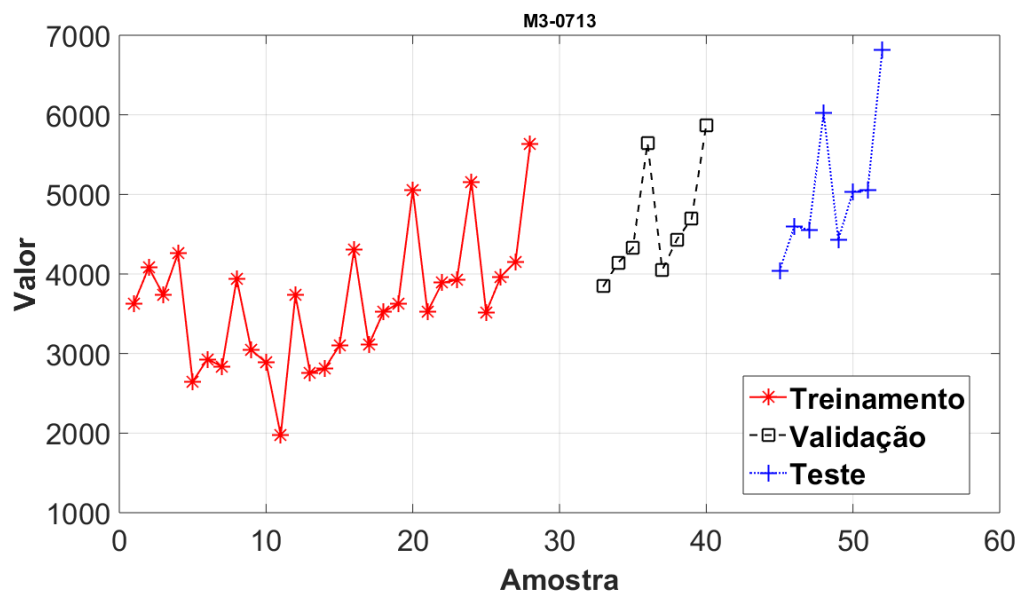


Figura 174. Conjuntos de Treinamento, Validação e Teste da Série M3-0713.

- **Série M3_0730**

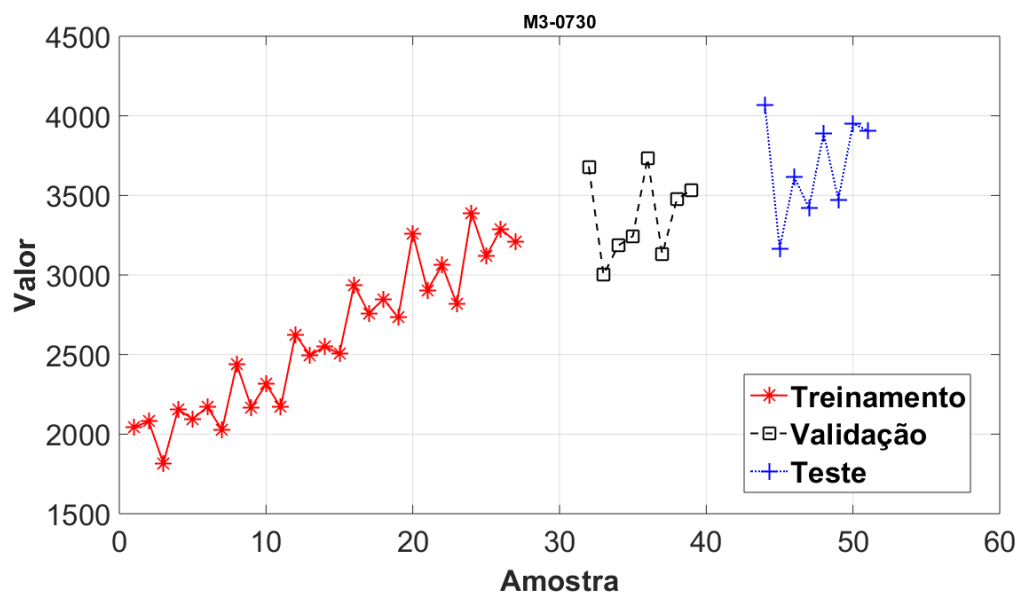


Figura 175. Conjuntos de Treinamento, Validação e Teste da Série M3-0730.

- **Série M3_0758**

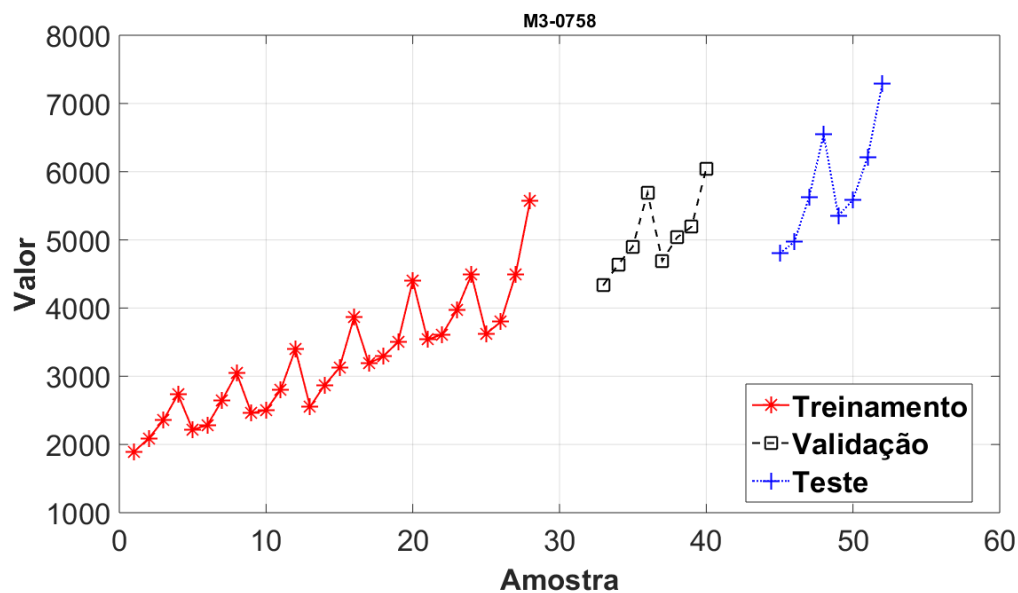


Figura 176. Conjuntos de Treinamento, Validação e Teste da Série M3-0758.

- **Série M3_0758**

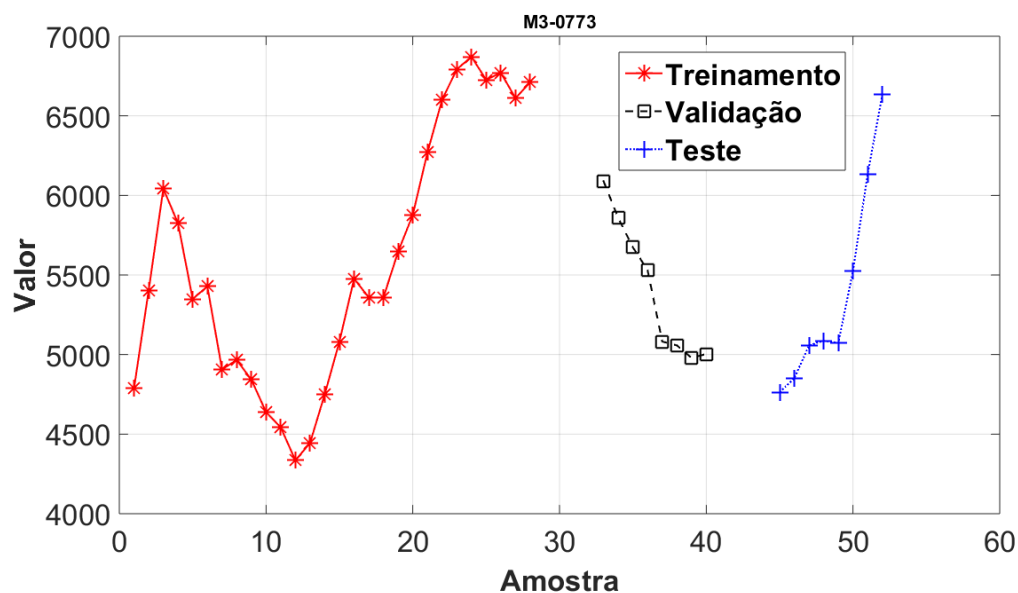


Figura 177. Conjuntos de Treinamento, Validação e Teste da Série M3-0773.

- **Série M3_0829**

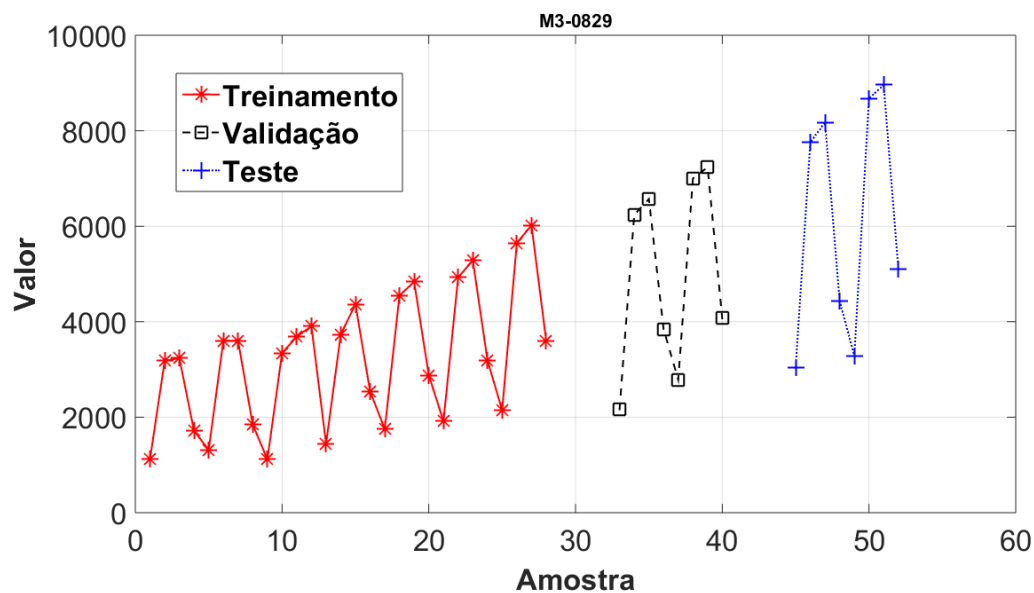


Figura 178. Conjuntos de Treinamento, Validação e Teste da Série M3-0829.

- **Série M3_0847**

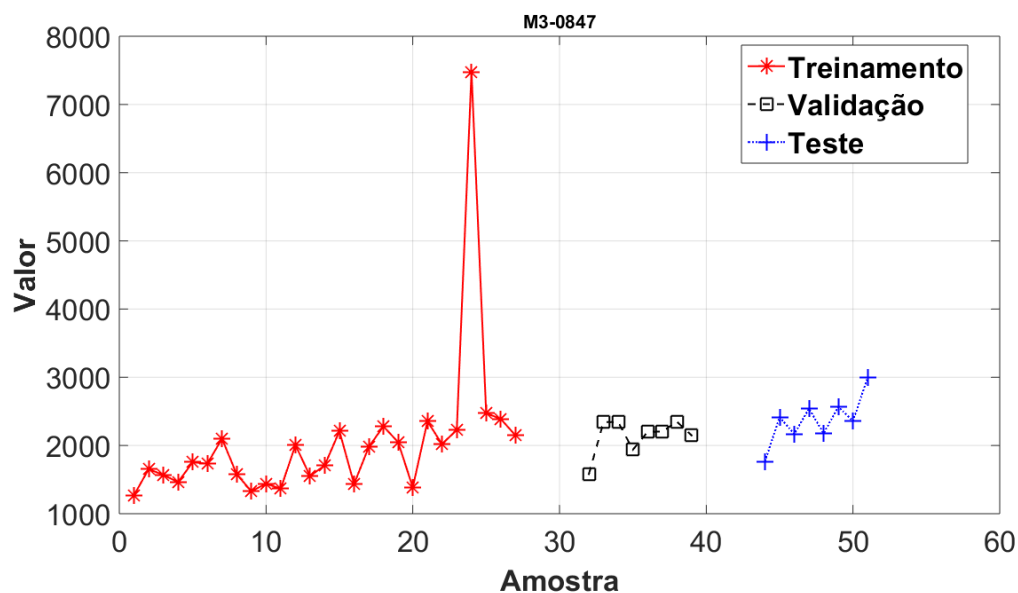


Figura 179. Conjuntos de Treinamento, Validação e Teste da Série M3-0847.

- **Série M3_0854**

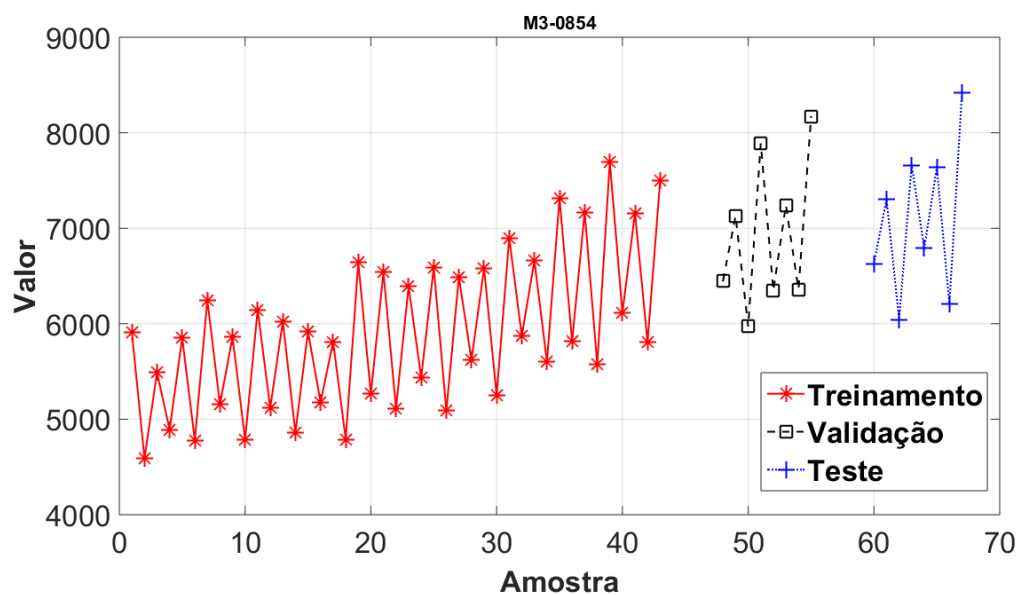


Figura 180. Conjuntos de Treinamento, Validação e Teste da Série M3-0854.

- **Série M3_0865**

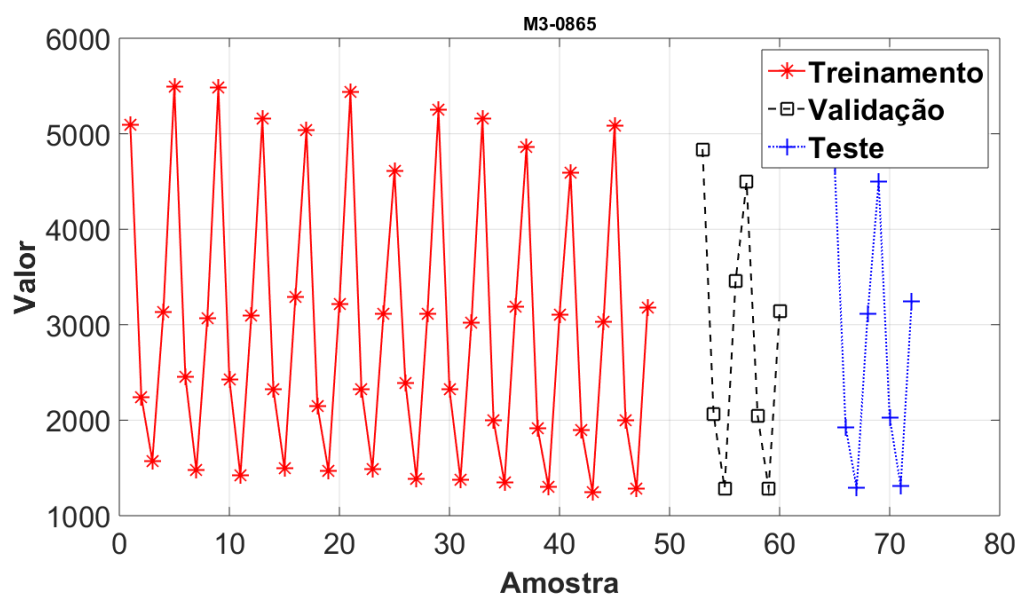


Figura 181. Conjuntos de Treinamento, Validação e Teste da Série M3-0865.

- **Série M3_0896**

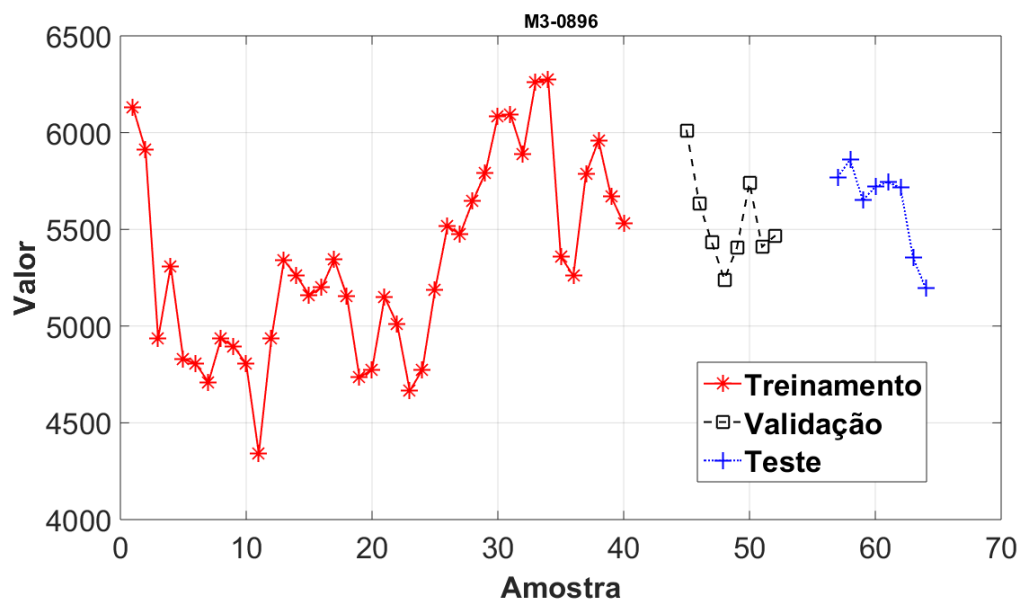


Figura 182. Conjuntos de Treinamento, Validação e Teste da Série M3-0896.

- **Série M3_1430**

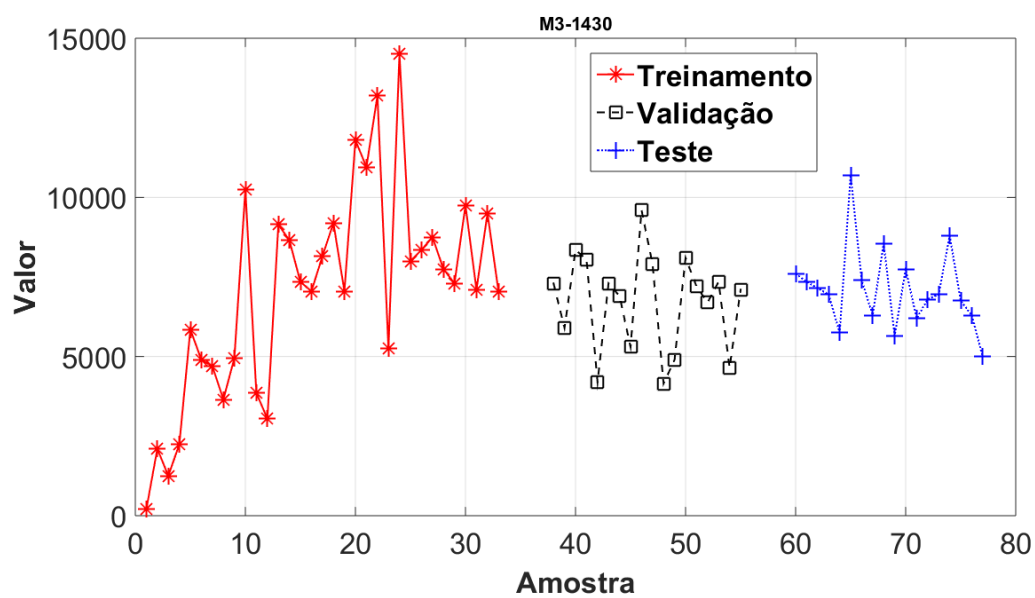


Figura 183. Conjuntos de Treinamento, Validação e Teste da Série M3-1430.

- **Série M3_1460**

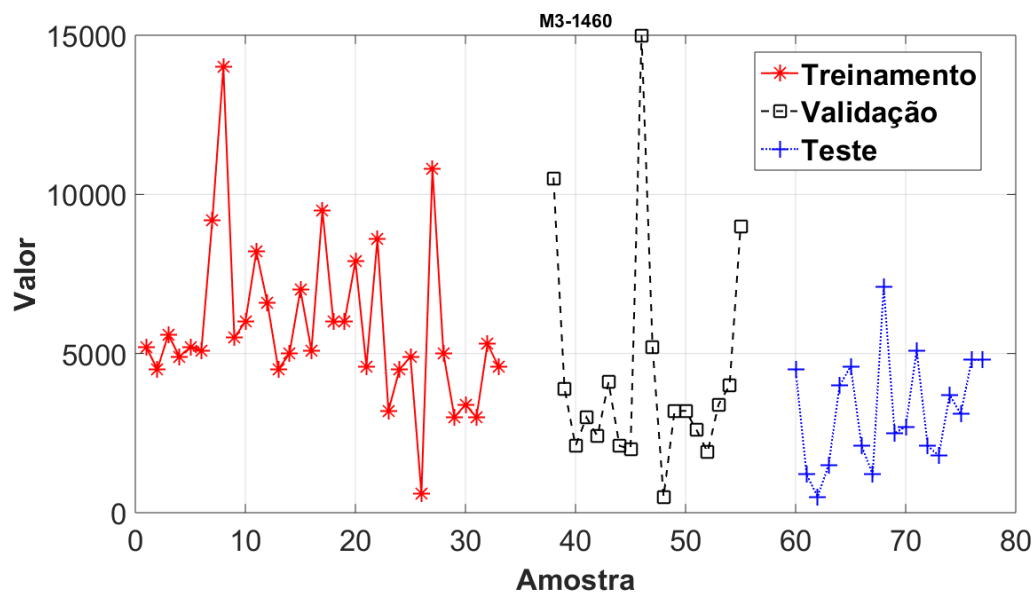


Figura 184. Conjuntos de Treinamento, Validação e Teste da Série M3-1460.

- **Série M3_1725**

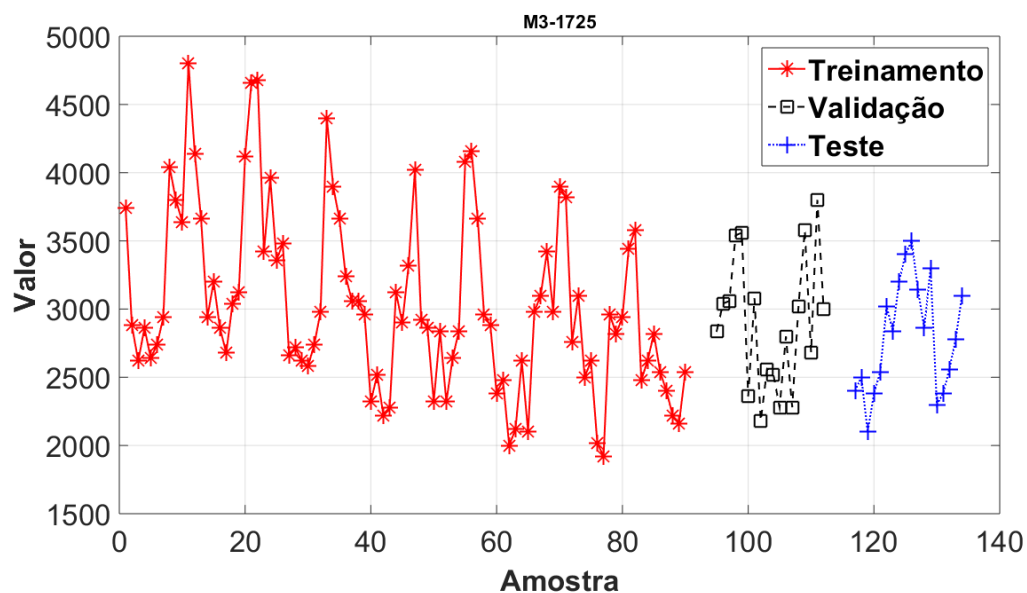


Figura 185. Conjuntos de Treinamento, Validação e Teste da Série M3-1725.

- **Série M3_1900**

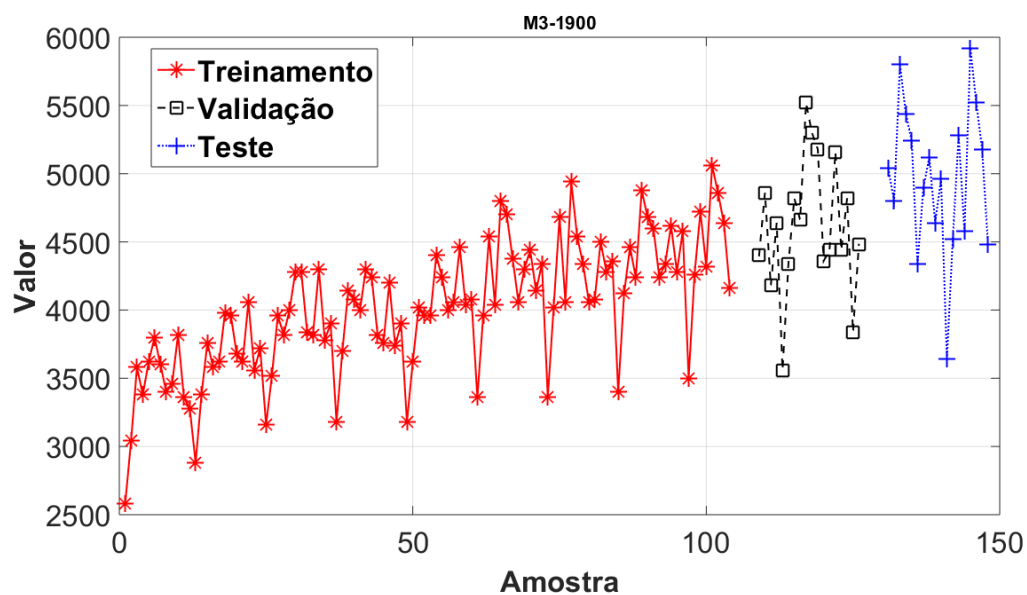


Figura 186. Conjuntos de Treinamento, Validação e Teste da Série M3-1900.

- **Série M3_1906**

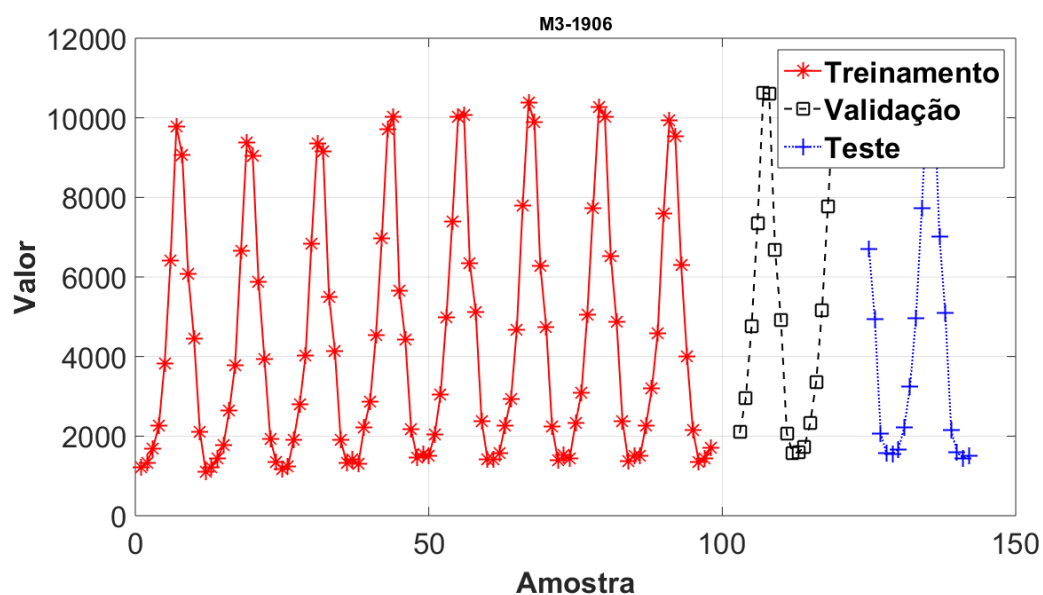


Figura 187. Conjuntos de Treinamento, Validação e Teste da Série M3-1906.

- **Série M3_1912**

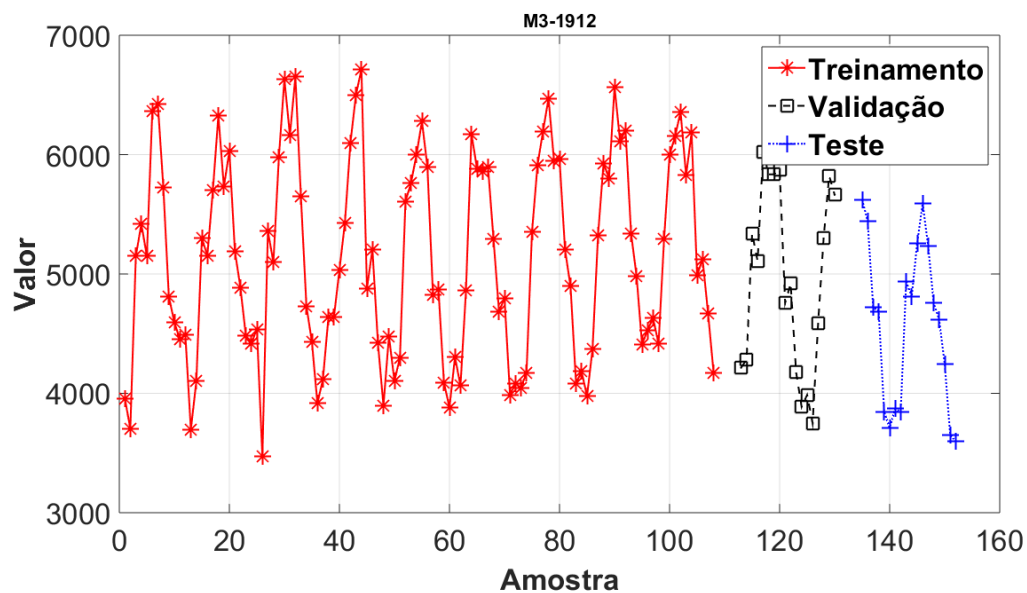


Figura 188. Conjuntos de Treinamento, Validação e Teste da Série M3-1912.

- **Série M3_1996**

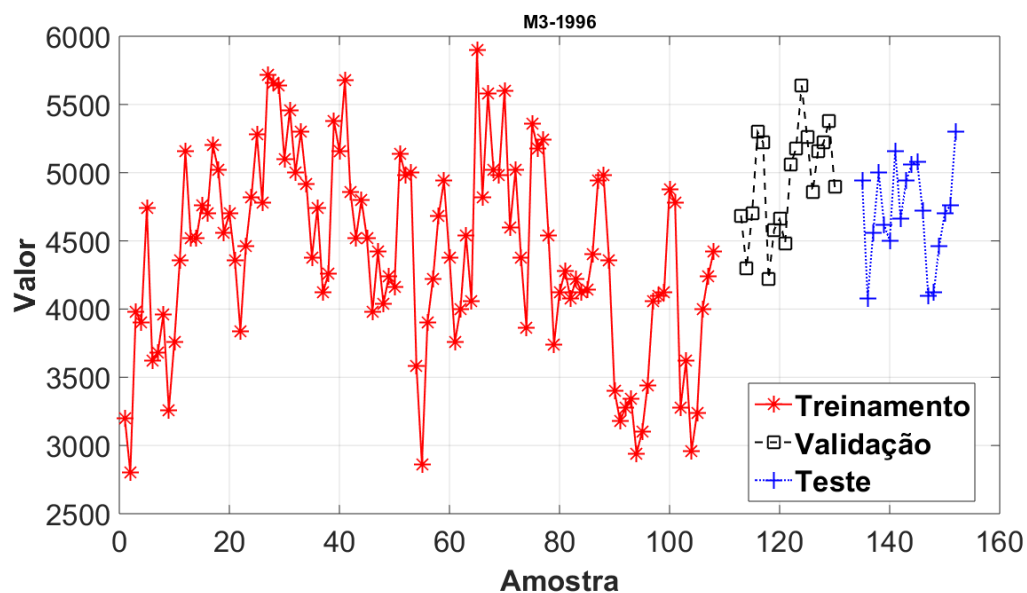


Figura 189. Conjuntos de Treinamento, Validação e Teste da Série M3-1996.

- **Série M3_2001**

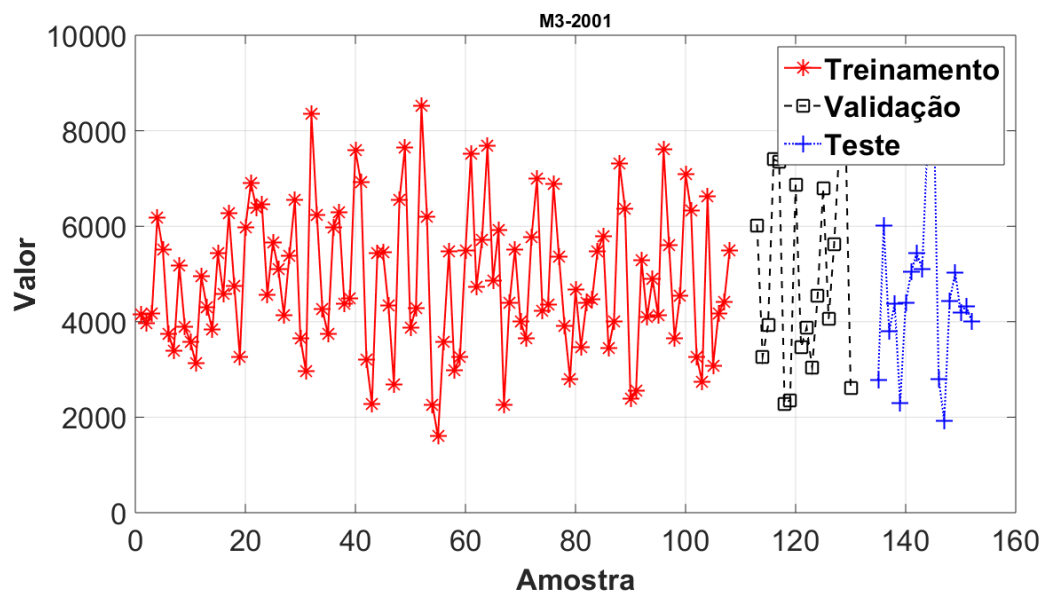


Figura 190. Conjuntos de Treinamento, Validação e Teste da Série M3-2001.

- **Série M3_2013**

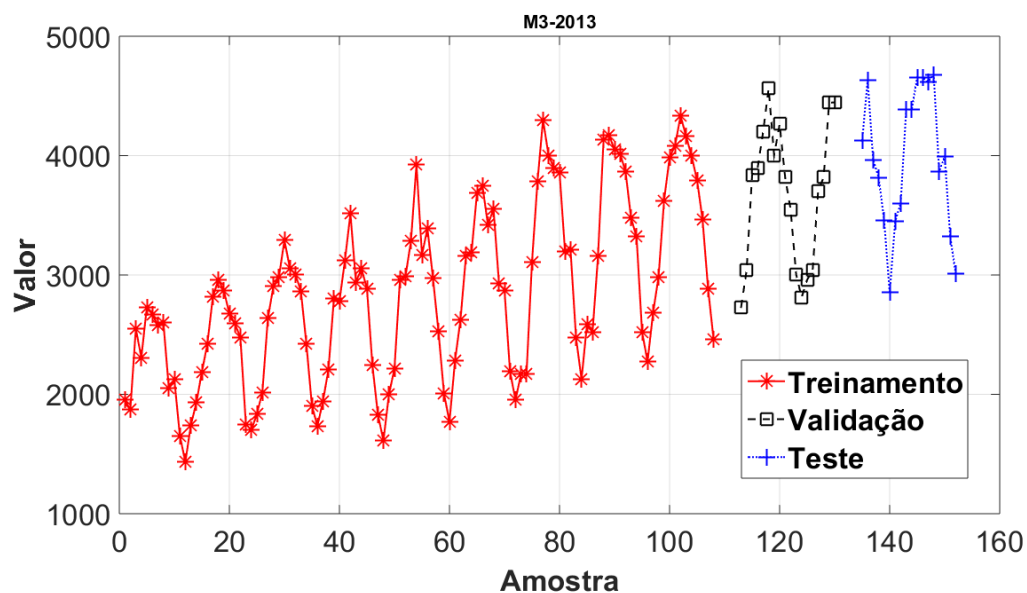


Figura 191. Conjuntos de Treinamento, Validação e Teste da Série M3-2001.

- **Série M3_2016**

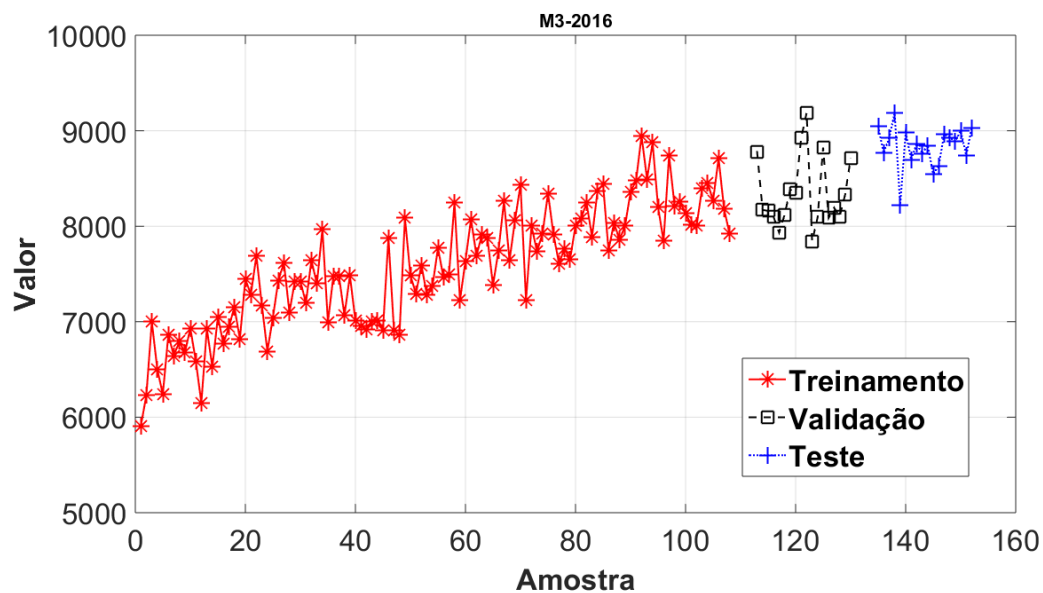


Figura 192. Conjuntos de Treinamento, Validação e Teste da Série M3-2016.

- **Série M3_2089**

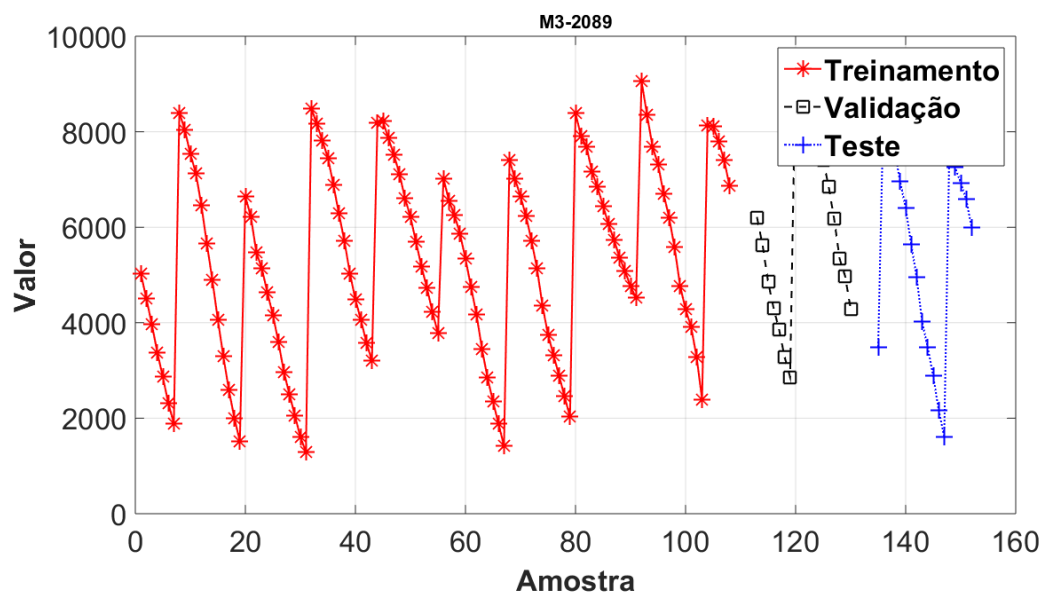


Figura 193. Conjuntos de Treinamento, Validação e Teste da Série M3-2089.

- **Série M3_2579**

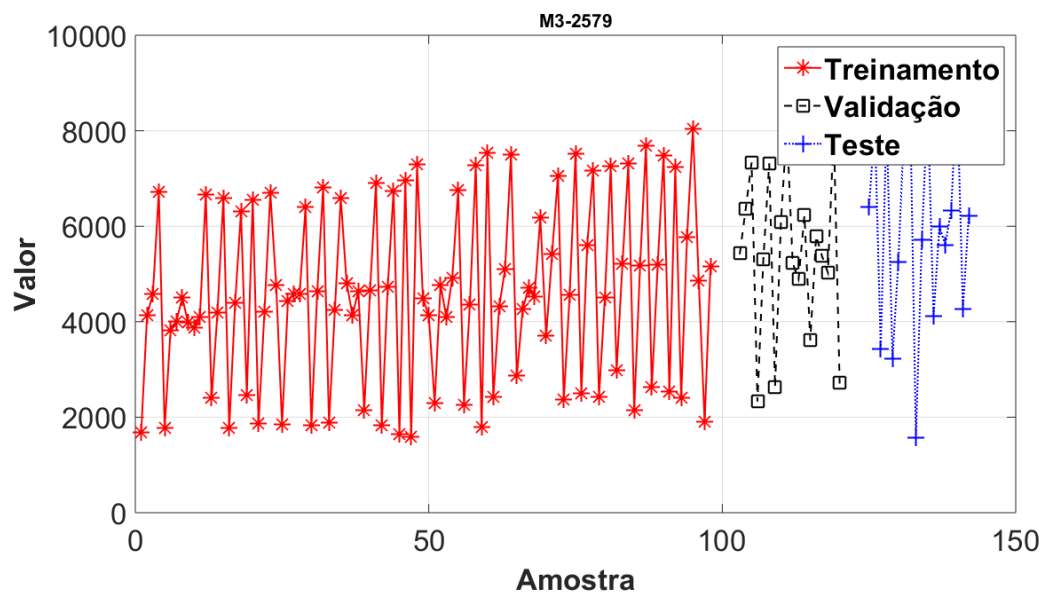


Figura 194. Conjuntos de Treinamento, Validação e Teste da Série M3-2579.