

5. Aplicação à Energia Natural Afluente

No Brasil, mais de 80% da energia elétrica gerada é de origem hidráulica (ONS, 2012). Por conta disto, a modelagem e previsão de ENA representa um papel fundamental. Para a aplicação da metodologia proposta nesta tese, as séries dos quatro subsistemas (Norte – N, Nordeste – NE, Sudeste/Centro-Oeste – SE e Sul – S) são estudadas com modelagem e previsão com e sem a filtragem SSA/MSSA. As séries de Energia Natural Afluente de janeiro de 1931 a dezembro de 2012 estão disponíveis no Operador Nacional do Sistema Elétrico (ONS). Os dados tem periodicidade mensal e são medidos em Megawatt médio (MWMed). Nas figuras 5.1, 5.2, 5.3 e 5.4 são apresentadas as séries de ENA dos quatro subsistemas do Sistema Interligado Nacional (SIN). Pode-se observar nestas figuras que o subsistema que abrange as regiões Sudeste e Centro-Oeste é o que possui a maior geração de energia hidroelétrica, enquanto o subsistema Norte apresenta menor geração. Mais detalhes sobre as séries são apresentados nas tabelas 5.1 e 5.2.

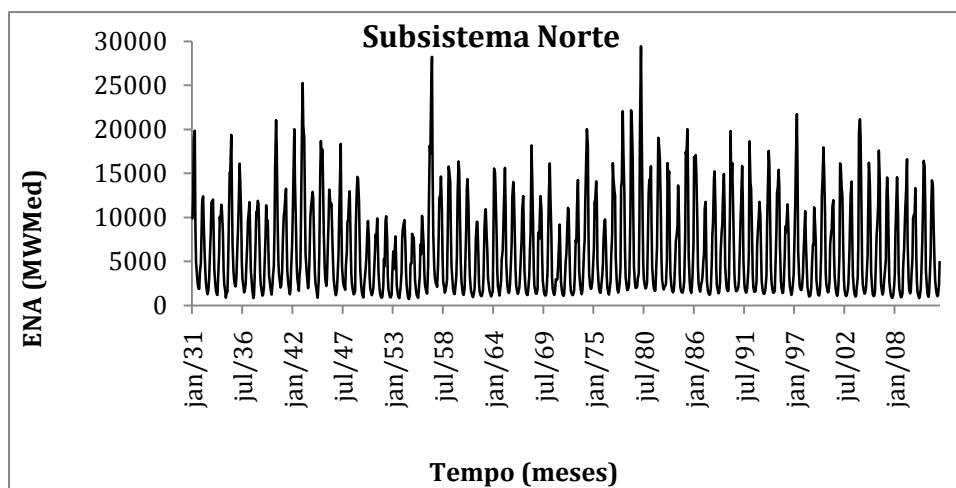


Figura 5.1. Série original das médias mensais de ENA do subsistema Norte (N).

Há uma uniformidade nos dados a exceção de alguns poucos picos.

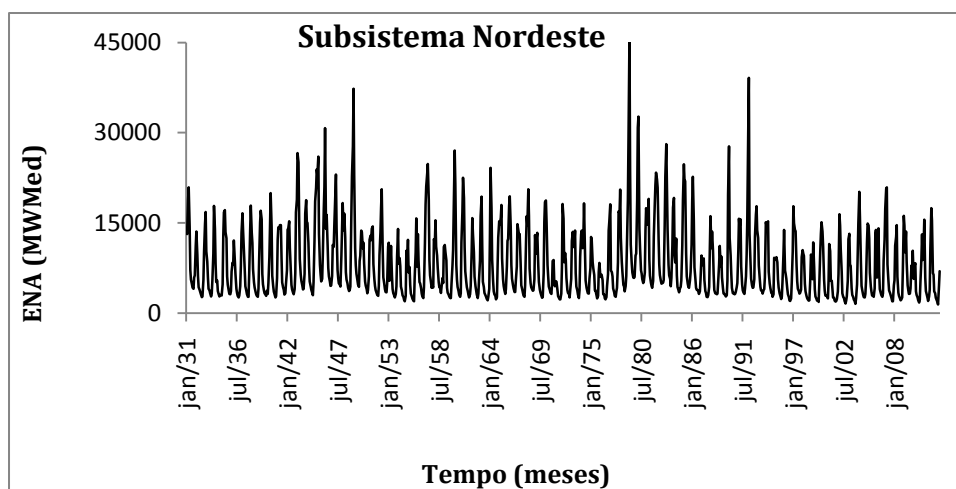


Figura 5.2. Série original das médias mensais de ENA do subsistema Nordeste (NE). Em geral, as médias mensais de ENA ficam abaixo de 20.000 MWMed sendo sua média 8.175 MWMed.

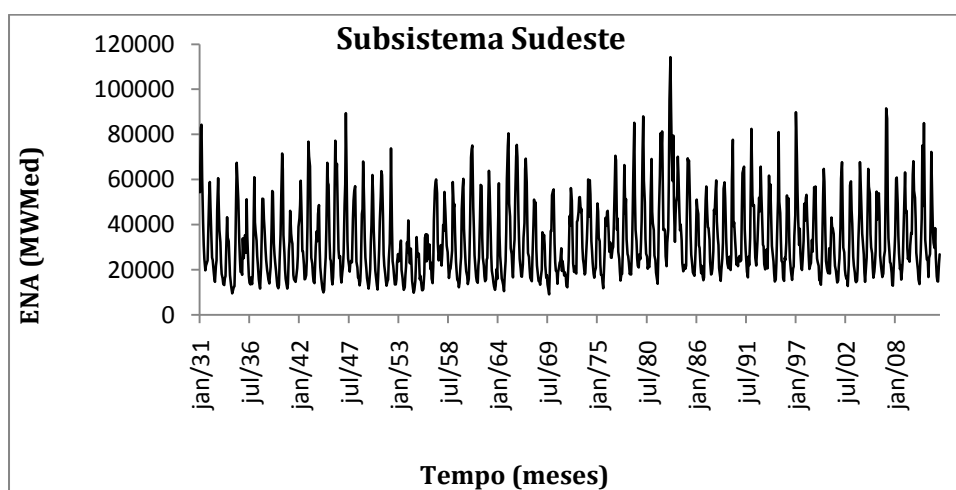


Figura 5.3. Série original das médias mensais de ENA do subsistema Sudeste/Centro-Oeste (SE). Neste subsistema, a média é bem superior às médias dos outros.

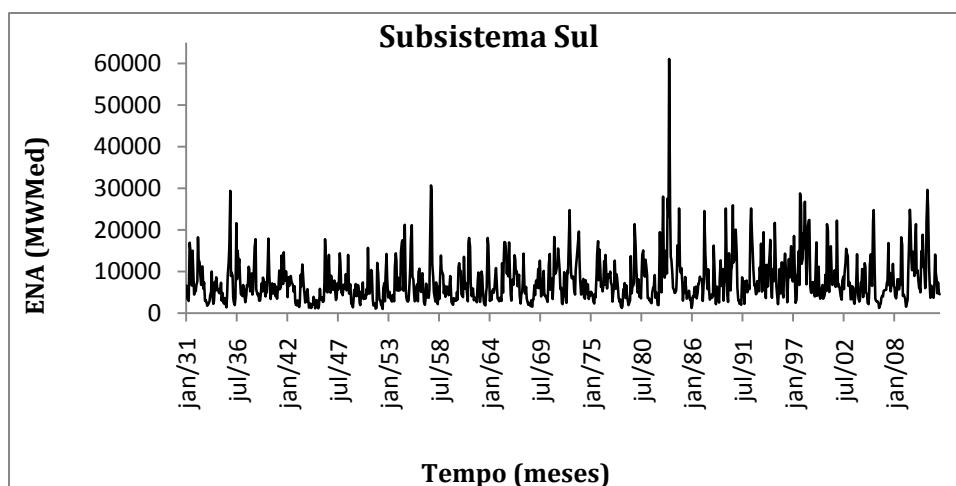


Figura 5.4. Série original das médias mensais de ENA do subsistema Sul (S).

Observa-se que uma medida em julho de 1983 bem superior as demais, que estão em geral abaixo de 20.000 MWMed.

Com o objetivo de conhecer estatisticamente as séries dos quatro subsistemas de ENA estudados, a tabela 5.1 apresenta algumas medidas estatísticas destas séries.

Tabela 5.1.: Medidas Estatísticas das quatro séries de ENA.

Estatísticas	Subsistemas de ENA			
	N	NE	SE	S
Tamanho	984	984	984	984
Média	6.276,42	8.175,30	32.993,65	7.753,11
Mediana	4.158,10	5.849,71	27.353,94	6.339,31
Mínimo	736,53	1.443,00	9.115,39	992,44
Máximo	29.424,81	46.262,92	114.307,50	61.043,77
Desvio Padrão	5.250,08	5.989,95	17.246,56	5.444,24
Assimetria	1,08	1,57	1,08	2,26
Curtose	0,56	3,45	0,82	10,99

A matriz de correlação entre os subsistemas é outra ferramenta de avaliação desta estrutura. A tabela 5.2 mostra esta matriz.

Tabela 5.2.: Correlação entre as séries de ENA dos quatro subsistemas.

Subsistemas	Subsistemas				
		N	NE	SE	S
	N	1	0,791953	0,698694	-0,294426
	NE	0,791953	1	0,795673	-0,286317
	SE	0,698694	0,795673	1	-0,086505
	S	-0,294426	-0,286317	-0,086505	1

Observa-se na tabela 5.2 que os subsistemas Norte, Nordeste e Sudeste/Centro-Oeste possuem entre si correlações fortes e positivas. Por outro lado, o subsistema Sul não possui nenhuma correlação positiva com os demais subsistemas.

Na aplicação aos quatro subsistema de ENA, foram feitas filtragens das séries originais via SSA e MSSA. Após as filtragens, as séries originais e filtradas foram modeladas via $PAR(p)$.

5.1. Modelagem $PAR(p)$ – MSSA

Antes da modelagem $PAR(p)$ para previsão e geração de cenários, as quatro séries dos subsistemas passam por uma filtragem baseada em MSSA. Para o caso MSSA, a matriz trajetória multidimensional obtida na fase de incorporação tem dimensão $(4L \times K)$. Como para cada série $T = 984$, o comprimento de janela escolhido foi $L = 492$ de acordo com os critérios preestabelecidos. Assim, a dimensão da matriz trajetória foi (1968×493) .

Desta forma, os L autovalores e seus respectivos autovetores U_l de XX' são obtidos, assim como os componentes principais $V_l = \frac{X'U_l}{\sqrt{\lambda_l}}$. No total há 492 autovetores com dimensão $4L \times 1 = 1.968 \times 1$ e 493 componentes principais com dimensão $K \times 1 = 493 \times 1$. Os nove primeiros componentes principais (vetores singulares) e sua contribuição para a série são apresentados na figura 5.5 e na figura 5.6 são apresentados alguns pares de vetores singulares que podem contribuir com a análise gráfica.

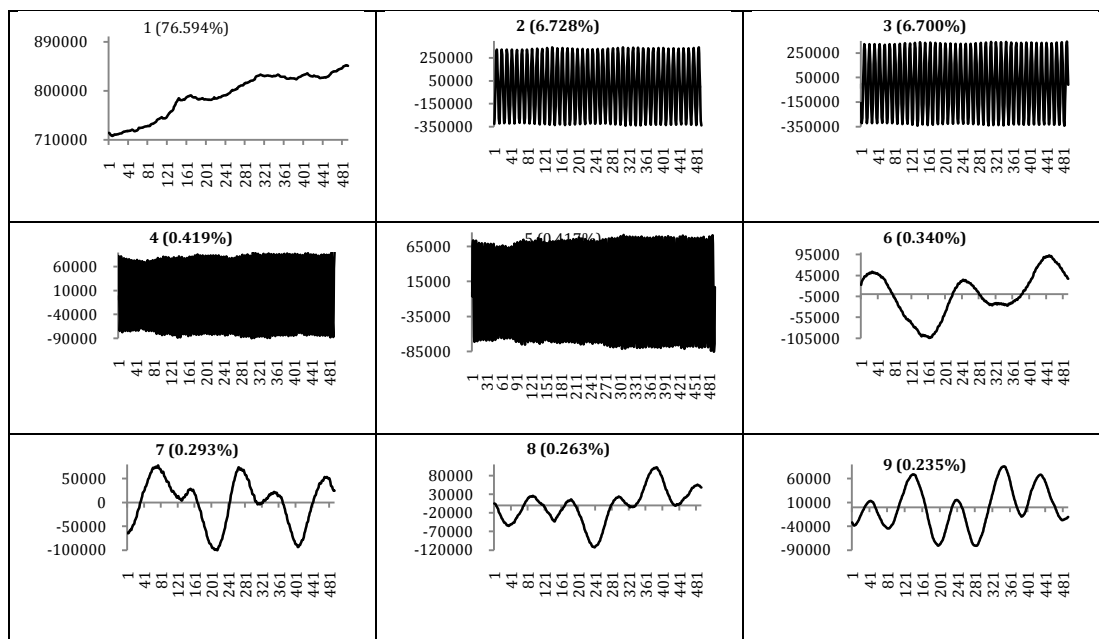


Figura 5.5: Os nove primeiros vetores singulares na filtragem MSA.

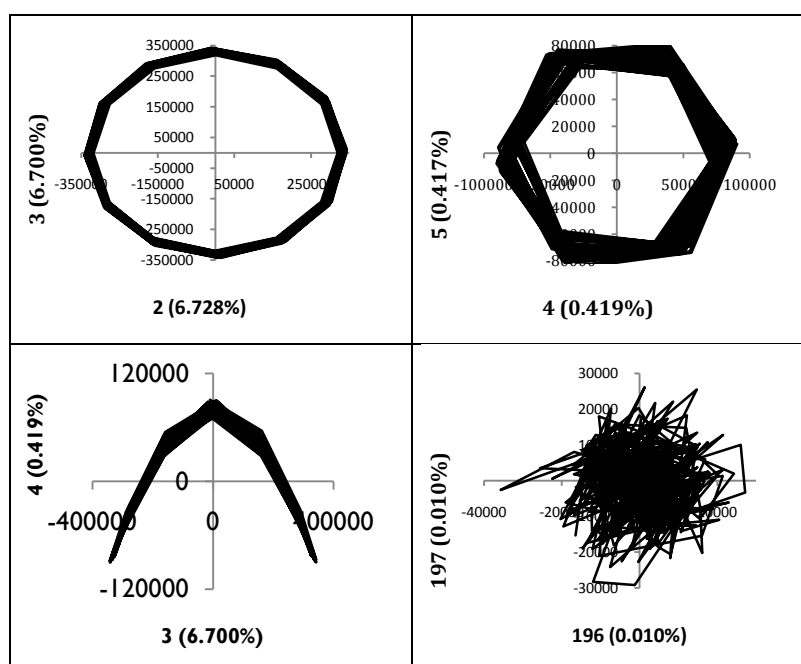


Figura 5.6: Alguns pares de vetores singulares na filtragem MSA.

Os gráficos de dispersão feitos a partir de pares de vetores singulares auxiliam na determinação das autotriplas que irão compor cada componente, sobretudo das componentes harmônica e ruidosa. Na figura 5.6 pode-se observar que os pares de vetores (2 e 3) e (4 e 5) tem um comportamento que mostram polígonos regulares conforme a teoria mostrada na seção 3.1.3.3. Na ocasião, o par (2 e 3) mostra um polígono regular de 12 lados, o que pode significar que os vetores singulares 2 e 3 sejam harmônicos de período 12. O gráfico de dispersão representado pelos vetores

(4 e 5) apresenta um polígono regular de 6 lados, levando a crer que as autotriplas associadas a estes vetores singulares são representantes da componente harmônica com período 6. Também pode-se observar que o gráfico de dispersão entre dois vetores harmônicos com períodos diferentes apresenta um comportamento diferenciado dos demais, como no par (3 e 4).

Para aprimorar esta análise gráfica dos vetores singulares, a análise de periodogramas se faz necessária na busca destas definições. As figuras 5.7 e 5.8 mostram, respectivamente, os periodogramas correlacionados aos pares (2 e 3) e (4 e 5). Tal análise confirma as suspeitas observadas nos gráficos de dispersão.

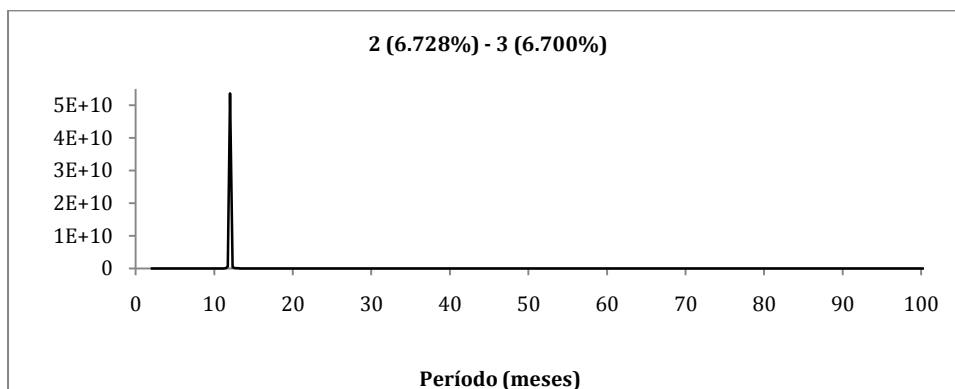


Figura 5.7: Periodograma dos vetores singulares (2 e 3) na filtragem MSSA.

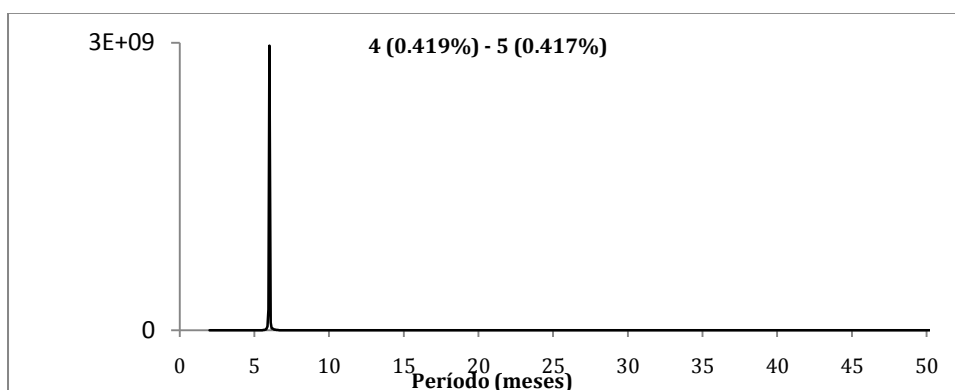


Figura 5.8: Periodograma dos vetores singulares (4 e 5) na filtragem MSSA.

O último diagrama de dispersão da figura 5.6 com os vetores (196 e 197) é um exemplo de par de vetores singulares que fazem parte da componente de ruído na decomposição MSSA.

Após a análise gráfica dos vetores singulares, as componentes foram selecionadas completando o processo de agrupamento MSSA. Tais componentes são apresentadas na tabela 5.3.

Tabela 5.3.: Componentes geradas na abordagem MSSA.

Subséries em cada componente	
Tendência	1, 6 - 11
Harmônica	2 - 5, 12 - 19
Ruído	20 - 492

Os gráficos das respectivas componentes de tendência nos quatro subsistemas são apresentados na figura 5.9, enquanto que os gráficos das componentes harmônicas estão na figura 5.10 e as componentes ruidosas estão representadas na figura 5.11.

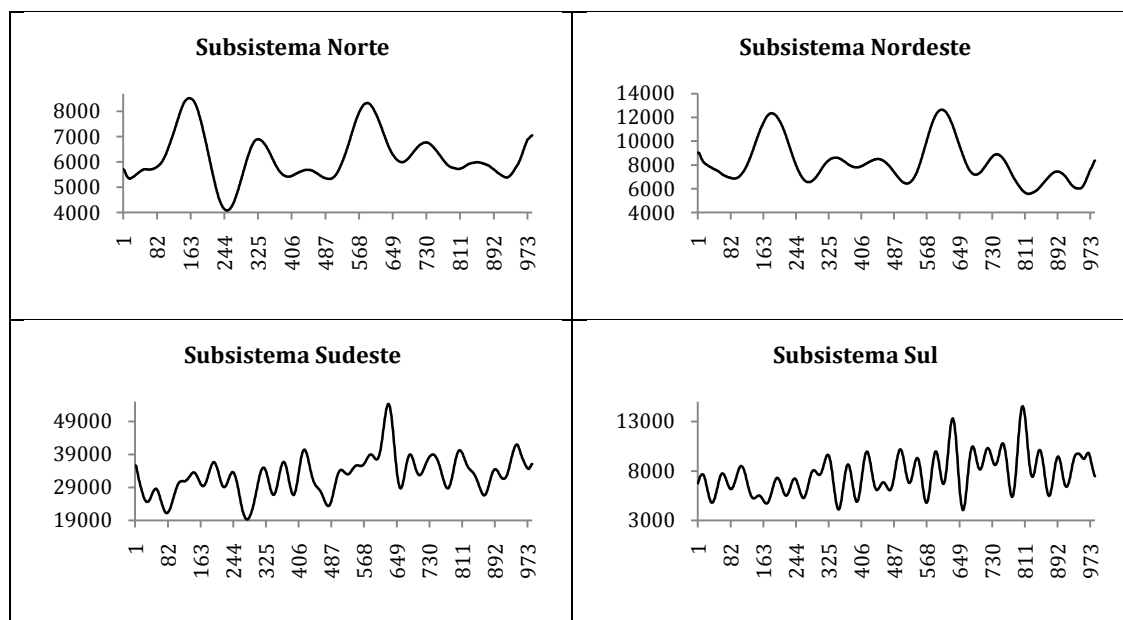


Figura 5.9: Componentes de tendência dos quatro subsistemas na abordagem MSSA.

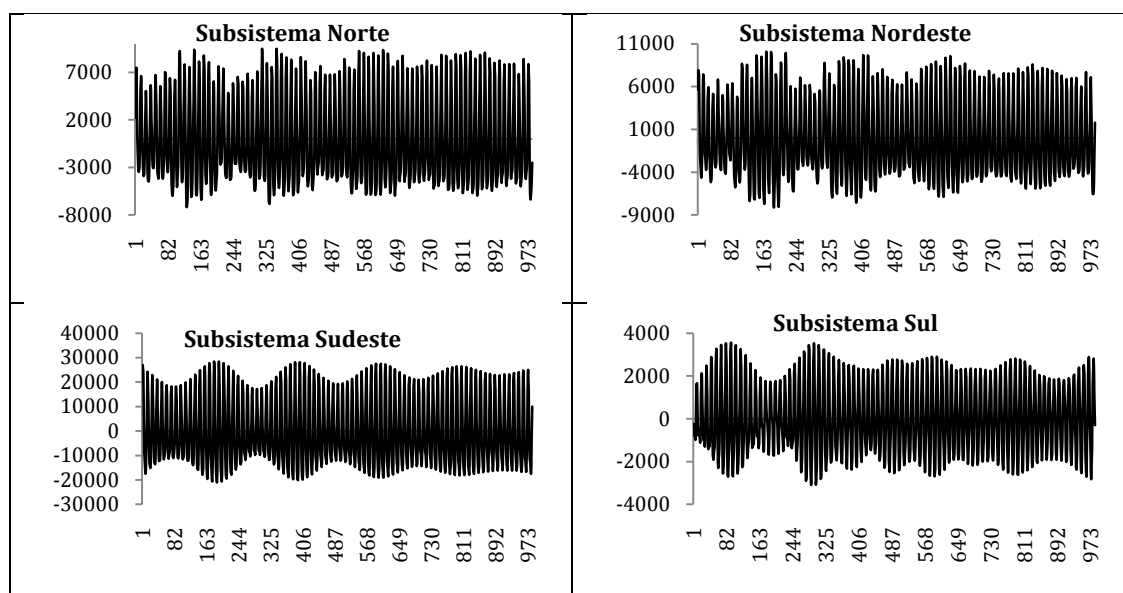


Figura 5.10: Componentes harmônicas dos quatro subsistemas na abordagem MSSA.

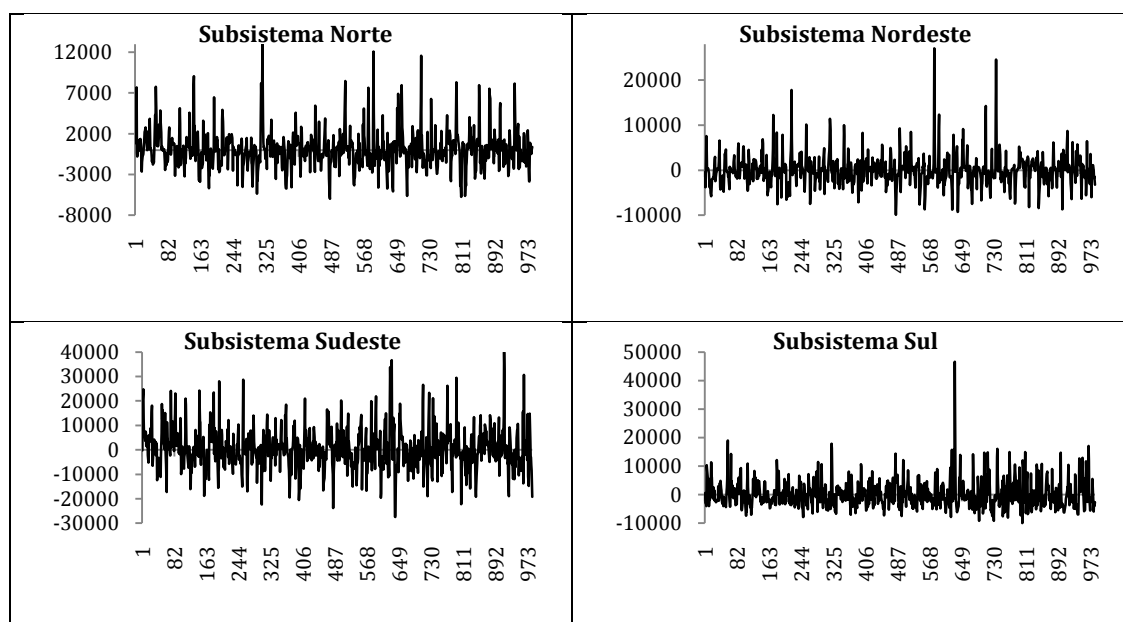


Figura 5.11: Componentes ruidosas dos quatro subsistemas na abordagem MSSA.

Com o objetivo de verificar se as componentes estão bem separadas, a correlação ponderada é aplicada entre elas diretamente na fase de agrupamento MSSA. A tabela 5.4 mostra os resultados desta correlação. Nesta tabela pode-se ver que as componentes estão bem separadas, indicando que haveria o mínimo de ruído nas componentes: tendência e harmônica.

Tabela 5.4.: Correlação ponderada entre as componentes MSSA.

Componentes	Tendência	Harmônica	Ruído
Tendência	1	0,001	0,011
Harmônica	0,001	1	0,042
Ruído	0,011	0,042	1

Com a finalidade de verificar e comprovar que a componente ruidosa realmente representa um ruído, o teste BDS foi aplicado a estas componentes em cada subsistema. A tabela 5.5 apresenta os resultados do teste BDS aplicado a série de ruídos dos quatro subsistemas. Os resultados nesta tabela mostram que a hipótese de independência não é rejeitada ao nível de 5% de significância em todas as dimensões medidas, comprovando serem ruidosas as componentes avaliadas neste teste.

Tabela 5.5.: Teste BDS para as séries de ruídos dos quatro subsistemas.

N	Dim	Estatística BDS	Estatística Z	Erro padrão	p-valor
	2	1.79E-7	0.012852	1.67E-5	0.9697
	3	5.30E-7	0.017004	3.74E-5	0.9864
	4	1.71E-6	0.020036	6.24E-5	0.9840
	5	1.71E-6	0.022483	9.13E-5	0.9821
	6	-9.92E-6	-0.096179	0.000123	0.9234
NE	Dim	Estatística BDS	Estatística Z	Erro padrão	p-valor
	2	3.09E-6	0.132011	4.29E-5	0.8950
	3	5.42E-6	0.100331	9.58E-5	0.9201
	4	1.46E-5	0.167668	0.000160	0.8668
	5	2.51E-5	0.196610	0.000234	0.8441
	6	3.45E-5	0.199968	0.000316	0.8415
SE	Dim	Estatística BDS	Estatística Z	Erro padrão	p-valor
	2	-4.13E-09	-1.060615	3.89E-09	0.2889
	3	-2.67E-13	-0.028212	9.45E-12	0.9775
	4	-1.73E-17	-0.001006	1.72E-14	0.9992
	5	-1.13E-21	-4.12E-5	2.74E-17	0.9999
	6	-7.37E-26	-1.83E-6	4.03E-20	0.9999
S	Dim	Estatística BDS	Estatística Z	Erro padrão	p-valor
	2	-1.41E-6	-0.025930	5.50E-05	0.9793
	3	-6.29E-6	-0.051989	0.000123	0.9585
	4	-2.03E-5	-0.062260	0.000205	0.9504
	5	-2.03E-5	-0.068674	0.000300	0.9452
	6	-3.15E-5	-0.078835	0.000406	0.9372

Depois de verificar as componentes ruidosas através do teste BDS nos quatro subsistemas, a mesma é removida da série e uma série filtrada é obtida somando as componentes tendência e harmônica. As figuras 5.12, 5.13, 5.14 e 5.15 apresentam a sobreposição da série original pela série filtrada via MSSA nos quatro subsistemas.

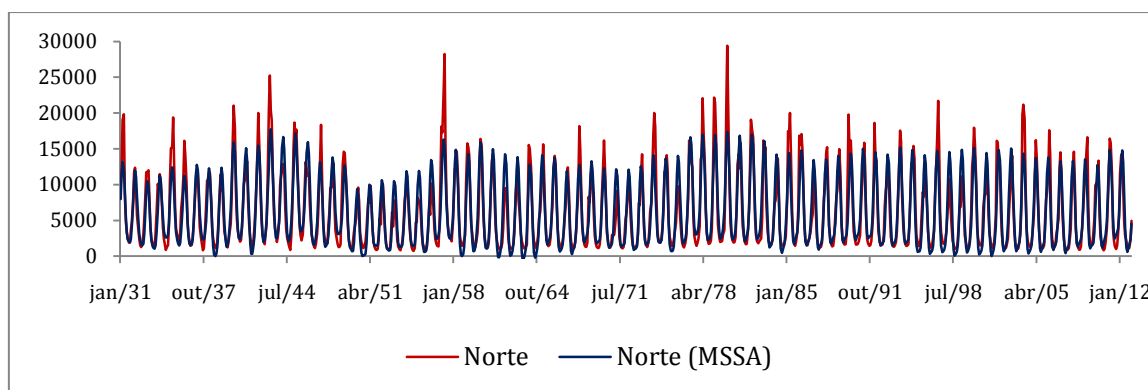


Figura 5.12: Série original e série filtrada via MSSA do subsistema Norte.

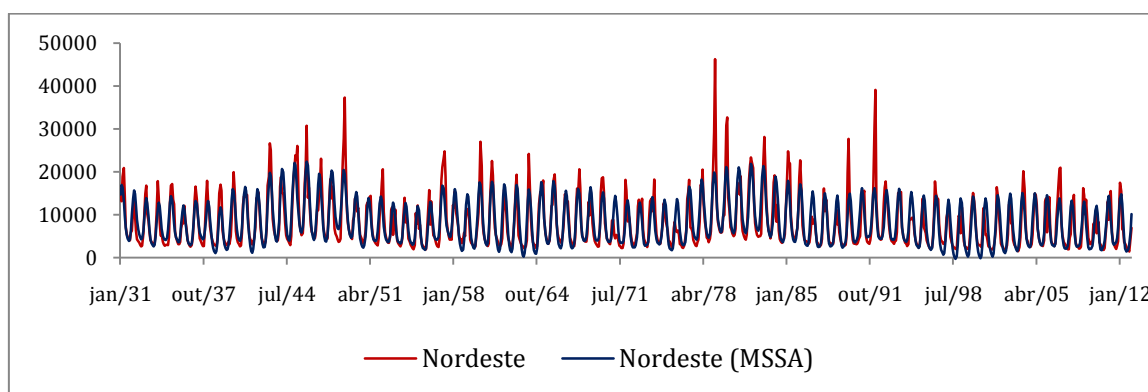


Figura 5.13: Série original e série filtrada via MSSA do subsistema Nordeste.

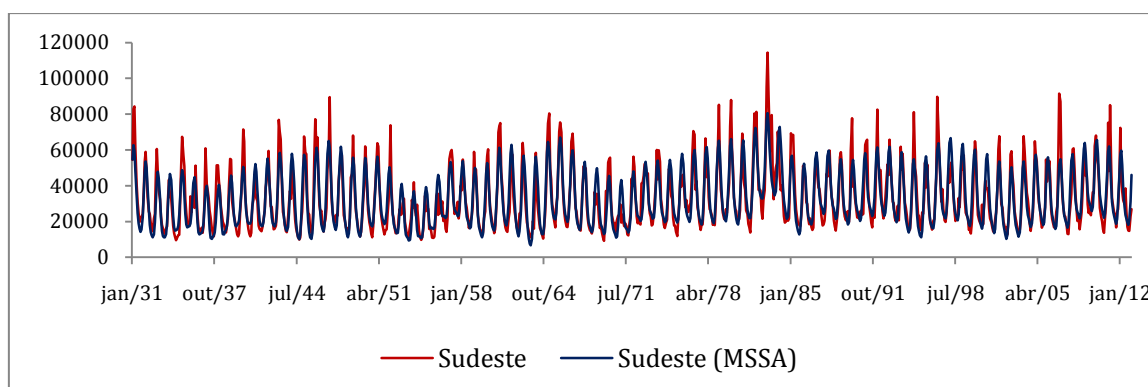


Figura 5.14: Série original e série filtrada via MSSA do subsistema Sudeste/Centro-Oeste.

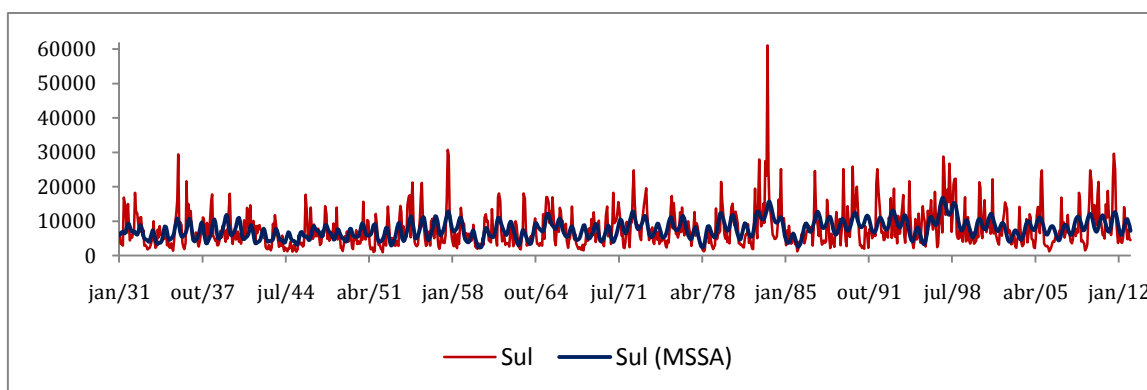


Figura 5.15: Série original e série filtrada via MSSA do subsistema Sul.

Com as séries dos quatro subsistemas originais e filtradas via MSSA, fez-se uma análise mensal da média (μ) e do desvio padrão (σ). Antes de entrar em período seco, a primeira coisa que se percebe nas figuras 5.16, 5.17, 5.18 e 5.19 é que na série original há uma heterocedasticidade, enquanto na série filtrada há uma homocedasticidade, ou seja, a filtragem via MSSA homogeneizou a variância.

A filtragem via MSSA contribui para a modelagem da série analisada, sobretudo quando se observa que a maior parte dos modelos de previsão admite a premissa de homocedasticidade.

No período úmido (dezembro a abril) há uma maior volatilidade que no período seco (maio a novembro), ou seja, há uma heterocedasticidade na série original, mas ao filtrar pela SSA a série se torna homocedástica, um requisito em muitas abordagens de previsão.

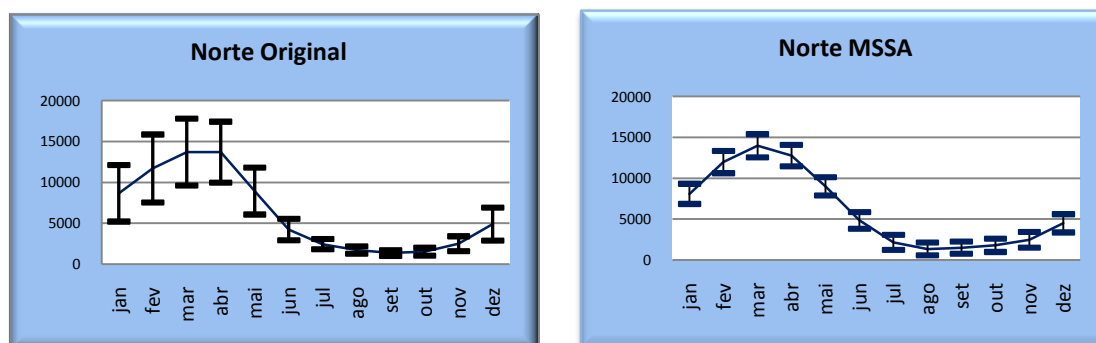


Figura 5.16: Comparação gráfica entre os intervalos ($\mu \pm \sigma$) na série original e na série filtrada MSSA do subsistema Norte.

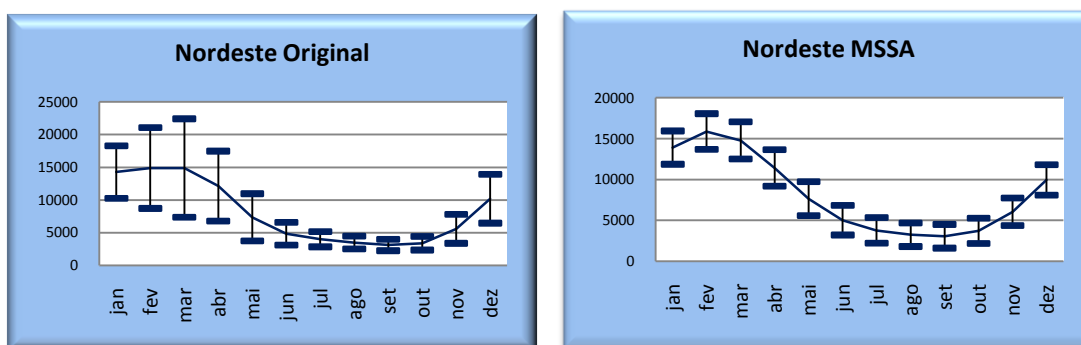


Figura 5.17: Comparação gráfica entre os intervalos ($\mu \pm \sigma$) na série original e na série filtrada MSSA do subsistema Nordeste.

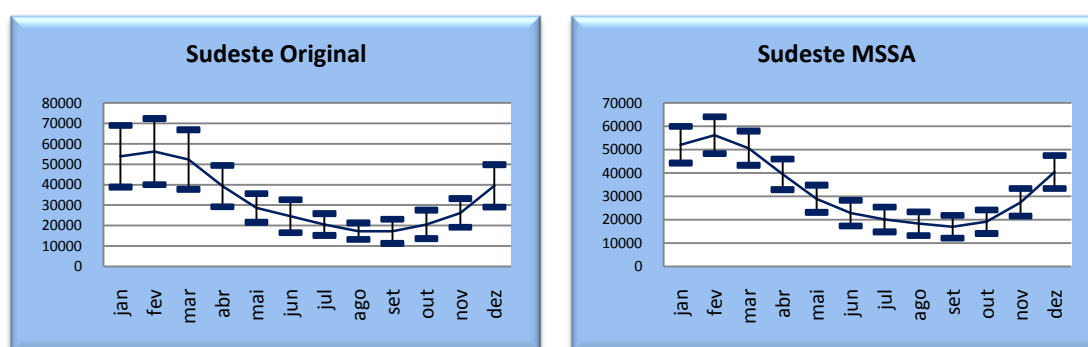


Figura 5.18: Comparação gráfica entre os intervalos ($\mu \pm \sigma$) na série original e na série filtrada MSSA do subsistema Sudeste/Centro-Oeste.

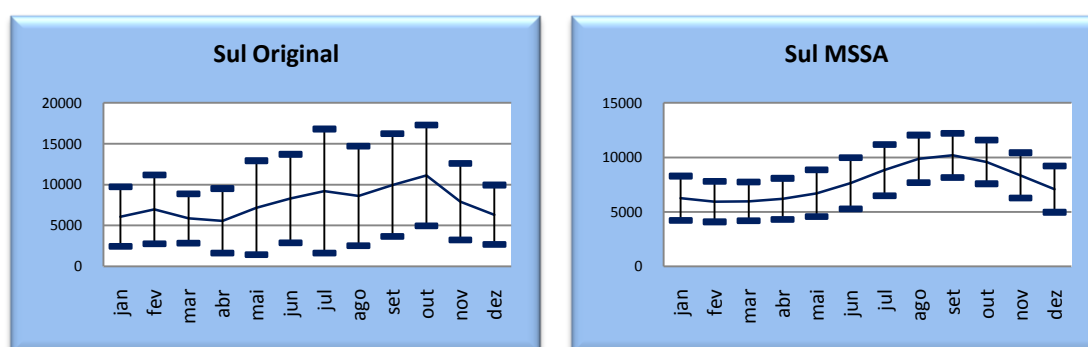


Figura 5.19: Comparação gráfica entre os intervalos ($\mu \pm \sigma$) na série original e na série filtrada MSSA do subsistema Sul.

Sobre as figuras 5.16 – 5.19 pode-se observar que, à exceção do subsistema Sul, que tem o seu pico no mês de setembro, os demais subsistemas tem seu pico durante o verão, entre os meses de fevereiro e março. Também nota-se que o subsistema Sudeste/Centro-Oeste tem a maiores médias de ENA, chegando a mais de

55.000 MWMed. A filtragem MSSA suavizou a curva e manteve as variâncias constantes.

De posse das séries originais e filtradas via MSSA nos quatro subsistemas, as mesmas foram modeladas via $PAR(p)$. Para avaliar a consistência dos modelos obtidos, foi aplicada a medida estatística MAPE, definida em (51), *in sample*. A tabela 5.6 apresenta resultados desta medida para os quatro subsistemas, respectivamente.

Tabela 5.6.: MAPE (%) *in sample* dos meses para os modelos $PAR(p)$ e $PAR(p)$ - MSSA nos quatro subsistemas.

MAPE (%)	Norte		Nordeste		Sudeste/Centro-Oeste		Sul	
Mês	$PAR(p)$	$PAR(p)$ MSSA	$PAR(p)$	$PAR(p)$ MSSA	$PAR(p)$	$PAR(p)$ MSSA	$PAR(p)$	$PAR(p)$ MSSA
Jan	21,3150	4,0486	20,087	4,0741	19,099	3,5581	43,863	6,1534
Fev	22,4870	3,2483	29,174	4,4951	20,425	3,7088	42,421	5,4864
Mar	14,0090	2,0893	24,617	4,4224	17,947	3,2184	31,303	8,8273
Abr	14,1460	2,3321	29,199	5,0078	12,334	3,2942	45,127	14,769
Mai	12,9600	2,8843	14,683	6,9024	10,379	4,6633	79,801	8,0755
Jun	10,5110	5,5560	7,271	10,9270	10,441	5,1704	55,224	5,1090
Jul	6,8260	4,9606	5,185	12,6700	8,007	5,5021	46,304	5,1961
Ago	5,7416	10,3510	4,522	12,3460	9,241	7,1534	67,320	4,3921
Set	8,3045	21,1750	6,771	11,9360	14,192	8,0904	51,134	4,2216
Out	13,2920	9,5091	13,626	5,9950	18,253	5,9805	48,867	3,2251
Nov	19,9250	5,7945	24,585	7,5249	13,677	5,5263	12,127	4,4839
Dez	22,5470	5,5226	25,193	4,2901	15,421	4,6130	46,571	7,2572

Conforme observado na tabela 5.6, a abordagem MSSA mostrou um poder preditivo maior na maioria dos meses avaliados por apresentar um valor de MAPE *in sample* menor que os apresentados pela modelagem feita com a série original, sendo superado apenas nos subsistemas Norte nos meses agosto e setembro e Nordeste nos meses junho, julho, agosto e setembro, que coincide com as figuras 5.16 e 5.17 no que se refere às variâncias menores nas séries originais destes dois subsistemas. O resultado mais significativo, onde a modelagem $PAR(p)$ -MSSA obteve melhor

desempenho, está no subsistema Sul, onde o MAPE nesta modelagem é muito menor que o MAPE da modelagem na abordagem sem a filtragem.

Em SOUZA (2013), o autor usa o encolhimento wavelet (MORETTIN, 1997) antes da modelagem PAR (p) e também verifica a medida MAPE no ajuste das séries de ENA. Na ocasião os dados observados variam de janeiro de 1931 a dezembro de 2010, enquanto que os dados medidos na filtragem MSSA variam de janeiro de 1931 a dezembro de 2012. Embora não sejam exatamente os mesmos anos, a ilustração é um exemplo da capacidade preditiva da abordagem MSSA diante de outra metodologia bastante usada na literatura. A tabela 5.7 apresenta a comparação desempenho da filtragem via MSSA contra o desempenho do encolhimento wavelet através desta medida mês a mês nos quatro subsistemas. É possível observar que o desempenho da filtragem MSSA se mostra mais eficiente que o do encolhimento wavelet, pois os valores de MAPE são bem inferiores no primeiro caso na maioria das situações.

Tabela 5.7.: MAPE (%) *in sample* dos meses para os modelos PAR(p) – Wavelet e PAR(p) - MSSA nos quatro subsistemas.

MAPE (%)	Norte		Nordeste		Sudeste/Centro-Oeste		Sul	
Mês	PAR(p) Wave	PAR(p) MSSA	PAR(p) Wave	PAR(p) MSSA	PAR(p) Wave	PAR(p) MSSA	PAR(p) Wave	PAR(p) MSSA
Jan	21,70	4,05	19,04	4,07	17,66	3,56	42,48	6,15
Fev	22,03	3,25	26,97	4,50	19,31	3,71	42,75	5,49
Mar	14,63	2,09	22,96	4,42	16,44	3,22	31,58	8,83
Abr	13,30	2,33	26,92	5,01	11,51	3,29	45,07	14,77
Mai	11,08	2,88	14,81	6,90	10,24	4,66	80,13	8,06
Jun	8,80	5,56	8,17	10,93	10,36	5,17	52,13	5,11
Jul	5,88	4,96	5,02	12,67	7,53	5,50	44,32	5,20
Ago	4,88	10,35	4,36	12,35	8,86	7,15	67,27	4,39
Set	7,38	21,17	6,56	11,94	12,85	8,09	50,80	4,22
Out	12,46	9,51	12,53	5,99	17,06	5,98	48,83	3,26
Nov	17,24	5,79	22,59	7,52	13,67	5,53	42,85	4,48
Dez	22,52	5,52	24,84	4,29	14,63	4,61	46,69	7,26

5.2. Geração de Cenários PAR(p) – MSSA

A geração dos cenários hidrológicos foi feita segundo os critérios estabelecidos na seção 2.4.7. Foram gerados 5.000 cenários para um horizonte de 5 anos (60 meses). Foram considerados cenários para as séries originais modeladas via PAR(p) e para as séries modeladas via PAR(p) – MSSA. Com o objetivo avaliar se as séries sintéticas possuem comportamento estatístico idêntico ao da série histórica, foram calculados a média e o desvio-padrão dos cenários gerados a partir dos modelos PAR(p) e PAR(p) – MSSA e para obter os cinco anos de série histórica, foi calculada a média e o desvio-padrão de cada mês nesta série e com isso, gerado um ano de série de médias históricas e o mesmo para o desvio-padrão. Para obter os 60 meses necessários para a comparação com as séries sintéticas, as séries de médias e desvios-padrão históricos foram replicadas 5 vezes. Desta forma, as séries que aparecem nas figuras 5.20 - 5.23 representando as séries de médias históricas apresentam 5 configurações iguais. Estas figuras apresentam comparações entre as médias históricas obtidas como descrito anteriormente, as médias e dos cenários obtidos através da modelagem PAR(p) e as médias dos cenários obtidos através da modelagem PAR(p) – MSSA.

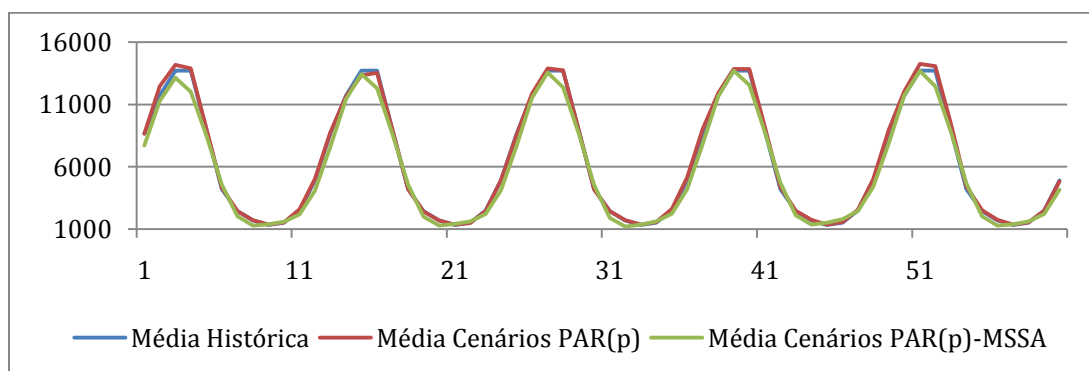


Figura 5.20: Comparação entre as médias dos cenários e a média histórica do subsistema Norte.

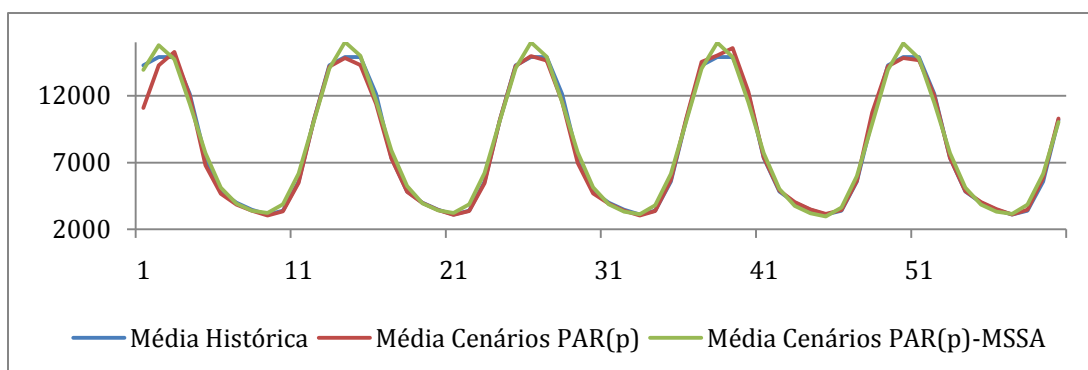


Figura 5.21: Comparação entre as médias dos cenários e a média histórica do subsistema Nordeste.

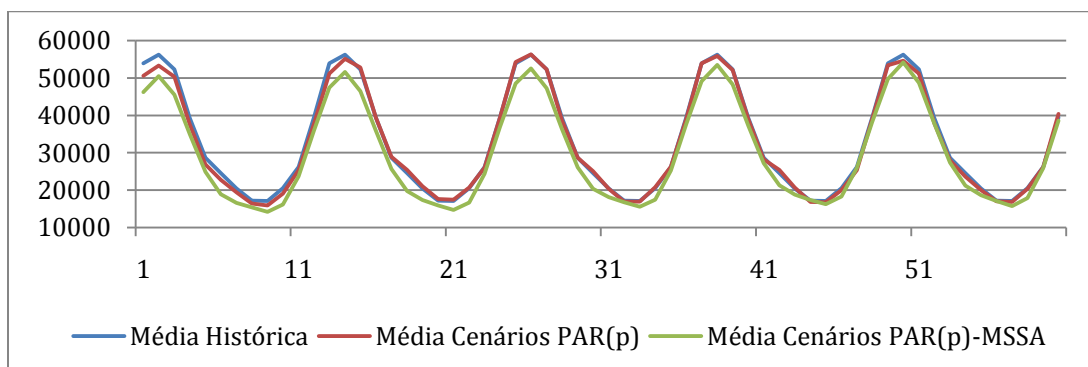


Figura 5.22: Comparação entre as médias dos cenários e a média histórica do subsistema Sudeste/Centro-Oeste.

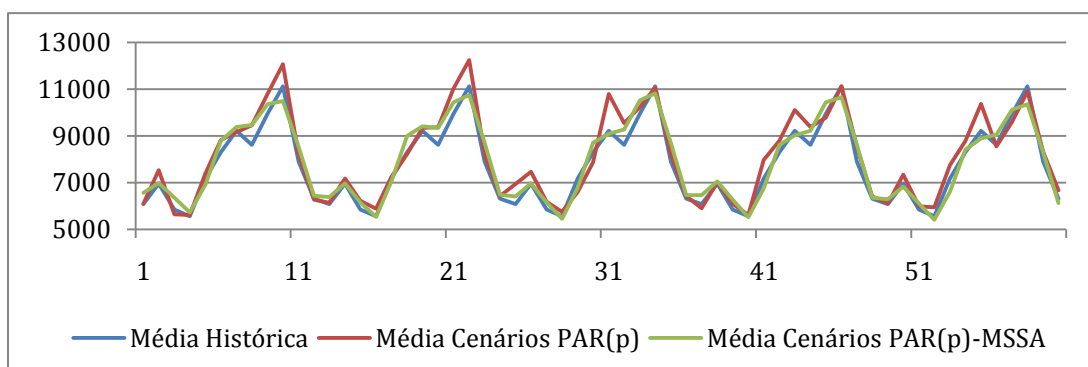


Figura 5.23: Comparação entre as médias dos cenários e a média histórica do subsistema Sul.

Com a filtragem MSSA, os desvios-padrão dos cenários gerados a partir da modelagem $PAR(p)$ – MSSA estão muito abaixo dos desvios-padrão da série histórica original. Por conta disso, para a comparação dos desvios-padrão obtidos através desta modelagem é necessário adicionar as séries de ruídos retiradas na filtragem MSSA aos cenários gerados nesta modelagem. Com isso, os desvios-padrão, que em princípio eram bem inferiores, se aproximam dos equivalentes

obtidos nas séries históricas. Para incluir estes ruídos nos cenários, a série de ruídos de comprimento 984 foi redistribuída mês a mês de modo a obter 82 séries de ruídos mensais. Em seguida, foi realizado um sorteio aleatório dos 5 anos de ruídos através do processo Bootstrap. Este procedimento foi feito 5.000 vezes. Com isso, obtém-se 5.000 séries de ruídos de comprimento 60 seguindo a mesma distribuição de probabilidade que a série original de ruídos extraídos na filtragem MSSA. Estes ruídos são adicionados aos cenários e os seus desvios-padrão são comparados com os desvios-padrão históricos e dos cenários gerados a partir da modelagem PAR(p). As figuras 5.24 – 5.27 mostram estas comparações entre desvios-padrão.

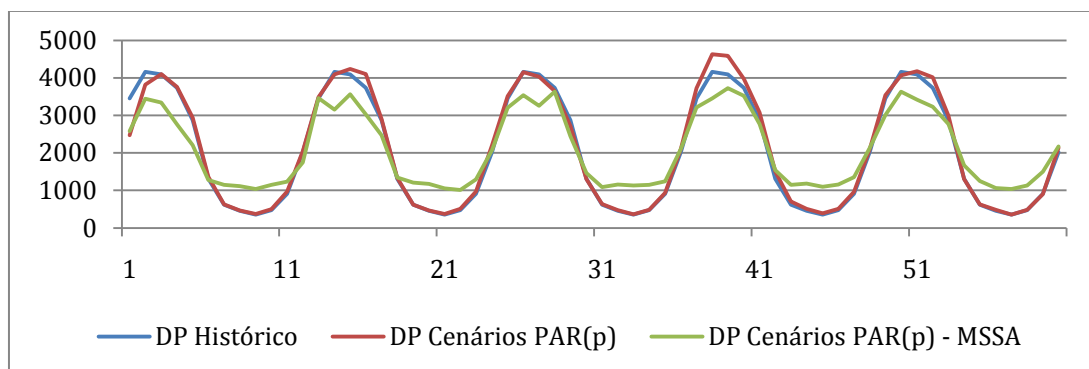


Figura 5.24: Comparação entre os desvios-padrão dos cenários e o desvio-padrão histórico do subsistema Norte.

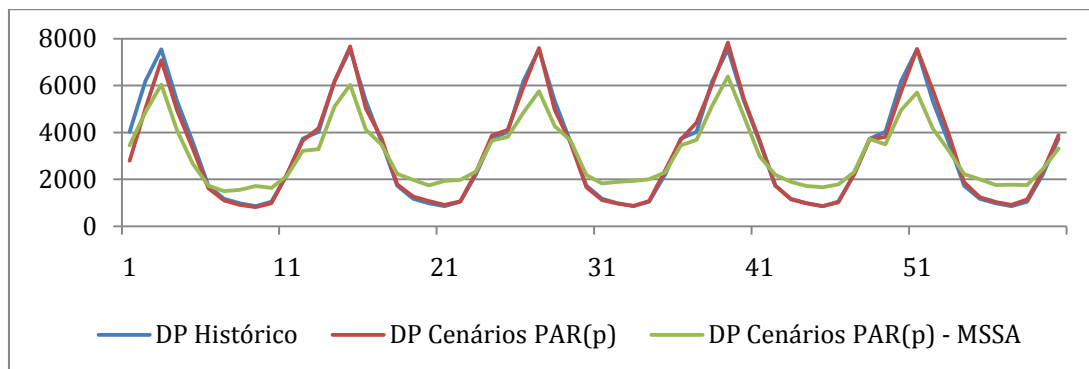


Figura 5.25: Comparação entre os desvios-padrão dos cenários e o desvio-padrão histórico do subsistema Nordeste.

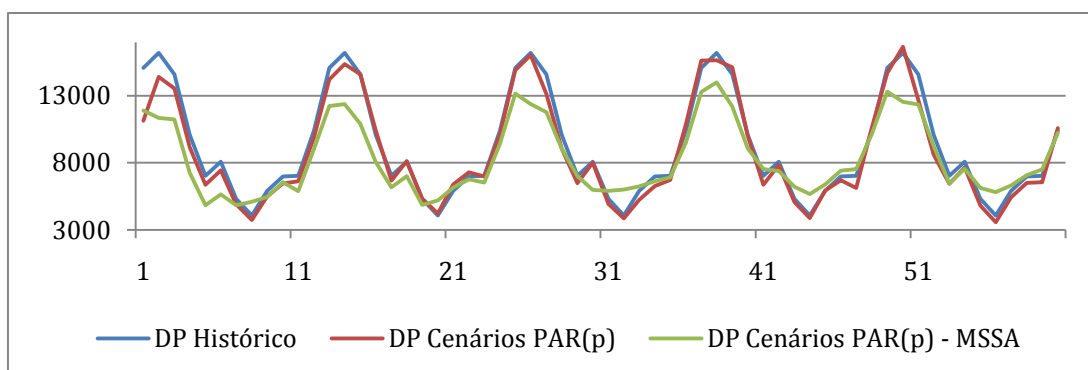


Figura 5.26: Comparação entre os desvios-padrão dos cenários e o desvio-padrão histórico do subsistema Sudeste/Centro-Oeste.

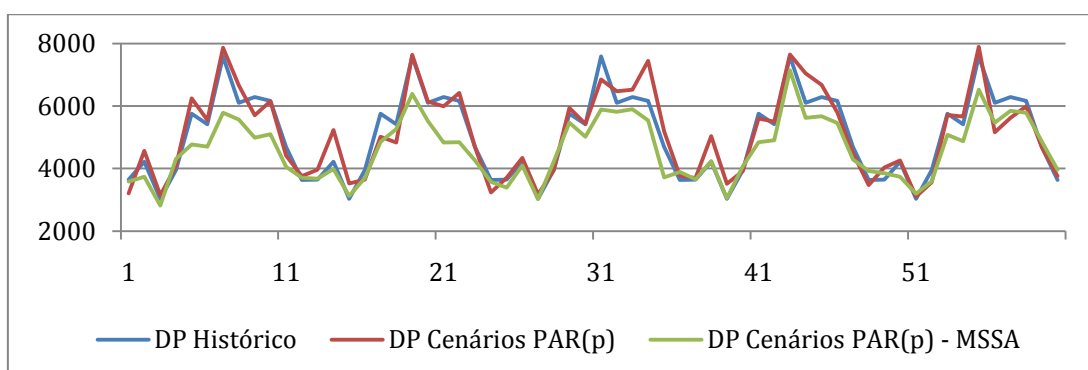


Figura 5.27: Comparação entre os desvios-padrão dos cenários e o desvio-padrão histórico do subsistema Sul.

As comparações entre as médias e os desvios-padrão nos quatro subsistemas são feitas por análise gráfica. Nesta análise busca-se a igualdade entre as médias e os desvios-padrão da série histórica e dos cenários obtidos com as modelagens PAR(p) sem e com a filtragem MSSA. Embora os gráficos apresentados nas figuras 5.20 – 5.27 mostrem que não há a igualdade, há uma boa similaridade entre a série histórica e as séries sintéticas através do modelo PAR(p) sem ou com a filtragem MSSA, com isso, pode-se dizer que os cenários reproduzem bem o histórico no primeiro caso e, no segundo caso, os cenários têm uma amplitude menor com a filtragem MSSA que os cenários gerados sem esta filtragem. Esta característica não é interessante na geração de cenários porque eles podem não contemplar todos os possíveis casos e esta situação está apresentada na análise do desempenho do modelo. No entanto, a média é conservada.

5.3. Análise de Desempenho do Modelo

A análise do desempenho do modelo é feita através de testes como definidos na seção 2.5. Testes de Média, Variância, Aderência e análises de sequências são feitos de acordo com os critérios estabelecidos naquela seção. O objetivo destes testes é o de verificar se as médias e as variâncias dos cenários gerados são estatisticamente iguais aos mesmos parâmetros medidos historicamente. O teste *t*-Student foi utilizado para testar a igualdade entre as médias e o teste de Levine para testar a igualdade entre as variâncias da série histórica original e o dos cenários gerados nas duas abordagens PAR(p) sem e com a filtragem MSSA. O nível de significância para os casos é de 5%. A tabela 5.8 apresenta os percentuais de não rejeição da hipótese nula de igualdade entre as médias e a tabela 5.9 apresenta os percentuais de não rejeição de hipótese nula de igualdade entre as variâncias dentre os 60 meses testados nos 5.000 cenários gerados a partir da modelagem PAR(p) e da modelagem PAR(p) – MSSA em comparação com as mesmas medidas nos dados históricos. Estes resultados indicam que as médias e variâncias dos cenários gerados a partir do modelo PAR(p) são estatisticamente iguais aos mesmos parâmetros medidos na série histórica na maior parte dos períodos. No entanto, embora reproduza a série histórica em boa parte dos períodos medidos, as séries sintéticas geradas a partir da modelagem PAR(p) – MSSA superam os percentuais de não rejeição da hipótese nula de igualdade apenas no que se refere à média do subsistema Nordeste e apresenta índices baixos nos subsistemas norte e Sul.

Tabela 5.8.: Percentuais de não rejeição da hipótese nula de igualdade entre as médias no teste *t*-Student em comparação com a série histórica.

Subsistema	Não rejeita H_0 (%)	
	PAR(p)	PAR(p) - MSSA
Norte	100	70
Nordeste	99	100
Sudeste/Centro-Oeste	100	99
Sul	97	65

Tabela 5.9.: Percentuais de não rejeição da hipótese nula de igualdade entre as variâncias no teste de Levene em comparação com a série histórica.

Subsistema	Não rejeita H_0 (%)	
	PAR(p)	PAR(p) - MSSA
Norte	99	79
Nordeste	97	72
Sudeste/Centro-Oeste	100	89
Sul	99	60

De acordo com os resultados mostrados nas tabelas 5.8 referente ao teste da média, os cenários gerados através do modelo $PAR(p)$ obtiveram melhores resultados em relação a reprodução da média histórica, visto que o teste não rejeita a hipótese nula de igualdade das médias em pelo menos 97% dos 60 meses de cenários enquanto no modelo $PAR(p) - MSSA$ este índice é de pelo menos 65%.

No que diz respeito à reprodução da variância da série histórica, os cenários gerados pela modelagem $PAR(p)$ obtiveram a taxa de não rejeição da hipótese nula de pelo menos 97% dos cenários e na modelagem $PAR(p) - MSSA$ este índice foi de pelo menos 60% e de no máximo 89% como mostrado na tabela 5.9. Este percentual de 60% de não rejeição da hipótese nula de igualdade das variâncias acontece no subsistema Sul, onde após a filtragem MSSA, a série suavizada apresenta um comportamento mais estacionário que a série original. Este fato pode ser observado na figura 5.15, que mostra a série do subsistema Sul original sobreposta por sua equivalente filtrada.

Em relação a este subsistema Sul, é possível verificar na figura 5.4 que ele tem um comportamento distinto dos demais. Este comportamento é confirmado na tabela 5.1, na assimetria e curtose. Como o MSSA trata os quatro subsistemas de forma igual, a filtragem do subsistema Sul pode ter excluído uma quantidade inadequada de ruído, conforme observado nas figuras 5.12 – 5.15, comprometendo os resultados.

Para avaliar a capacidade de reprodução de períodos críticos que os cenários podem ter, os testes de sequências são aplicados conforme seção 2.5.4. A ideia é observar se os cenários reproduzem as distribuições de probabilidade das variáveis aleatórias: comprimento, soma e intensidade da sequência referente a sua respectiva distribuição na série histórica. Para isso, são aplicados os testes de Kolmogorov-Smirnov para a soma e para a intensidade e Qui-Quadrado para o comprimento da sequência. As tabelas 5.10 e 5.11 apresentam os resultados das análises de sequências negativas para os cenários obtidos a partir das modelagens $PAR(p)$ sem e com a filtragem MSSA.

Tabela 5.10.: Teste das Sequencias Negativas para o modelo PAR(p)

Subsistema	Comprimento	Soma	Intensidade
	Valor crítico: 3,84	p -valor min:0,05	p -valor min:0,05
N	2,33	0,99812	0,25267
NE	0,86	0,80479	0,65092
SE	0,18	0,83240	0,77302
S	0,02	0,96319	0,03753

Tabela 5.11.: Teste das Sequencias Negativas para o modelo PAR(p) - MSSA

Subsistema	Comprimento	Soma	Intensidade
	Valor crítico: 3,84	p -valor min:0,05	p -valor min:0,05
N	0,60	0,27713	0,46908
NE	2,97	0,00055	0,00024
SE	1,58	0,14515	0,33490
S	6,35	0,07471	0,00277

Os testes apresentados nas tabelas 5.10 e 5.11 mostram que os cenários gerados pela modelagem PAR(p) não foram capazes de reproduzir a variável intensidade do subsistema Sul. Por sua vez, os cenários gerados através da modelagem PAR(p) com os dados filtrados via MSSA tem capacidade de reproduzir os períodos críticos de seca nos subsistemas Norte e Sudeste mas não são capazes de reproduzir as variáveis soma e intensidade no subsistema Nordeste e as variáveis comprimento e intensidade no subsistema Sul. Os resultados mostram um mau desempenho da modelagem proposta no que diz respeito à geração de cenários sintéticos. Este fato leva a uma nova abordagem apresentada na seção 5.4.

5.4. Abordagem PAR(p) – MSSA/SSA com filtragem do subsistema Sul separadamente

Como foi apresentado na figura 5.4 e confirmado na estatística de curtose na tabela 5.1 e na tabela 5.2 de correlação entre os subsistemas, o subsistema Sul apresenta comportamento diferenciado dos demais. A abordagem MSSA executa de forma igual todas as séries temporais envolvidas na filtragem. Como o procedimento desta abordagem foi feito utilizando os quatro subsistemas simultaneamente, é provável que os resultados tenham sido prejudicados por conta desta diferença de comportamento. Sendo assim, uma nova abordagem é apresentada tratando o

subsistema Sul de forma diferenciada dos demais subsistemas. Na ocasião, os subsistemas Norte, Nordeste e Sudeste/Centro-Oeste são filtrados através da abordagem MSSA usando o valor da janela de defasagem $L = 492$ e o subsistema Sul é filtrado via SSA separadamente usando o mesmo comprimento de janela.

Após a análise gráfica dos vetores singulares na filtragem MSSA com os três subsistemas N, NE e SE, as componentes que representam a série filtrada e a série de ruídos são obtidas. Os vetores que compõem estas componentes são apresentados na tabela 5.12. Pode-se perceber que, diferentemente da filtragem considerando os quatro subsistemas, a componente ruidosa apresenta uma quantidade substancialmente menor de vetores singulares, o que indica que na filtragem anterior, o sinal obtido carregava componentes ruidosas para a série filtrada. A partir da série de ruídos MSSA, a séries de ruídos de cada subsistema é obtida e testada via BDS. Os resultados deste teste são apresentados na tabela 5.13.

Tabela 5.12.: Componentes geradas na abordagem MSSA para os subsistemas Norte, Nordeste e Sudeste/Centro-Oeste.

Vetores em cada componente	
Série Filtrada	1-130, 133-140, 143-151, 154, 201, 202
Ruído	131, 132, 141, 142, 152, 153, 155-200, 203-492

Tabela 5.13.: Teste BDS para as séries de ruídos dos subsistemas Norte, Nordeste e Sudeste/Centro-Oeste na filtragem MSSA.

	Dim	Estatística BDS	Estatística Z	Erro padrão	p-valor
N	2	-4.04E-06	-0.346957	1.16E-05	0.7286
	3	-1.63E-05	-0.624830	2.60E-05	0.5321
	4	-2.42E-05	-0.557253	4.35E-05	0.5774
	5	-1.55E-05	-0.243479	6.36E-05	0.8076
	6	-2.72E-05	-0.316557	8.60E-05	0.7516
NE	2	1.28E-07	0.006924	1.84E-05	0.9945
	3	4.53E-06	0.110231	4.11E-05	0.9122
	4	7.67E-07	0.053964	6.87E-05	0.9911
	5	5.42E-06	0.053964	0.000100	0.9570
	6	-2.26E-06	-0.016622	0.000136	0.9867
SE	2	-9.89E-09	-1.594679	6.20E-09	0.1108
	3	-9.90E-13	-0.063449	1.56E-11	0.9494
	4	-9.94E-17	-0.003383	2.94E-14	0.9973
	5	-9.03E-21	-0.000186	4.85E-17	0.9999
	6	-8.95E-25	-1.21E-05	7.39E-20	1.0000

Após a determinação do ruído em cada subsistema, as séries filtradas são obtidas. As figuras 5.28 - 5.30 apresentam as séries originais de cada subsistema sobrepostas pelas respectivas séries filtradas via MSSA¹.

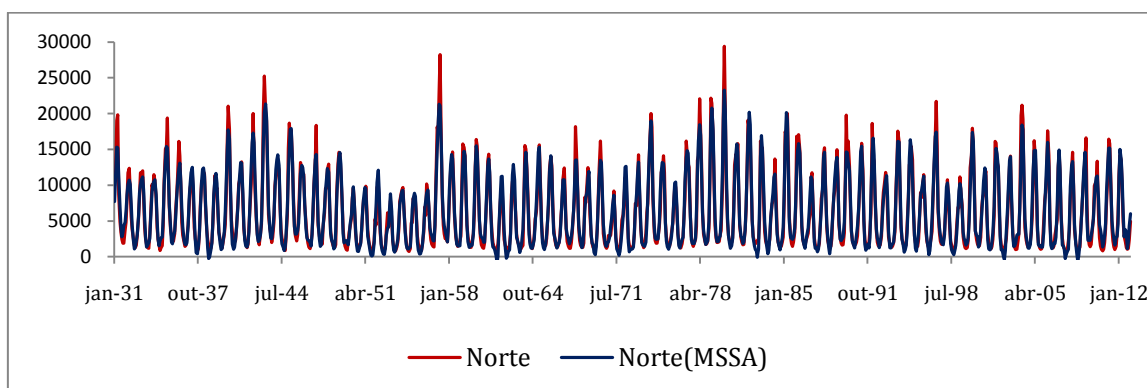


Figura 5.28: Comparação gráfica entre a série original e a série filtrada através da nova abordagem MSSA no subsistema Norte.

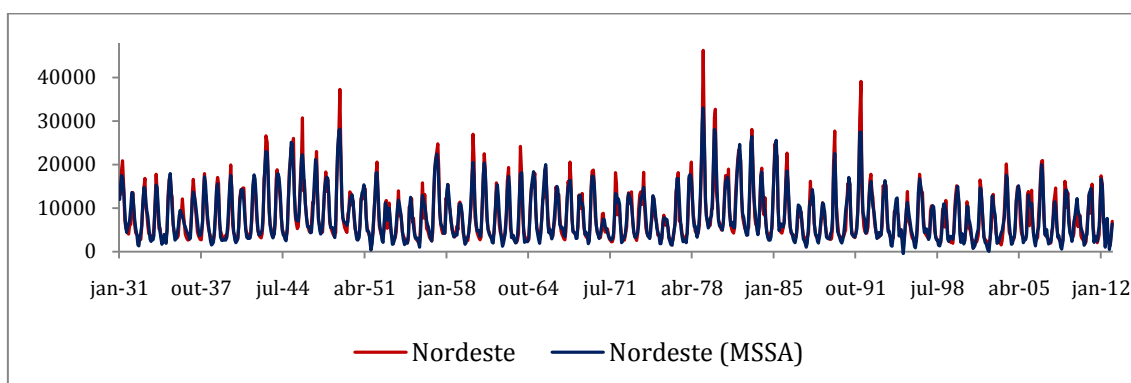


Figura 5.29: Comparação gráfica entre a série original e a série filtrada através da nova abordagem MSSA no subsistema Nordeste.

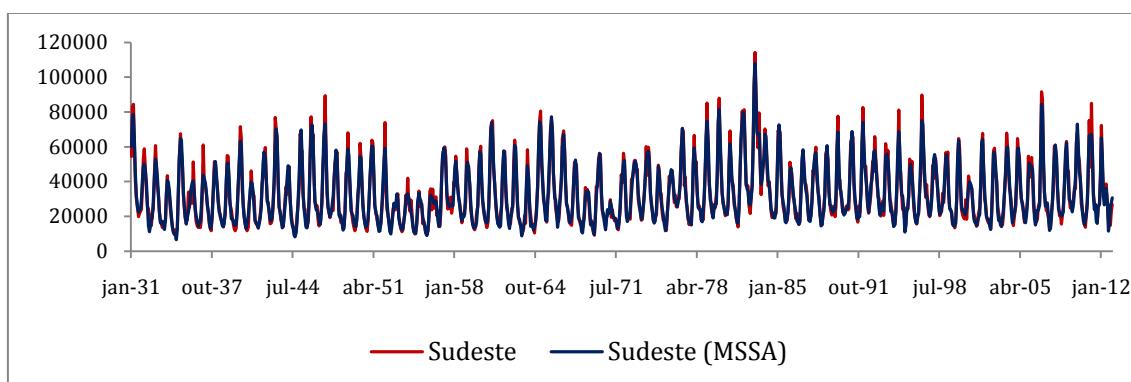


Figura 5.30: Comparação gráfica entre a série original e a série filtrada através da nova abordagem MSSA no subsistema Sudeste/Centro-Oeste.

¹ Nesta nova abordagem MSSA foram considerados apenas os subsistemas Norte, Nordeste e Sudeste/Centro-Oeste.

A série temporal dos dados de ENA do subsistema Sul foi filtrada separadamente via SSA. A análise gráfica também foi feita e as componentes da série filtrada e da série de ruídos são obtidas e apresentadas na tabela 5.14. Assim como foi observado nos outros subsistemas, o número de vetores que compõem a componente ruidosa no subsistema Sul é muito inferior ao número de vetores desta componente observados na tabela 5.3. Esta componente ruidosa é testada via BDS e os resultados deste teste são apresentados na tabela 5.15.

Tabela 5.14.: Componentes geradas na abordagem SSA para o subsistema Sul.

Vetores em cada componente	
Série Filtrada	1-103, 135, 136
Ruído	104-134, 137-492

Tabela 5.15.: Teste BDS para a série de ruído do subsistema Sul na filtragem SSA.

Dim	Estatística BDS	Estatística Z	Erro padrão	p-valor
2	9.20E-07	0.051323	1.79E-05	0.9591
3	2.76E-06	0.068889	4.01E-05	0.9451
4	3.44E-06	0.051328	6.70E-05	0.9591
5	5.03E-06	0.051302	9.80E-05	0.9591
6	-4.94E-06	-0.037204	0.000133	0.9703

A figura 5.31 apresenta a série original do subsistema Sul sobreposta pela série filtrada obtida na abordagem SSA. Pode-se verificar que nesta nova abordagem a série filtrada se aproxima mais da série original em comparação com a série filtrada via MSSA mostrado na figura 5.15.

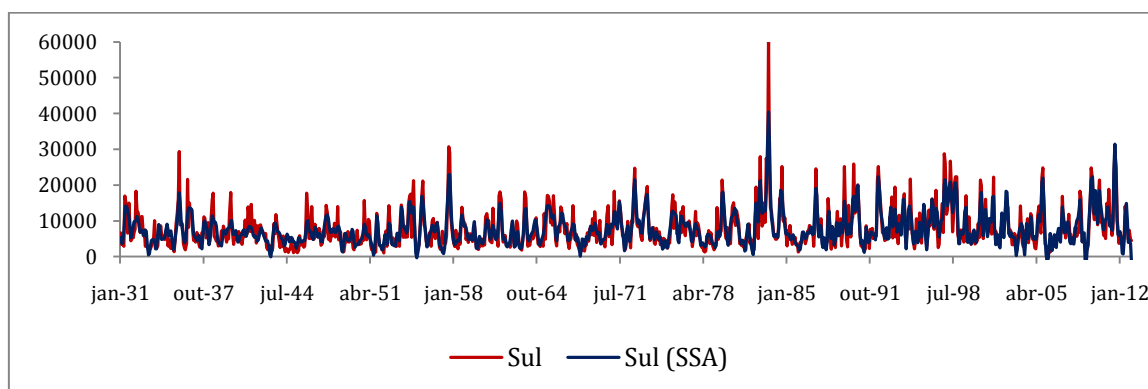


Figura 5.31: Comparação gráfica entre a série original e a série filtrada através da nova abordagem SSA no subsistema Sul.

Com as séries dos três subsistemas N, NE e SE filtradas via MSSA e a série do subsistema S filtrada separadamente via SSA, a modelagem PAR(p) sem e com a filtragem através da abordagem SSA é feita e o MAPE *in sample* mês a mês é calculado para todos os subsistemas. A tabela 5.16 apresenta os resultados do MAPE para todas as situações estudadas.

Tabela 5.16.: MAPE (%) *in sample* dos meses para os modelos PAR(p), PAR(p)-SSA para o subsistema Sul e PAR(p) - MSSA para os demais subsistemas.

MAPE (%)	Norte		Nordeste		Sudeste/Centro-Oeste		Sul	
Mês	PAR(p)	PAR(p) MSSA	PAR(p)	PAR(p) MSSA	PAR(p)	PAR(p) MSSA	PAR(p)	PAR(p) SSA
Jan	21,3150	10,064	20,087	7,690	19,099	6,548	43,863	17,127
Fev	22,4870	6,3033	29,174	7,824	20,425	6,283	42,421	15,813
Mar	14,0090	4,6202	24,617	8,199	17,947	5,092	31,303	6,220
Abr	14,1460	4,325	29,199	9,087	12,334	5,505	45,127	43,878
Mai	12,9600	5,505	14,683	13,094	10,379	7,064	79,801	20,065
Jun	10,5110	8,765	7,271	14,224	10,441	6,681	55,224	14,441
Jul	6,8260	15,189	5,185	13,979	8,007	6,786	46,304	21,558
Ago	5,7416	10,348	4,522	20,281	9,241	9,832	67,320	28,783
Set	8,3045	44,657	6,771	21,945	14,192	11,075	51,134	15,748
Out	13,2920	37,865	13,626	63,768	18,253	8,955	48,867	15,240
Nov	19,9250	1,2914	24,585	15,806	13,677	8,562	12,127	17,195
Dez	22,5470	10,912	25,193	9,066	15,421	8,211	46,571	18,719

Assim como na tabela 5.6, os resultados mostrados na tabela 5.16 também apresentam um melhor desempenho dos modelos com as séries filtradas através da abordagem Singular Spectrum Analysis para a maioria dos meses nos quatro subsistemas. Com as séries modeladas via PAR(p), são gerados os cenários de ENA para os quatro subsistemas.

Para as análises do desempenho dos cenários gerados a partir da modelagem PAR(p), são comparadas as médias e os desvios-padrão dos dados históricos com as médias e os desvios-padrão dos cenários obtidos através da modelagem PAR(p) para as séries de cada um dos quatro subsistemas sem e com a filtragem. As figuras 5.32 –

5.35 apresentam as comparações entre as médias históricas e as médias dos cenários nas duas abordagens sem e com a filtragem.

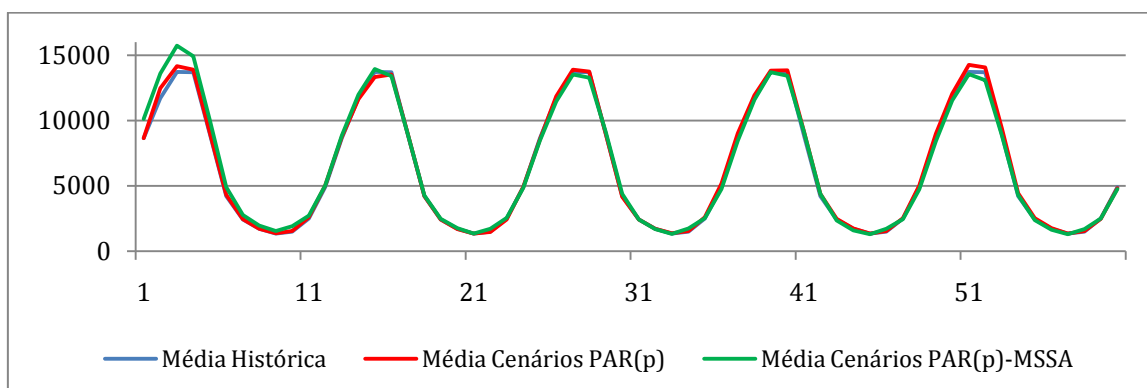


Figura 5.32: Comparação entre as médias dos cenários e a média histórica do subsistema Norte com a nova abordagem MSSA.

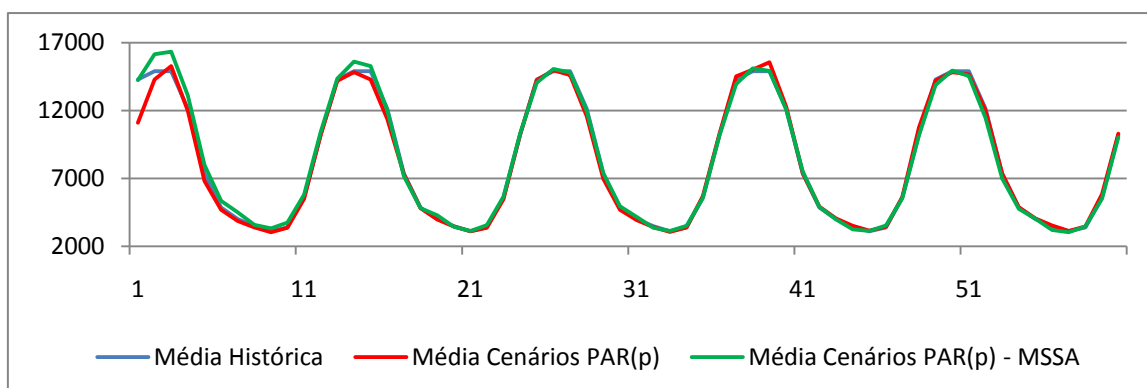


Figura 5.33: Comparação entre as médias dos cenários e a média histórica do subsistema Nordeste com a nova abordagem MSSA.

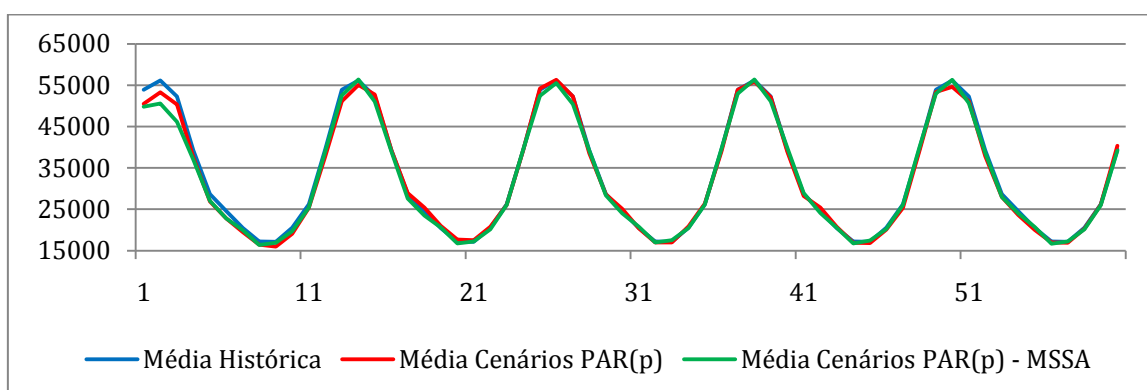


Figura 5.34: Comparação entre as médias dos cenários e a média histórica do subsistema Sudeste/Centro-Oeste com a nova abordagem MSSA.

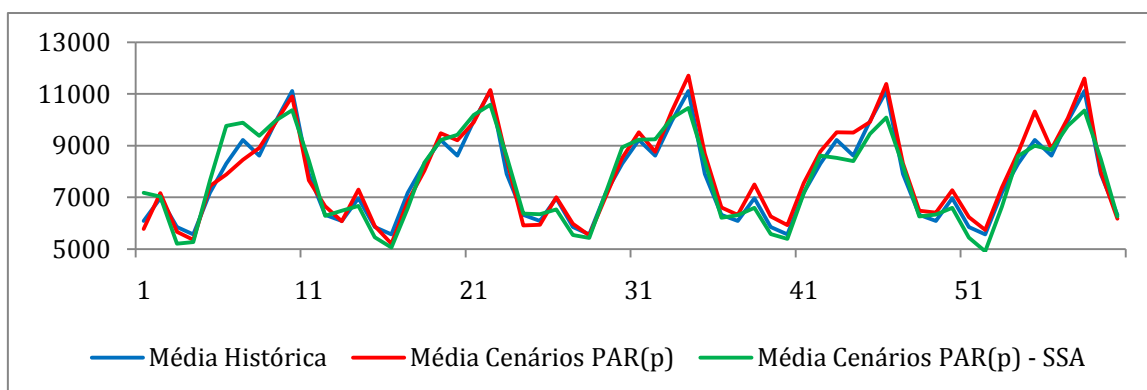


Figura 5.35: Comparação entre as médias dos cenários e a média histórica do subsistema Sul com a abordagem SSA.

Nesta nova abordagem, as médias dos cenários gerados a partir das séries filtradas via MSSA e SSA estão bem mais próximas das médias históricas quando comparadas com as mesmas medidas nas figuras 5.20 – 5.23.

Assim como foi comentado anteriormente na seção 5.2, os desvios-padrão dos cenários gerados através da abordagem PAR(p) – MSSA/SSA² são menores que os desvios sem a filtragem. Então, os ruídos anteriormente removidos nesta filtragem são adicionados aos cenários de modo a obter novos valores de desvio-padrão. A forma com a qual estes ruídos são inseridos nos cenários foi discutida na seção 5.2 e é usada neste mesmo procedimento com a nova abordagem MSSA/SSA. As figuras 5.36 – 5.39 apresentam a comparação entre os desvios-padrão históricos e os cenários gerados a partir da modelagem PAR(p) sem e com a filtragem partindo desta nova abordagem. Pode-se perceber que os desvios-padrão dos cenários gerados a partir da modelagem com a filtragem MSSA/SSA se aproximam mais dos desvios-padrão das séries históricas.

² MSSA/SSA considera os subsistemas Norte, Nordeste e Sudeste/Centro-Oeste filtrados via MSSA e o subsistema Sul filtrado via SSA.

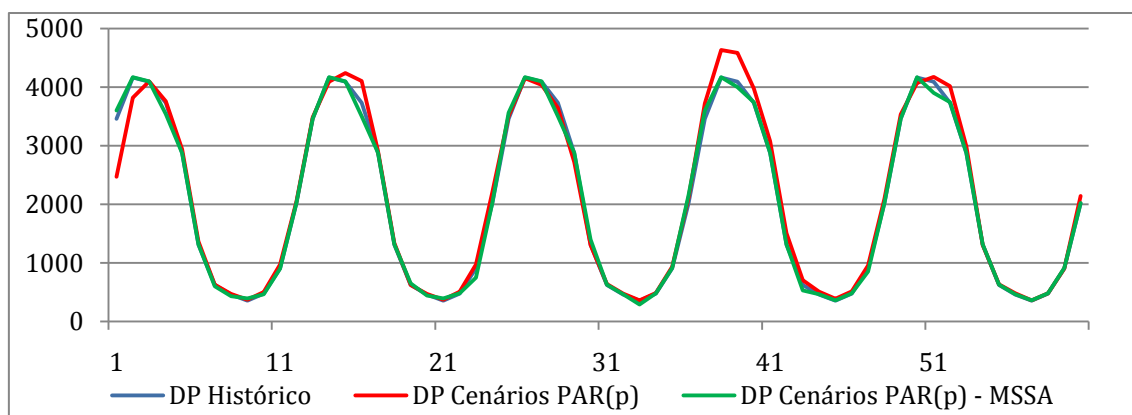


Figura 5.36: Comparação entre os desvios-padrão dos cenários e o desvio-padrão histórico do subsistema Norte sob a nova abordagem MSSA e a inclusão dos ruídos.

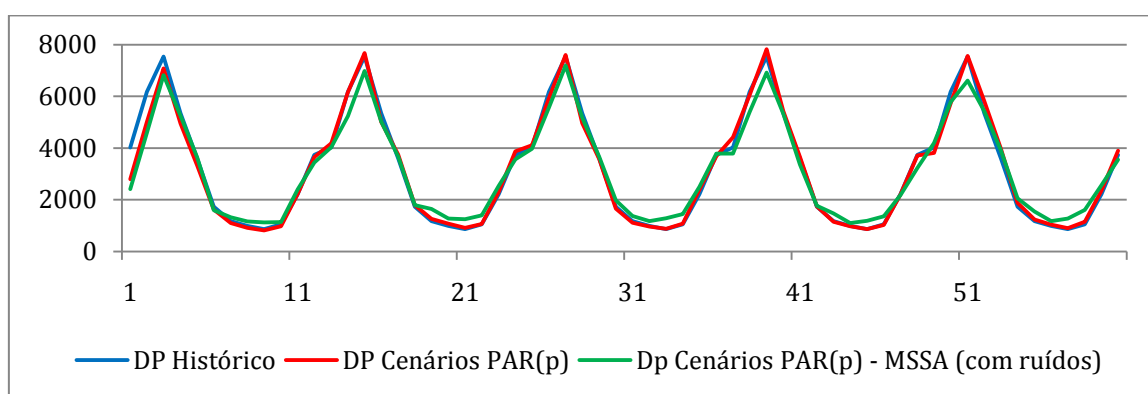


Figura 5.37: Comparação entre os desvios-padrão dos cenários e o desvio-padrão histórico do subsistema Nordeste sob a nova abordagem MSSA e a inclusão dos ruídos.

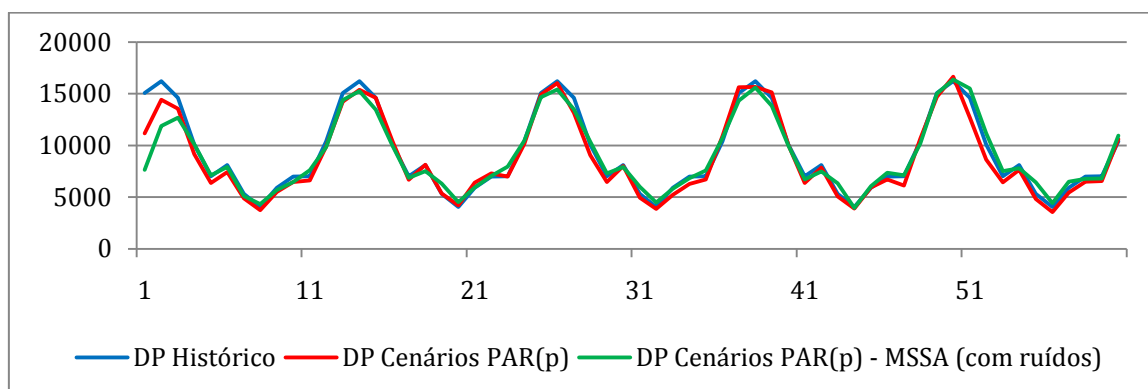


Figura 5.38: Comparação entre os desvios-padrão dos cenários e o desvio-padrão histórico do subsistema Sudeste/Centro-Oeste sob a nova abordagem MSSA e a inclusão dos ruídos.

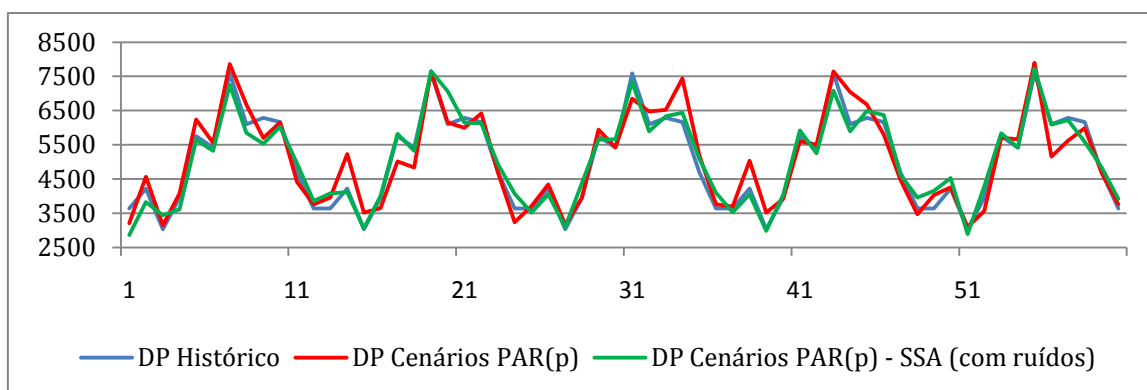


Figura 5.39: Comparação entre os desvios-padrão dos cenários e o desvio-padrão histórico do subsistema Sul sob a abordagem SSA e a inclusão dos ruídos.

A partir destas análises gráficas das médias e desvios-padrão dos cenários gerados na modelagem PAR(p) das séries de ENA filtradas através do processo MSSA/SSA pode-se concluir que os cenários desta forma gerados reproduzem a série histórica uma vez que se busca a similaridade nestes gráficos. Para a análise do desempenho do modelo, os testes aplicados sobre as médias e variâncias são os mesmos apresentados na seção 5.3 com as suas respectivas características. As tabelas 5.17 e 5.18 apresentam os resultados do teste *t*-Student de igualdade entre as médias e os resultados do teste de Levene de igualdade entre as variâncias.

Tabela 5.17.: Percentuais de não rejeição da hipótese nula de igualdade entre as médias no teste *t*-Student em comparação com a série histórica com a nova abordagem.

Subsistema	Não rejeita H_0 (%)	
	PAR(p)	PAR(p) – MSSA/SSA
Norte	100	85
Nordeste	99	100
Sudeste/Centro-Oeste	100	94
Sul	97	99

Tabela 5.18.: Percentuais de não rejeição da hipótese nula de igualdade entre as variâncias no teste de Levene em comparação com a série histórica com a nova abordagem.

Subsistema	Não rejeita H_0 (%)	
	PAR(p)	PAR(p) – MSSA/SSA
Norte	100	97
Nordeste	99	97
Sudeste/Centro-Oeste	100	97
Sul	97	97

Os resultados apresentados nas tabelas 5.17 e 5.18 mostram uma melhora no desempenho do modelo $PAR(p)$ com a filtragem MSSA/SSA, sobretudo no subsistema Sul que obteve desempenho abaixo de 70% quando os quatro subsistemas haviam sido filtrados simultaneamente na filtragem MSSA.

Na avaliação da capacidade de os cenários reproduzirem os períodos críticos de seca, o teste das sequências negativas é aplicado para os cenários gerados a partir do modelo $PAR(p)$ – MSSA/SSA. A tabela 5.19 apresenta os resultados deste teste. Os parâmetros e critérios são os mesmos utilizados anteriormente e apresentados nas tabelas 5.10 e 5.11. Nesta nova abordagem, os cenários gerados não são capazes de reproduzir os períodos críticos apenas na variável Soma do subsistema NE e o desempenho no subsistema Sul é melhor que o mesmo avaliado pelos cenários gerados a partir do modelo $PAR(p)$ sem filtragem, conforme apresentado na tabela 5.10, onde os mesmos não tem a capacidade de reproduzir a variável Intensidade.

Tabela 5.19.: Teste das Sequências Negativas para o modelo $PAR(p)$ – MSSA/SSA

Subsistema	Comprimento Valor crítico:3,84	Soma p -valor min: 0,05	Intensidade p -valor min: 0,05
N	1,06	0,76024	0,70471
NE	0,16	0,01220	0,07203
SE	0,03	0,37662	0,23918
S	0,80	0,10420	0,41750