UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA

MATHEUS MINGATOS FERNANDES GEMIGNANI

Cenários sintéticos de radiação
solar para estudos energéticos

São Paulo
2018

MATHEUS MINGATOS FERNANDES GEMIGNANI

**Cenários sintéticos de radiação
solar para estudos energéticos**

Versão Corrigida

(Versão original encontra-se na unidade que
aloja o Programa de Pós-graduação)

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para
obtenção do título de Doutor em
Ciências

Área de Concentração:
Sistemas de Potência

Orientador:
Prof. Dr. Nelson Kagan

São Paulo
2018

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e alterado em relação à versão original, sob responsabilidade única do autor e com anuência de seu orientador.

São Paulo, 18 de agosto de 2018

Assinatura do autor:

Assinatura do orientador:

Ficha catalográfica

Gemignani, Matheus Mingatos Fernandes

Cenários sintéticos de radiação solar para estudos energéticos / M. M. F. Gemignani – versão corr. – São Paulo, 2018.

158 p.

Tese (Doutorado) – Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Energia e Automação Elétricas.

1.Radiação solar 2.Cenários sintéticos 3.Modelo estocástico 4. Leilão de energia I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Energia e Automação Elétricas II.t

Dedico este trabalho a todos que, direta e indiretamente, me apoiaram nos períodos mais difíceis.

AGRADECIMENTOS

Ao Prof. Nelson Kagan, pela orientação, compreensão e importantes contribuições.

Ao Prof. Luiz Lebensztajn, pelo apoio e costumeira boa vontade.

Aos Professores Mauricio Salles e André Gimenes, pelas valiosas contribuições no exame de qualificação.

Ao Prof. Marcos Gouvêa, pelas discussões para elaboração deste trabalho e em outros temas de interesse comum, pelo incentivo e pela amizade.

Ao ENERQ – Centro de Estudos em Regulação e Qualidade de Energia, da Escola Politécnica da USP, por propiciar a realização deste trabalho.

À Companhia Energética de Petrolina, em especial ao Sr. Hélio Takeno, pela disponibilização de dados e pelo suporte na realização da pesquisa.

Aos colegas Guido Rostegui e Luís Lourenço, pelas discussões, pela colaboração na pilotagem do MatLab e pelas significativas contribuições na pesquisa.

Ao meu irmão Thomaz, pelas aulas de matemática e estatística.

Aos meus pais, Silvio e Suzana, e à Beatriz, pelo apoio e suporte nos períodos nebulosos e pela compreensão dos momentos de ausência.

Aos colegas do Enerq, pela ótima convivência neste período.

Aos meus amigos, em especial Daniel Brooke, Daniel Carillo, Daniel Yoshio, Ewerton Guarnier, Gabriel Vergara, Ivo Cyrillo, Manoel Medeiros, Marcelo Lopes, Mário Harada, Thiago Daud e Wagner Silva, pelas diversas divagações dos últimos anos.

Finalmente, a todos aqueles que de alguma forma contribuíram para o êxito deste trabalho e que a memória não me permitiu listar.

“Pain is temporary. It may last a minute, or an hour, or a day, or a year, but eventually it will subside and something else will take its place. If I quit, however, it lasts forever.”

(Lance E. Armstrong)

RESUMO

GEMIGNANI, M. M. F. **Cenários sintéticos de radiação solar para estudos energéticos.** 2018. 158 f. Tese (Doutorado) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2018.

Esta tese apresenta os resultados de pesquisa sobre geração de séries sintéticas de radiação solar para estudos energéticos, realizada através do uso de modelos estocásticos e com o propósito de desenvolver método para aplicações práticas no setor elétrico. Para tanto, inicialmente foi levantado o estado da arte do tema, com revisão da literatura de séries temporais e de processos estocásticos, suas particularidades e potencialidades, complementado pela contextualização do uso de cenários no setor elétrico nacional, especialmente na operação e planejamento do sistema hidrotérmico, e por experiências internacionais na modelagem do recurso solar. A modelagem das séries utilizou dados reais de localidades do nordeste brasileiro e foi desenvolvida através do método de Box-Jenkins, realizando-se estudos de alternativas para cada uma de suas etapas. O pré-tratamento dos dados foi avaliado por três estratégias de remoção da tendência das séries e na estimativa dos coeficientes dos modelos foram comparados os métodos de Yule-Walker e dos mínimos quadrados. As análises consideraram quatro opções de modelos autorregressivos e os períodos horário, diário e mensal. O modelo autorregressivo convencional com intervalo mensal, identificado como o mais adequado para aplicação em estudos energéticos, e sua variação periódica foram implementados e avaliados com maior profundidade. Este estudo complementar considerou diferentes ordens de atraso e realizou comparações dos resultados por três métodos de cálculo do erro. O modelo desenvolvido com estrutura autorregressiva periódica de primeira ordem apresentou resultados satisfatórios e significativamente superiores aos dos demais modelos. Por fim, este modelo foi empregado na geração de séries sintéticas, criando 1.000 cenários de radiação solar mensal, posteriormente aplicados em modelo de contrato de venda de energia para avaliação de estratégias de participação em leilões, em análise de riscos de suprimento e em estimativa probabilística da receita esperada por parques geradores.

Palavras-chave: radiação solar, cenários sintéticos, modelos estocásticos, leilões de energia.

ABSTRACT

GEMIGNANI, M. M. F. **Solar radiation synthetic sequences for energy studies**. 2018. 158 p. Thesis (Doctorate degree) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2018.

This thesis presents the results of a research on the generation of solar radiation synthetic sequences for energy studies, carried out through the use of stochastic models and with the purpose of developing a method for practical applications in the electric sector. In order to do so, the state of the art was devised through a review of the literature of time series and stochastic processes, their particularities and potentialities, complemented by the contextualization of the use of scenarios in the national electricity sector, especially in the hydrothermal system operation and planning, and international experiences in modeling the solar resource. The series modeling used real data from localities in the Brazilian Northeast and was developed through the Box-Jenkins method, carrying out alternative studies for each of its stages. The data pre-treatment has been evaluated by three strategies for the series trend removal and by the methods of least squares and of Yule-Walker for the estimation of the model coefficients. The analysis considered four options of autoregressive models and hourly, daily and monthly periods. The conventional autoregressive model with monthly interval, identified as the most suitable for applications in energy studies, and its periodic variation were implemented and evaluated in greater depth. This complementary study considered different orders of delay and made comparisons of the results for three error calculation methods. The model developed with periodic autoregressive structure of first order presented results that are satisfactory and significantly superior than the other models. Finally, this model was used in the generation of synthetic series, creating 1,000 scenarios of monthly solar radiation, to be later applied in a model of power purchase agreement to evaluate strategies for auctions bidding, analysis of supply risks and probabilistic estimation of the expected revenue of power plants.

Keywords: solar radiation; synthetic sequences; stochastic models; energy auctions.

LISTA DE FIGURAS

Figura 1. Capacidade instalada da geração (MW).....	23
Figura 2. Modelo do gerador hidrelétrico.....	24
Figura 3. Rendimento da turbina em Foz do Areia	25
Figura 4. Histórico de vazões no Rio Madeira em Porto Velho.....	27
Figura 5. Diagrama de análise de séries temporais	29
Figura 6. Autocorrelação e distribuição da velocidade do vento	33
Figura 7. Radiação diária média em localidade de NE.....	35
Figura 8. Irradiância no mês de janeiro	35
Figura 9. Medição da corrente em um resistor ideal	36
Figura 10. Histórico da temperatura diária em maio de 2016, em São Paulo	37
Figura 11. Cotação do dólar, 2015	39
Figura 12. Diferenciação da cotação do dólar	40
Figura 13. Ruído branco, simulação de média nula e variância unitária.....	41
Figura 14. Processo AR de ordem 2	42
Figura 15. Processo MA de ordem 2	43
Figura 16. Processo ARMA de ordem (2,2).....	44
Figura 17. Fluxograma representativo do método de Box-Jenkins	46
Figura 18. Análise preliminar de média, valor máximo, média dos mínimos e desvio padrão para cada mês.....	53
Figura 19. Séries de energia diária incidente para o mês de julho.	54
Figura 20. Séries de energia diária incidente para o mês de dezembro.....	54
Figura 21. Verificação da distribuição da energia total incidente para o mês de julho.....	55
Figura 22. Verificação da distribuição da energia total incidente para o mês de dezembro. ...	55
Figura 23. Tendência e remoção de tendência para o mês de julho	57
Figura 24. Tendência e remoção de tendência para o mês de dezembro.....	58
Figura 25. Retirada de tendência pelo máximo diário, mês de julho	59
Figura 26. Retirada de tendência pelo máximo diário, mês de dezembro.....	59
Figura 27. Etapa adicional para a retirada de média e tendência pelo máximo diário	60
Figura 28. Retirada de tendência pela média mensal	60
Figura 29. Métodos de geração de séries sintéticas da radiação diária (janeiro).....	65
Figura 30. Métodos de geração de séries sintéticas da radiação diária (fevereiro)	65

Figura 31. Métodos de geração de séries sintéticas da radiação diária (março).....	65
Figura 32. Métodos de geração de séries sintéticas da radiação diária (abril)	66
Figura 33. Métodos de geração de séries sintéticas da radiação diária (maio).....	66
Figura 34. Métodos de geração de séries sintéticas da radiação diária (junho).....	66
Figura 35. Métodos de geração de séries sintéticas da radiação diária (julho)	67
Figura 36. Métodos de geração de séries sintéticas da radiação diária (agosto)	67
Figura 37. Métodos de geração de séries sintéticas da radiação diária (setembro)	67
Figura 38. Métodos de geração de séries sintéticas da radiação diária (outubro)	68
Figura 39. Métodos de geração de séries sintéticas da radiação diária (novembro)	68
Figura 40. Métodos de geração de séries sintéticas da radiação diária (dezembro).....	68
Figura 41. Métodos de geração de séries sintéticas de estrutura AR (MMQ), para a energia total diária (janeiro)	71
Figura 42. Métodos de geração de séries sintéticas de estrutura ARMA (MMQ) para a energia total diária (janeiro)	72
Figura 43. Métodos de geração de séries sintéticas de estrutura ARI (MMQ) para a energia total diária (janeiro)	72
Figura 44. Métodos de geração de séries sintéticas de estrutura ARIMA (MMQ) para a energia total diária (janeiro)	72
Figura 45. Séries sintéticas da radiação mensal a partir da estrutura AR.....	74
Figura 46. Séries sintéticas da radiação mensal a partir da estrutura ARI	75
Figura 47. Séries sintéticas da radiação mensal a partir da estrutura ARMA	75
Figura 48. Séries sintéticas da radiação mensal a partir da estrutura ARIMA.....	75
Figura 49. Métodos para geração de séries sintéticas horária com estrutura AR.....	77
Figura 50. Histogramas normalizados, 2 mil séries de radiação diária, mês de julho.....	79
Figura 51. Histogramas normalizados, 2 mil séries de radiação diária, mês de dezembro	80
Figura 52. Histogramas normalizados, 2 mil séries sintéticas de radiação mensal	82
Figura 53. Histogramas normalizados, 2 mil séries de radiação horária, mês de julho.	84
Figura 54. Histogramas normalizados, 2 mil séries de radiação horária, mês de dezembro....	85
Figura 55. Resíduos do modelo AR de ordens 1, 2 e 3.	92
Figura 56. Radiação mensal observada e prevista, modelo AR(3).....	93
Figura 57. Desvio padrão dos resíduos mensais do modelo PAR(3)	95
Figura 58. Distribuição dos resíduos PAR(3) em janeiro e abril.....	95
Figura 59. Radiação mensal observada e prevista, modelo PAR(3)	96
Figura 60. Erros dos modelos (pu)	97

Figura 61. Contabilização da energia nos contratos de compra de energia (PPA).....	101
Figura 62. Exemplo de séries sintéticas.....	102
Figura 63. Distribuição das séries de energia incidente	102
Figura 64. Histograma das séries de energia incidente	103
Figura 65. Receita de energia para o primeiro ano.....	104
Figura 66. Receita de energia para o vigésimo ano.....	104
Figura 67. Evolução da redução de receita por ano.....	105
Figura 68. Evolução da redução de receita agregada para 20 anos	106
Figura 69. Redução de receita acumulada anualmente, por contrato	107
Figura 70. Redução de receita acumulada, relativa, por contrato.....	108
Figura 71. Distribuição de receita, por contrato	109
Figura 72. Distribuição de receita, relativa, por contrato	110
Figura 73. Distribuição da redução de receita, acumulada no contrato de 20 anos.....	110
Figura 74. Distribuição da redução de receita relativa, acumulada no contrato de 20 anos...	111
Figura 75. Histograma da receita relativa entre os contratos de 0,88 e 1,00 pu.....	112
Figura 76. Histograma da receita relativa entre os contratos de 0,88 e 0,87 pu.....	112
Figura 77. Tendências mensais.....	124
Figura 78. Dados de 2015 retirados de tendência.....	125
Figura 79. Tendências para todos os meses de acordo com o método 2	126
Figura 80. Retirada de tendência pela média mensal	127
Figura 81. Influência do tamanho do conjunto de dados na estimativa do modelo, JAN-MAR	129
Figura 82. Influência do tamanho do conjunto de dados na estimativa do modelo, ABR-JUN	130
Figura 83. Influência do tamanho do conjunto de dados na estimativa do modelo, JUL-SET	131
Figura 84. Influência do tamanho do conjunto de dados na estimativa do modelo, OUT-DEZ	132
Figura 85. Métodos de geração de séries horárias com MMQ, estrutura AR (janeiro).....	133
Figura 86. Métodos de geração de séries horárias com MMQ, estrutura AR (fevereiro)	134
Figura 87. Métodos de geração de séries horárias com MMQ, estrutura AR (março).....	134
Figura 88. Métodos de geração de séries horárias com MMQ, estrutura AR (abril)	134
Figura 89. Métodos de geração de séries horárias com MMQ, estrutura AR (maio).....	135
Figura 90. Métodos de geração de séries horárias com MMQ, estrutura AR (junho).....	135

Figura 91. Métodos de geração de séries horárias com MMQ, estrutura AR (julho)	135
Figura 92. Métodos de geração de séries horárias com MMQ, estrutura AR (agosto)	136
Figura 93. Métodos de geração de séries horárias com MMQ, estrutura AR (setembro)	136
Figura 94. Métodos de geração de séries horárias com MMQ, estrutura AR (outubro)	136
Figura 95. Métodos de geração de séries horárias com MMQ, estrutura AR (novembro)	137
Figura 96. Métodos de geração de séries horárias com MMQ, estrutura AR (dezembro)	137
Figura 97. Métodos de geração de séries horárias com MMQ, estrutura ARMA (janeiro) ...	137
Figura 98. Métodos de geração de séries horárias com MMQ, estrutura ARMA (fevereiro) 138	
Figura 99. Métodos de geração de séries horárias com MMQ, estrutura ARMA (março)	138
Figura 100. Métodos de geração de séries horárias com MMQ, estrutura ARMA (abril)	138
Figura 101. Métodos de geração de séries horárias com MMQ, estrutura ARMA (maio)	139
Figura 102. Métodos de geração de séries horárias com MMQ, estrutura ARMA (junho) ...	139
Figura 103. Métodos de geração de séries horárias com MMQ, estrutura ARMA (julho)	139
Figura 104. Métodos de geração de séries horárias com MMQ, estrutura ARMA (agosto) ..	140
Figura 105. Métodos de geração de séries horárias com MMQ, estrutura ARMA (set)	140
Figura 106. Métodos de geração de séries horárias com MMQ, estrutura ARMA (out)	140
Figura 107. Métodos de geração de séries horárias com MMQ, estrutura ARMA (nov)	141
Figura 108. Métodos de geração de séries horárias com MMQ, estrutura ARMA (dez)	141
Figura 109. Métodos de geração de séries horárias com MMQ, estrutura ARI (janeiro)	141
Figura 110. Métodos de geração de séries horárias com MMQ, estrutura ARI (fevereiro) ...	142
Figura 111. Métodos de geração de séries horárias com MMQ, estrutura ARI (março)	142
Figura 112. Métodos de geração de séries horárias com MMQ, estrutura ARI (abril)	142
Figura 113. Métodos de geração de séries horárias com MMQ, estrutura ARI (maio)	143
Figura 114. Métodos de geração de séries horárias com MMQ, estrutura ARI (junho)	143
Figura 115. Métodos de geração de séries horárias com MMQ, estrutura ARI (julho)	143
Figura 116. Métodos de geração de séries horárias com MMQ, estrutura ARI (agosto)	144
Figura 117. Métodos de geração de séries horárias com MMQ, estrutura ARI (set)	144
Figura 118. Métodos de geração de séries horárias com MMQ, estrutura ARI (out)	144
Figura 119. Métodos de geração de séries horárias com MMQ, estrutura ARI (nov)	145
Figura 120. Métodos de geração de séries horárias com MMQ, estrutura ARI (dez)	145
Figura 121. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (janeiro) 145	
Figura 122. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (fev)	146
Figura 123. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (março) ..	146
Figura 124. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (abril)	146

Figura 125. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (maio)...	147
Figura 126. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (junho)..	147
Figura 127. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (julho)...	147
Figura 128. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (agosto)	148
Figura 129. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (set)	148
Figura 130. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (out).....	148
Figura 131. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (nov).....	149
Figura 132. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (dez)	149
Figura 133. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura AR.....	151
Figura 134. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura AR.....	152
Figura 135. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARI.....	153
Figura 136. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARI.....	154
Figura 137. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARMA.....	155
Figura 138. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARMA.....	156
Figura 139. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARIMA	157
Figura 140. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARIMA	158

LISTA DE TABELAS

Tabela 1. Previsibilidade de geração	23
Tabela 2. Modelos AR, retirada de tendência pela média diária.....	63
Tabela 3. Modelos AR, retirada de tendência pela máxima diária.....	63
Tabela 4. Modelos AR, retirada de tendência pela média mensal.....	63
Tabela 5. Modelos para geração de séries sintéticas com estrutura AR.....	70
Tabela 6. Modelos para geração de séries sintéticas com estrutura ARMA	70
Tabela 7. Modelos para geração de séries sintéticas com estrutura ARI	70
Tabela 8. Modelos para geração de séries sintéticas com estrutura ARIMA	71
Tabela 9. Parâmetros dos modelos anuais	74
Tabela 10. Parâmetros dos modelos horários com estrutura AR e ARI.....	76
Tabela 11. Parâmetros dos modelos horários com estrutura ARMA e ARIMA	76
Tabela 12. Erro quadrático médio dos modelos diários para Petrolina	80
Tabela 13. Erro quadrático médio dos modelos diários para Pombal	81
Tabela 14. Erro quadrático médio dos modelos mensais	82
Tabela 15. Erro quadrático médio dos modelos horários para Petrolina.....	85
Tabela 16. Erro quadrático médio dos modelos horários para Pombal	86
Tabela 17. Série mensal de radiação, pré-tratada para remoção de tendência.	91
Tabela 18. Parâmetros do modelo PAR(3) mensal.....	94
Tabela 19. Erros médios dos modelos	97

LISTA DE SIGLAS, ABREVIATURAS E TERMOS

ANEEL	Agência Nacional de Energia Elétrica
AR	Processo autorregressivo – <i>Auto Regressive</i>
ARIMA	Processo autorregressivo integrado de média móvel – <i>Auto Regressive Integrated Moving Average</i>
ARMA	Processo autorregressivo de média móvel – <i>Auto Regressive Moving Average</i>
CEPEL	Centro de Pesquisas de Energia Elétrica, da Eletrobras
DECOMP	Modelo de Determinação da Coordenação da Operação a Médio Prazo
ELETROBRAS	Centrais Elétricas Brasileiras – Companhia pública controladora de diversas geradoras, transmissoras, distribuidoras e centro de pesquisa em energia elétrica
EPE	Empresa de Pesquisa Energética
FV	Fotovoltaico
HydroLab	Software de suporte para o planejamento e operação de sistemas hidrotérmicos de geração elétrica, desenvolvido pela Unicamp
Irradiação - I	Radiação solar incidente, integrada durante determinado intervalo de tempo (Wh/m^2) ¹
Irradiância - G	Radiação solar incidente, instantânea (W/m^2)
MA	Processo de média móvel – <i>Moving Average</i>
MME	Ministério de Minas e Energia
NEWAVE	Modelo de Planejamento da Operação de Sistemas Hidrotérmicos Interligados de Longo e Médio Prazo
ONS	Operador Nacional do Sistema Elétrico
RNA	Redes Neurais Artificiais
SDDP	Software de despacho hidrotérmico estocástico com restrições de rede, da empresa PSR
SIN	Sistema Interligado Nacional
SUISHI-O	Modelo de Simulação a Usinas Individualizadas para Subsistemas Hidrotérmicos Interligados

¹ As terminologias “irradiação” e “irradiância” são dadas pela (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2013).

SUMÁRIO

LISTA DE FIGURAS	8
LISTA DE TABELAS	13
LISTA DE SIGLAS, ABREVIATURAS E TERMOS	14
1 INTRODUÇÃO	18
1.1 CONSIDERAÇÕES GERAIS	18
1.2 OBJETIVO	20
1.3 ORGANIZAÇÃO DO DOCUMENTO	20
2 CENÁRIOS SINTÉTICOS NO SETOR ELÉTRICO BRASILEIRO	22
2.1 CONTEXTUALIZAÇÃO	22
2.2 USO DE SÉRIES SINTÉTICAS	25
2.3 PROCESSO ESTOCÁSTICO	28
2.4 OUTRAS APLICAÇÕES.....	32
3 SÉRIES TEMPORAIS	35
3.1 CONTEXTUALIZAÇÃO	35
3.2 CLASSIFICAÇÃO DOS PROCESSOS	36
<i>Processos Determinísticos.....</i>	<i>36</i>
<i>Processos Estocásticos</i>	<i>36</i>
3.3 PROCESSOS ESTACIONÁRIOS.....	37
<i>Estacionariedade por períodos</i>	<i>39</i>
3.4 CLASSES DE PROCESSOS ESTOCÁSTICOS E PROPRIEDADES	40
<i>Ruído branco</i>	<i>41</i>
<i>AR - Autorregressivo</i>	<i>41</i>
<i>MA - Média móvel.....</i>	<i>42</i>
<i>ARMA - Autorregressivo com média móvel</i>	<i>43</i>
<i>ARIMA - Autorregressivo Integrado com Média Móvel</i>	<i>44</i>
3.5 MÉTODO DE BOX-JENKINS.....	45
3.6 GERAÇÃO DE CENÁRIOS DE RADIAÇÃO SOLAR	48
4 ESTIMATIVA DE PARÂMETROS DOS MODELOS.....	51
4.1 ANÁLISE DOS DADOS	52
<i>Considerações</i>	<i>52</i>
<i>Análise preliminar</i>	<i>52</i>
4.2 PRÉ-PROCESSAMENTO DOS DADOS	55

<i>Método 1: retirada de tendência pela média diária</i>	56
<i>Método 2: retirada da tendência pelo máximo diário</i>	58
<i>Método 3: retirada da tendência pela média mensal</i>	60
<i>Comentários</i>	61
4.3 GERAÇÃO DE SÉRIES SINTÉTICAS PELO MÉTODO DE YULE-WALKER	61
<i>Método</i>	61
<i>Estimativa dos parâmetros dos modelos</i>	62
<i>Análise da distribuição das séries</i>	64
4.4 GERAÇÃO DE SÉRIES SINTÉTICAS PELO MÉTODO DOS MÍNIMOS QUADRADOS	69
<i>Método</i>	69
<i>Estimativa dos modelos</i>	69
<i>Análise da distribuição das séries</i>	71
<i>Comentários</i>	73
4.5 MODELO PARA GERAÇÃO DE SÉRIES DE RADIAÇÃO MENSAL	74
4.6 MODELO PARA GERAÇÃO DE SÉRIES DE RADIAÇÃO HORÁRIA	75
4.7 RESULTADOS	78
<i>Modelos para geração de séries sintéticas de radiação diária</i>	78
<i>Modelos para geração de séries sintéticas de radiação mensal</i>	81
<i>Modelos para geração de séries sintéticas de radiação horária</i>	83
4.8 CONCLUSÕES PARCIAIS	86
5 APRIMORAMENTO DO MODELO	88
5.1 ESTIMATIVAS DO MODELO AR	89
5.2 MODELO AUTORREGRESSIVO PERIÓDICO - PAR E ESTIMATIVAS	93
5.3 VERIFICAÇÃO	96
5.4 COMENTÁRIOS	98
6 APLICAÇÕES PRÁTICAS EM LEILÕES E EM RISCOS DE GERAÇÃO	99
6.1 REGRAS DE LEILÕES DE ENERGIA	99
6.2 GERAÇÃO DE SÉRIES SINTÉTICAS	101
6.3 PREMISSAS	103
6.4 RECEITA ESPERADA	103
6.5 DEFINIÇÃO DO MELHOR LANCE DE ENERGIA	105
6.6 PERDA DE RECEITA	107
6.7 PROBABILIDADES DE GERAÇÃO ANUAL	108
6.8 VARIAÇÃO DA RECEITA ESPERADA	110
7 CONCLUSÕES	113
7.1 CONSIDERAÇÕES FINAIS	113

7.2	CONTRIBUIÇÕES.....	115
7.3	TÓPICOS PARA ULTERIOR DESENVOLVIMENTO	116
7.4	PUBLICAÇÕES REALIZADAS.....	117
REFERÊNCIAS		119
ANEXO I – REMOÇÃO DA TENDÊNCIA PELA MÉDIA DIÁRIA		124
ANEXO II – REMOÇÃO DA TENDÊNCIA PELO MÁXIMO DIÁRIO.....		126
ANEXO III – REMOÇÃO DA TENDÊNCIA PELA MÉDIA MENSAL.....		127
ANEXO IV – INFLUÊNCIA DO TAMANHO DO CONJUNTO DE DADOS NA ESTIMATIVA DOS PARÂMETROS DO MODELO AR.....		128
ANEXO V – ANÁLISE DAS SÉRIES GERADAS PELO MÉTODO DOS MÍNIMOS QUADRADOS		133
ANEXO VI – ANÁLISE DAS SÉRIES DE RADIAÇÃO HORÁRIA (MMQ)		150

1 INTRODUÇÃO

1.1 CONSIDERAÇÕES GERAIS

O sistema elétrico brasileiro possui parque gerador com significativa participação de fontes renováveis, cujos insumos dependem de fatores não controláveis e da região geográfica em que se encontram as usinas. A diversidade de características dos recursos naturais e a dimensão continental do sistema torna difícil sua previsibilidade, consequentemente criando um desafio significativo e único para a operação eletro-energética.

Em vista deste desafio, o setor elétrico desenvolveu nas últimas décadas metodologias para melhor aproveitar os insumos disponíveis, otimizando tanto o despacho individual como o do conjunto das usinas do Sistema Interligado Nacional – SIN. Atualmente a estratégia consolidada, e adotada pelo governo para o planejamento e a operação do SIN, é o estudo energético baseado em diversos cenários de geração, criados de forma estocástica, e utilizando apenas a geração de fonte hídrica, dada sua significativa representatividade e importância.

Entretanto, a geração hidrelétrica tem perdido representatividade na matriz energética, sendo que, apesar de ainda corresponder a mais de dois terços da capacidade instalada, outras fontes começaram a ganhar destaque, seja comercialmente ou pelos potenciais novos desafios que possam apresentar ao sistema.

A primeira fonte com este destaque é a energia eólica², que foi introduzida localmente através do Programa de Incentivo às Fontes Alternativas de Energia Elétrica – PROINFA e ganhou evidência a partir de 2010 por meio dos leilões de energia do mercado regulado. A fonte passou a despertar preocupação do setor assim que ganhou alguma representatividade, principalmente por conta de sua intermitência, da não despachabilidade e, mais recentemente, da dificuldade de previsão energética.

Por iniciativa da Agência Nacional de Energia Elétrica – ANEEL, através de seu programa de pesquisa e desenvolvimento, foi realizado o Projeto Estratégico “Modelo de otimização do despacho hidrotérmico”, que apesar de ter outro foco, contou com contribuições de alguns dos

² Importante notar que a fonte mais representativa após a hidrelétrica é a térmica, porém esta não apresenta o mesmo crescimento relativo e tem insumo de fácil previsibilidade, não implicando nos desafios citados.

pesquisadores no estudo do insumo eólico. Esse tema também é estudado pelo Centro de Pesquisas de Energia Elétrica – CEPEL, responsável pela implementação do modelo oficial do governo para o planejamento e operação do sistema – NEWAVE.

Atualmente acredita-se que a próxima fonte a apresentar desafios ao SIN é a energia solar, cuja capacidade instalada é ainda pouco representativa, porém possui elevada expectativa de crescimento. Essa perspectiva surge do enorme potencial solar disponível no território brasileiro, da evolução dos compromissos com a sustentabilidade e da constante redução do custo de investimento. Assim sendo, este crescimento aliado às características da geração, como a não despachabilidade e a sazonalidade, tornar-se-á um próximo desafio ao setor elétrico, no planejamento e na operação tanto do sistema interligado como das redes de distribuição de energia.

A superação deste desafio, entretanto, carece de informações primordiais como o conhecimento da energia incidente; porém, a disponibilidade de dados solares, assim como acontece em muitas outras áreas, é um problema crítico no Brasil. A maioria das estações meteorológicas existentes é particular e tem fins principalmente direcionados para a agricultura, sem o registro específico da radiação. Isto ocorre tanto pelo elevado custo dos equipamentos adicionais necessários, como pelo pequeno benefício agregado ao proprietário.

Do ponto de vista governamental também não houve iniciativa em se ter estações próprias. As poucas existentes e com manutenção adequada estão instaladas em aeroportos, mas igualmente não possuem medição da radiação por ser uma informação desnecessária para suas necessidades.

Como estratégia, assim como praticado com a energia eólica, foram impostas exigências de medição aos empreendedores do setor elétrico, que são obrigados a possuir estação meteorológica própria, medindo previamente e durante a operação de seus parques, e entregar periodicamente os dados à Empresa de Pesquisas Energética – EPE. Os dados entregues não são atualmente disponibilizados ao público. Há, entretanto, diversos estudos solarimétricos promovidos pelo Instituto Nacional de Pesquisas Espaciais – INPE, utilizando, majoritariamente, medições de radiação por satélites que são, via de regra, menos precisas que as medições terrestres.

Em contrapartida às carências supracitadas, há questionamentos mal resolvidos sobre a real necessidade de medição da radiação para a realização de estudos de geração, tanto de

dimensionamento elétrico como de previsão da produção energética. Especialistas internacionais, em geral, defendem a utilização única de dados provenientes de modelos baseados em medições de satélites, enquanto que a regulamentação brasileira exige medições detalhadas, realizadas por equipamentos específicos instalados em solo, apesar de não apresentar estudos para embasar e justificar tal demanda.

Por fim, dentro do contexto apresentado, destaca-se a necessidade de se aprofundar o conhecimento da radiação solar do ponto de vista das necessidades e práticas da geração de energia elétrica no setor elétrico brasileiro.

1.2 OBJETIVO

A proposta deste trabalho é contribuir com o conhecimento do recurso solar, com fins aplicados à geração de energia elétrica no contexto brasileiro, estudando probabilisticamente o insumo e suas aplicações e oferecendo subsídios para estudos energéticos, em especial para avaliação de estratégias de comercialização. Para tanto, foram desenvolvidas duas frentes principais: modelos de séries temporais e aplicações práticas.

Após a obtenção e análise de consistência de dados solarimétricos, foram avaliados modelos existentes de séries temporais, como autorregressivos com ou sem componente de média móvel, e o período de análise dos dados, propondo-se modelo para a geração de séries sintéticas de radiação solar para uma localidade com medição disponível, com fins de planejamento e operação energéticos.

Uma vez obtidas séries sintéticas que representem possíveis cenários de radiação, foram conduzidas análises aplicadas, como avaliações de incertezas de geração e de riscos de contratos de longo prazo, tanto do ponto de vista do gerador como do mercado comprador da energia, oferecendo importantes subsídios para a tomada de decisões e estratégias de negociação.

1.3 ORGANIZAÇÃO DO DOCUMENTO

A pesquisa desenvolvida e os resultados obtidos são apresentados nos capítulos seguintes. Inicialmente, o Capítulo 2 contextualiza a utilização de cenários probabilísticos no setor nacional, realizado pela geração de séries sintéticas obtidas a partir da modelagem estocástica.

Em seguida foi realizada a revisão bibliográfica de séries temporais, apresentada no Capítulo 3, com o detalhamento dos tipos de processo existentes, da influência da estacionariedade, das propriedades das estruturas autorregressivas e do método utilizado para a estimativa dos modelos. A revisão é concluída com o levantamento do estado da arte da modelagem da radiação solar, identificando os principais trabalhos sobre o tema, seus avanços e tendências na obtenção de modelos.

O Capítulo 4 apresenta o estudo de diferentes métodos para a aplicação dos modelos autorregressivos com dados de radiação solar, avaliando, além das estruturas dos modelos, a retirada de tendências e sazonalidades, a estimativa pelos métodos de Yule-Walker e dos mínimos quadrados e a periodicidade de análise. Após a estimativa e avaliação de mais de 200 modelos, são destacados os mais adequados para possíveis aplicações práticas.

O Capítulo 5 complementa o estudo dos modelos com foco em séries de longo prazo, introduzindo nova estrutura de modelo para séries mensais, a avaliação de sua dependência temporal e a verificação por dois métodos de comparação adicionais, recomendando o modelo ideal desta periodicidade para estudos futuros.

Este modelo é utilizado em aplicações práticas no Capítulo 6, onde são geradas séries sintéticas de radiação solar com periodicidade mensal, posteriormente avaliadas por modelo de contrato de comercialização de energia, com implicações nos riscos de suprimento, na estratégia de venda da energia e na variação da receita esperada.

Por fim, o Capítulo 7 apresenta as conclusões da pesquisa, suas principais contribuições e sugestões para futuros desenvolvimentos.

2 CENÁRIOS SINTÉTICOS NO SETOR ELÉTRICO BRASILEIRO

O modelo atualmente utilizado em estudos de planejamento e operação do setor elétrico brasileiro é o principal ponto de partida para o levantamento do estado da arte desta pesquisa, tanto por suas particularidades como por ser base para a aplicação dos resultados previstos. Deste modo, este capítulo consolida as principais informações levantadas neste tema e relevantes para a contextualização e o desenvolvimento da pesquisa.

2.1 CONTEXTUALIZAÇÃO

A necessidade da geração de cenários energéticos artificiais surgiu da demanda do setor elétrico de se utilizar critérios probabilísticos na simulação do sistema, permitindo a análise de diversas condições operativas e melhor avaliação dos riscos e da confiabilidade do suprimento. Essa necessidade é destacada, entre outros trabalhos, em (MACEIRA, PENNA e DAMÁZIO, 2006): “em estudos energéticos, critérios de suprimento são baseados em índices de risco, estimados a partir da simulação da operação energética do sistema para diversos cenários de afluências aos aproveitamentos hidroelétricos ou subsistemas”.

Ou seja, o objetivo inicial de melhorar a avaliação do atendimento elétrico e energético do sistema, através de simulações e análise de riscos, demanda a modelagem probabilística de diversas de suas variáveis de entrada, como, por exemplo, os diversos insumos energéticos e os riscos de falhas na geração e transmissão. Desta forma, quanto mais variáveis modeladas, melhor será o resultado da simulação do sistema completo.

Entretanto, a modelagem ideal, de todas as variáveis disponíveis, é fisicamente inviável, tanto pelo enorme volume de trabalho necessário para o estudo específico de cada variável, como pelas exigências de processamento computacional que seriam demandadas. Adicionalmente, uma parcela dos dados de entrada não agregaria resultados significativos por não serem representativos ou por não apresentarem variabilidade estatística.

A Figura 1 ilustra a capacidade instalada da geração no Brasil, onde se pode observar que a capacidade instalada hidrelétrica é bastante superior às outras fontes e responde como a principal responsável pelo atendimento elétrico. Do ponto de vista energético é ainda mais significativa, pois tem despacho prioritário e muito mais frequente. As fontes renováveis

principais são de tecnologia mais recente e ainda não são representativas, porém se encontram em franco crescimento.

Figura 1. Capacidade instalada da geração (MW)

Fonte: (MINISTÉRIO DE MINAS E ENERGIA; EMPRESA DE PESQUISA ENERGÉTICA, 2017)

Quanto à previsibilidade do insumo, resumida na Tabela 1, as fontes que dependem da natureza são desafios para o sistema, pois possuem componentes aleatórios significativos e não controláveis, com exceção feita às hidrelétricas com capacidade de regularização, que podem reduzir os efeitos da ausência ou excesso de afluência.

Tabela 1. Previsibilidade de geração

Fonte	Previsibilidade do insumo	Importância da previsão
Hidro	aleatória / difícil (redução da capac. regularização)	alto impacto (representatividade e aleatoriedade)
Termo	fácil	pequena
Eólica	aleatória / difícil (tendências sazonais e regionais)	pequena, porém crescente
Solar	aleatória / difícil (nebulosidade, particulados, poluição...)	tecnologia em franco crescimento

Fonte: (MINISTÉRIO DE MINAS E ENERGIA; EMPRESA DE PESQUISA ENERGÉTICA, 2015)

Desta forma, destaca-se a importância da previsão do insumo hídrico, pela sua aleatoriedade e pela representatividade do potencial hidrelétrico. As fontes eólica e solar, apesar de ainda pouco impactantes, apresentam desafio em sua previsibilidade, embora já se tenha registro de

difficultades operativas em locais com grande concentração de parques eólicos, criando desafios para a estabilidade do sistema elétrico.

Convencionalmente, e partindo deste contexto, no sistema brasileiro apenas os cenários hidrológicos são modelados estatisticamente, principalmente devido à sua significativa representatividade e dificuldade de previsão. A operação térmica, por outro lado, tem como seu insumo – seja gás, óleos ou urânio – uma variável de disponibilidade bem definida e apresenta riscos de suprimento bastante baixos, sendo modelada – não estatisticamente – apenas por meio de taxas de indisponibilidade, programada e forçada, oriundas da manutenção e operação das plantas de geração. Cabe lembrar que a geração térmica oriunda da biomassa tem forte componente sazonal, mas tem contribuição irrelevante na produção de energia no Brasil. As fontes eólica e solar não são modeladas como geradores, mas consideradas de forma simplista e incompleta como demandas negativas e contínuas.

Destaca-se que os modelos de geração são compostos pela representação matemática do gerador e do seu insumo, porém a modelagem probabilística só se aplica à disponibilidade do insumo. Por exemplo, como se observa na Figura 2 e na Figura 3, e nas Equações [1] e [2], referentes a parte do equacionamento de um gerador hidrelétrico, o modelo do gerador em si é composto por diversos coeficientes e equacionamentos, sendo representado de forma puramente determinística. Este modelo, apesar de relativamente complexo, depende de conhecimento consolidado e tem resultados precisos, não apresentando desafios quando se trata da simulação da operação de um sistema elétrico. Por outro lado, o insumo energético – no caso, a afluência que alimenta o reservatório – dadas as suas características naturais e aleatórias, demanda o tratamento probabilístico e passa a ser o grande desafio para o desenvolvimento do modelo.

Figura 2. Modelo do gerador hidrelétrico

Fonte: Apresentação da disciplina “Formação de Preços e Comercialização de Energia”, Prof. Dorel S. Ramos, 2016, EPUSP

$$P_g(x, q, v) = k \cdot \eta(x, q, v) \cdot h_l(x, q, v) \cdot q \quad [1]$$

$$h_l(x, q, v) = h_{mon}(x) - h_{jus}(q + v) - pc(x, u) \quad [2]$$

Onde: as variáveis são explicitadas na Figura 2.

Fonte: Apresentação “O Sistema Interligado Nacional”, FEEC-Unicamp

2.2 USO DE SÉRIES SINTÉTICAS

O estudo da geração elétrica com foco no insumo energético cria uma primeira questão quanto ao processo de definição dos dados: por que criar modelos de geração sintética e não utilizar simplesmente o histórico de afluições?

Esse assunto foi bastante discutido por especialistas do setor elétrico brasileiro nas décadas de 1970 e 1980 (e eventualmente ainda é), tendo sido realizados diversos estudos baseados na geração de cenários artificiais. A seguir são destacadas as principais pesquisas realizadas neste âmbito.

Os primeiros estudos, como (BARRETO e GUERRERO-SALAZAR, 1975) e (GUERRERO-SALAZAR, *et al.*, 1977), propõem a aplicação no Brasil de séries sintéticas e destacam as vantagens de seu uso em relação ao de dados históricos, apresentando resultados das experiências realizadas na geração de séries de afluição mensais. Em seguida, (KELMAN, *et*

al., 1979) realiza uma aplicação prática na avaliação energética de bacias, alimentando simuladores de usinas com as vazões obtidas artificialmente.

Em (KELMAN, OLIVEIRA e PEREIRA, 1979) e, poucos anos depois, em (PEREIRA, *et al.*, 1984), os autores propuseram modelos estocásticos de vazões mensais baseados na desagregação de vazões anuais. O primeiro realiza a modelagem multivariada, na qual “obtidas as séries geradas anuais, que são dependentes entre si, processa-se a desagregação em separado de cada série”, e o segundo contribui adicionando análises em novas localidades e a correção de valores de vazões negativas.

Ainda na década de 1980, (MACEIRA, KELMAN e DAMÁZIO, 1987) e (KELMAN, 1987) fazem críticas ao modelo vigente na época, propondo, respectivamente, metodologias para a avaliação de modelos PAR (explicado na Seção 2.3) e o uso de séries sintéticas para o cálculo da energia firme.

Além das modelagens com base anual e mensal, outros trabalhos focaram no estudo da estocasticidade diária. (KELMAN, 1979) estudou a precipitação baseado em medições em campo nos EUA, concluindo que a geração sintética de dados de estações confiáveis é melhor que o uso dos registros históricos. Em (NASCIMENTO e KELMAN, 1995) esse foco é estudado com outra abordagem, apresentando “um modelo de geração de chuvas diárias em que a ocorrência do processo é modelada pela abordagem das sequências alternadas de dias secos e chuvosos”. O autor faz ainda um complemento em (KELMAN, 1980), desenvolvendo um modelo para vazões diárias, que conta com representações independentes de fatores hidrológicos externos à bacia e seus reservatórios.

Outra aplicação relevante da geração sintética de aflúências é no controle de cheias e dimensionamento de barragens e de vertedouros. Os trabalhos (KELMAN, DAMÁZIO e COSTA, 1985), (KELMAN, 1984), (KELMAN e DAMAZIO, 1983) e (DAMÁZIO, *et al.*, 1983) abordam de formas diversas essas questões, propondo modelos de geração de vazões sem a necessidade de dados de precipitação, com diferentes tempos de retorno ou com o objetivo de estudar as variáveis de controle de um reservatório. Além de suas particularidades, todos os trabalhos apresentam conceitos importantes para a análise do insumo energético em questão, especialmente quanto à modelagem estatística dos dados.

A consolidação dos estudos e das discussões sobre o uso de cenários sintéticos é bem justificada no *Manual de Referência do Modelo NEWAVE* em (ELETROBRAS; CEPEL): “o registro de

afluências observado no passado (chamado de série histórica) é, no entanto, insuficiente para compor uma amostra de tamanho necessário para estimar índices de risco com incertezas aceitáveis”. Da mesma forma, em publicação mais recente, (MACEIRA, PENNA e DAMÁZIO, 2006) destaca que “o histórico de afluências compõe uma amostra insuficiente para estimar índices de risco com incertezas aceitáveis”

Em outras palavras, a série histórica de vazões (exemplo do Rio Madeira na Figura 4), atualmente com mais de 80 anos de registro para diversas usinas brasileiras, não é capaz de simular os possíveis cenários de geração de forma completa e com precisão adequada.

Figura 4. Histórico de vazões no Rio Madeira em Porto Velho

Fonte: (MACEIRA, PENNA e DAMÁZIO, 2006)

Como alternativa (ELETROBRAS; CEPEL) destaca que “as características básicas da série histórica podem ser capturadas por modelos estocásticos capazes de produzir séries sintéticas de afluências, diferentes da série histórica mas igualmente prováveis”.

Ou seja, é possível se produzir um conjunto infinito de séries artificiais, baseado em um conjunto finito – e pequeno – de séries históricas, porém com igual probabilidade de ocorrência e conservando as informações que a caracterizam, permitindo melhor avaliação da operação da própria usina (otimização do despacho, controle de cheias, regularização etc.) e dos sistemas que ela compõe (otimização da operação e do planejamento, da geração e da transmissão).

Atualmente, o setor elétrico brasileiro faz uso de diferentes modelos de planejamento da operação energética, conforme o nível de detalhamento necessário para a representação do sistema, o período de estudo e as restrições computacionais. Independentemente do modelo utilizado, a aleatoriedade das vazões é sempre considerada, sendo apenas adaptada conforme o caso. Os modelos de planejamento citados são NEWAVE (médio prazo, com reservatórios

agregados), SUIISHI-O (desdobramento do NEWAVE, com usinas individualizadas), DECOMP (curto prazo), todos desenvolvidos pela ELETROBRAS-CEPEL e com utilização imposta por influência governamental para os órgãos públicos federais. Outros modelos são utilizados em paralelo pelo mercado, como o SDDP da PSR e o HYDROLAB da Unicamp, porém sem status oficial

Os três modelos do CEPEL “simulam um grande número de séries hidrológicas, calculando assim, índices probabilísticos de desempenho do sistema para cada estágio de simulação” (MACEIRA, PENNA e DAMÁZIO, 2006). Essas séries hidrológicas, em todos os casos, são geradas por um outro modelo, chamado GEVAZP.

O GEVAZP realiza basicamente o mesmo processo de geração de afluências para todos os modelos de planejamento, diferenciando-se principalmente nas aplicações do DECOMP, para o qual os cenários hidrológicos são agrupados por técnicas específicas, criando um conjunto de cenários capaz de representar a estocasticidade das vazões. Esta estratégia é utilizada para reduzir o número de cenários e, conseqüentemente, as necessidades de processamento computacional.

2.3 PROCESSO ESTOCÁSTICO

Um processo estocástico é aquele que possui ao menos um termo aleatório, fazendo com que sua saída não possa ser determinada diretamente a partir dos dados de entrada, mas seja possível de ser estimada se conhecidas as probabilidades de suas variáveis. Adicionalmente, há o conceito de séries temporais, onde as variáveis analisadas são registradas a cada determinado intervalo de tempo.

Aplicando este conceito no caso hidrelétrico, temos um processo estocástico da natureza que é responsável pela criação de séries temporais de vazões. Os resultados deste processo, apesar de apresentarem características típicas sazonais, geográficas e com diversas interdependências, entre outras, não podem ser previstos exatamente através do uso de equações e modelos matemáticos determinísticos por conterem uma ou mais variáveis aleatórias.

A análise das séries temporais do histórico de vazões, através do estudo das distribuições de probabilidades, permite a representação do processo estocástico através de um modelo, ajustado para gerar dados – no caso, séries sintéticas – com as mesmas características e a mesma

probabilidade de ocorrência da série histórica. Esse processo, também descrito e exemplificado em (COSE - UNICAMP, 2010) e (PROMON, 1992), é ilustrado na Figura 5. Destaca-se que as séries geradas podem nunca ocorrer na realidade, mas refletem as diversas realizações possíveis caracterizadas pela análise estocástica da série histórica.

Em (ELETROBRAS; CEPEL) há uma descrição complementar: “um processo estocástico é totalmente descrito pelo conjunto de todas as séries temporais que o compõe ou pela distribuição de probabilidades conjunta de todas as variáveis aleatórias envolvidas. Como na realidade não está disponível nem uma coisa nem outra, a modelagem de séries sintéticas tem por objetivo ajustar um modelo pelo qual acredita-se que a série histórica tenha sido produzida e a partir dele gerar séries sintéticas que representam as séries temporais que podem ser amostradas pelo processo físico que se está observando, o processo estocástico”.

Figura 5. Diagrama de análise de séries temporais

No ajuste do modelo tem-se por costume procurar simplificações, como a identificação de características constantes ou estacionárias³. No caso das vazões, tem-se que “o processo de afluições anuais é considerado estacionário e o processo de afluições mensais é considerado não estacionário” (ELETROBRAS; CEPEL), pois a média da vazão anual em determinado ponto de um rio não sofre variações significativas com o passar dos anos – salvo em casos onde há interferências externas à natureza – e a vazão mensal possui características completamente diferentes conforme o período do ano.

Para identificar as estacionariedades e relações entre os dados da amostra, são calculadas propriedades estatísticas diversas, como a média [3], o desvio padrão [4], a covariância [5] e a correlação [6].

$$\mu = \frac{\sum_{i=1}^n z_i}{n} \quad \sigma = \sqrt{\frac{\sum_{i=1}^n (z_i - \mu)^2}{n}} \quad [3]; [4]$$

$$\gamma(j) = \frac{\sum_{i=1+j}^n (z_i - \mu)(z_{i-j} - \mu)}{n} \quad \rho(j) = \frac{\gamma(j)}{\sigma^2} \quad [5]; [6]$$

Onde:

- μ média da amostra
- z_i valor amostral da observação i
- n quantidade de dados da amostra
- σ desvio padrão da amostra
- $\gamma(j)$ covariância de ordem j
- $\rho(j)$ correlação de ordem j

A média e o desvio padrão são operações bastante convencionais e representam, respectivamente, o valor médio amostral e o grau de dispersão da amostra em torno da média. Esses valores são resultados simples e calculados apenas uma vez por período estudado, porém seus resultados são base para toda a análise da série temporal.

³ Em um processo estacionário as probabilidades da variável em estudo não se alteram com o tempo, ou seja, a variável possui média, desvio padrão, covariâncias, etc., porém estas são constantes para toda a série. Detalhamento adicional é apresentado no Capítulo 3.

O cálculo da covariância permite identificar a dependência temporal entre duas observações da variável amostrada, espaçadas por um intervalo de tempo fixo (ordem). Em uma série de afluições com valores observados anualmente, a covariância de ordem um representa a dependência típica da vazão entre dois anos consecutivos.

A correlação é apenas a padronização da ordem de grandeza da covariância para a base “1”, ou seja, uma correlação igual a um (de primeira ordem) indica que a vazão anual é completamente dependente do ano anterior e a correlação igual a zero significa total independência entre vazões de anos consecutivos.

Ao alterar o período de estudo para intervalos com características cíclicas, como a sazonalidade anual, lunar, ou diária, entre muitas outras, a análise é realizada novamente para cada período e para diversas ordens, coerentes com as características da periodicidade do ciclo. No caso da amostragem mensal, por exemplo, são calculadas as médias e os desvios padrão para cada mês, como nas Equações [7]; [8].

$$\mu_m = \frac{\sum_{i=1}^n z_{12(i-1)+m}}{n} \quad \sigma_m = \sqrt{\frac{\sum_{i=1}^n (z_{12(i-1)+m} - \mu_m)^2}{n}} \quad [7]; [8]$$

Onde:

m mês

μ_m média do mês m da amostra

σ_m desvio padrão do mês m da amostra

O cálculo da correlação em períodos cíclicos não apresenta a mesma simplicidade da média e do desvio padrão, pois é necessário calcular diversas ordens de correlação, para cada intervalo do período. No exemplo da amostragem mensal, ter-se-ia covariâncias e correlações diversas para cada um dos 12 meses do ano, [9] e [10]. O conjunto de valores de correlações de um mesmo mês é chamado de função de autocorrelação⁴ do mês.

⁴ Destaca-se que o termo autocorrelação (assim como autocovariância) nada mais é do que correlação de uma série temporal com ela mesma, porém deslocada no tempo por um intervalo pré-estabelecido, conhecido por ordem da correlação e representado na equação [5]; [6] pelo termo j .

$$\gamma_m(k) = \frac{\sum_{i=1}^n (z_{12(i-1)+m} - \mu_m)(z_{12(i-1)+m-k} - \mu_m)}{n} \quad [9]$$

$$\rho_m(k) = \frac{\gamma_m(k)}{\sigma_m \sigma_{m-k}} \quad [10]$$

Onde:

$\gamma_m(k)$ covariância de ordem k para o mês m

$\rho_m(k)$ correlação de ordem j para o mês m

No caso de séries de afluência mensais, como as utilizadas no setor elétrico, é comum se utilizar modelos do tipo PAR(p), conhecidos por modelos autorregressivos periódicos. Este modelo nada mais é que uma formulação autorregressiva com parâmetros de comportamento periódico.

Partindo desta caracterização estocástica das séries, a metodologia padrão de análise faz uso de técnicas originalmente propostas por (BOX, JENKINS e REINSEL, 2008) e consiste de três etapas principais: escolher a ordem do modelo, estimar os seus parâmetros e testar os resultados. O Capítulo 3 detalha essas etapas após apresentar a contextualização dos conceitos e práticas convencionais para o estudo de séries temporais.

2.4 OUTRAS APLICAÇÕES

A modelagem dos insumos energéticos impactantes na geração elétrica nacional ainda é bastante limitada ao recurso hídrico, sendo este o único realmente implementado nos modelos oficiais do setor. Entretanto, outros estudos já foram e ainda são conduzidos com a intenção de melhor entender as outras fontes primárias.

Dentre as iniciativas realizadas, destacam-se os estudos de universidades dentro dos projetos do programa de P&D Estratégico nº 001/08 (ANEEL - AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA, 2008), que, apesar de ter como objetivo a “modelagem da otimização do despacho hidrotérmico”, contou com contribuições de alguns dos executores no estudo do comportamento da velocidade dos ventos. Esses estudos, entretanto, ainda não são aplicados extensivamente.

Do lado governamental, nota-se que há interesse no mapeamento eólico por parte da EPE, por meio de seu Sistema de Acompanhamento das Medições Anemométricas (MINISTÉRIO DE MINAS E ENERGIA; EMPRESA DE PESQUISA ENERGÉTICA), com base em informações exigidas dos empreendedores que são utilizadas em publicações com histórico e com caracterização da produção energética, sem fins de modelagem do insumo.

Por fim, o órgão oficial para o desenvolvimento dos modelos utilizados no setor elétrico, o CEPEL, já realizou alguns estudos com esse enfoque, como em (MACEIRA, *et al.*, 2011). Esse trabalho trata especificamente da modelagem estatística de ventos para a utilização no planejamento e operação, justificando que o modelo atual (média histórica) “pode não refletir corretamente os valores mínimos e máximos, a correlação temporal”, e ressalta a necessidade do “desenvolvimento de novos modelos e ferramentas para análise dos impactos da estocasticidade dos ventos”, com “análise exploratória das séries históricas (...), visando a identificação de padrões de sazonalidade e demais fatos (...) que descrevem o comportamento do processo estocástico que se pretende modelar e que são fundamentais para a especificação dos modelos de previsão e geração de cenários sintéticos”.

Por fim, após apresentar suas motivações e potencialidades, o trabalho de (MACEIRA, *et al.*, 2011) realiza a análise exploratória de uma série de velocidade de vento de um ano, apresentando algumas estatísticas obtidas e resultados dos testes de aderência das distribuições de probabilidade Normal, Lognormal, Gama, Beta e Weibull. A Figura 6 apresenta a função de autocorrelação horária e diária e o histograma ajustado à distribuição Weibull, escolhida como a melhor representação para a série estudada.

Fonte: (MACEIRA, *et al.*, 2011)

Destaca-se algumas semelhanças das justificativas e objetivos do trabalho citado com a presente pesquisa, diferindo principalmente quanto ao insumo analisado. Entretanto, este estudo pretende realizar, além da modelagem da radiação solar como um processo sob incerteza, uma pesquisa mais aprofundada das correlações temporais (utilizando maior quantidade de dados, tanto locais como baseados em diferentes períodos de registro), o desenvolvimento e implementação de modelo para a geração de séries sintéticas e aplicações práticas para avaliação de potencial, de riscos de projetos e de comercialização de energia, entre outros.

3 SÉRIES TEMPORAIS

3.1 CONTEXTUALIZAÇÃO

Uma série temporal é um conjunto de observações de uma determinada variável registradas ao longo do tempo, como, por exemplo, as taxas de juros mensais em um ano e as médias diárias de temperatura dentro de um mês. A Figura 7 e a Figura 8 ilustram exemplos aplicados ao insumo solar: a radiação diária média e a irradiância durante, respectivamente, um ano e um mês, em localidade no interior do nordeste brasileiro.

Figura 7. Radiação diária média em localidade de NE

Fonte: Autor.

Figura 8. Irradiância no mês de janeiro

Fonte: Autor.

Alternativamente, como citado no capítulo anterior, pode-se definir uma série temporal como “o registro de uma realização de um processo estocástico”. Essa definição do tipo do processo provém de uma série de classificações já consolidadas e bem definidas na literatura de engenharia de sistemas, como em (BOX, JENKINS e REINSEL, 2008) e (MADSEN, 2007), e de econometria, (ENDERS, 2014).

A seguir, este capítulo faz uma revisão bibliográfica sobre a classificação de processos, em especial os estocásticos, analisando os modelos para sua representação, comparando as classes existentes, suas propriedades características e o método indicado para a modelagem de séries temporais.

3.2 CLASSIFICAÇÃO DOS PROCESSOS

Os processos são classificados conforme suas características, utilizando diversos critérios como linearidade, aleatoriedade, continuidade e variância no tempo, entre outros. Destes, a primeira análise convencionalmente utilizada é a determinação da existência de componentes aleatórios no processo, diferenciando-o entre determinístico ou estocástico.

Processos Determinísticos

Um processo classificado como determinístico é aquele que conhecida a sua entrada é possível determinar sua saída, podendo ser descrito matematicamente através de equações sem componentes aleatórios. Um exemplo simples deste processo é a corrente elétrica provocada pela aplicação de uma diferença de potencial em um resistor ideal, sendo o seu modelo o equacionamento da Primeira Lei de Ohm, ilustrada na Figura 9. Neste tipo de processo, uma vez conhecidos os dados de entrada, sempre é possível determinar exatamente qual será o resultado.

Figura 9. Medição da corrente em um resistor ideal

Fonte: Autor.

Processos Estocásticos

Complementarmente ao processo determinístico é classificado o processo estocástico, no qual o conhecimento da entrada não permite a determinação exata da saída, mas apenas sua estimativa baseada nas características probabilísticas que regem seu comportamento. Um

exemplo cotidiano deste processo é a temperatura ambiente, ilustrada na Figura 10 com os valores históricos diários da cidade de São Paulo em maio de 2016.

Os processos estocásticos são muitas vezes associados a processos da natureza, mas não se limitam a estes. Atividades controladas diretamente por humanos também são enquadradas desta forma, como as cotações de ações e de moedas no mercado financeiro ou a disponibilidade da rede celular durante a celebração do Ano Novo.

Comparando a Figura 9 e a Figura 10, pode-se notar a clara relação de proporcionalidade no caso da corrente medida quando se conhece a tensão aplicada no resistor, enquanto que na série histórica da temperatura não se verifica uma relação clara entre os valores de tempo e temperatura.

De tal modo, este tipo de processo é descrito através de funções densidade de probabilidade ou equações que contenham termos aleatórios, não sendo possível determinar exatamente a saída, mas sim uma previsão de resultado com uma incerteza associada.

Consolidando, um processo com saída bem determinada por uma equação sem termos aleatórios é determinístico e um processo com equações dependentes de ao menos um termo aleatório é estocástico.

3.3 PROCESSOS ESTACIONÁRIOS

A estacionariedade de um processo pode ser interpretada como uma “forma de equilíbrio estatístico estável”, refletida por “uma série que se desenvolve aleatoriamente no tempo, em torno de uma média constante” (FEIJÓ, 2010). Em termos mais técnicos, define-se um processo

estocástico como estritamente estacionário se as suas funções de distribuição de probabilidade conjuntas permanecem constantes mesmo quando o processo é transladado no tempo.

A definição deste tipo de classificação, apesar de simples, implica em uma análise árdua para a obtenção das funções de autocorrelação e para a avaliação de suas invariâncias no tempo, tornando o investimento em recursos de análise para a plena determinação da estacionariedade normalmente injustificável.

Entretanto, apesar de ser apenas uma etapa de classificação e de apresentar a dificuldade citada, esta análise é necessária e, no âmbito de análise de séries temporais e processamento de sinais, trabalha-se com uma definição alternativa. Essa estratégia restringe a análise da estacionariedade às propriedades estatísticas cujo cálculo é mais simples do que a obtenção das funções de distribuição conjunta de probabilidades.

Estas propriedades são a média, também nomeada na literatura como “valor esperado”, e a função de autocovariância. As equações, discutidas anteriormente no Capítulo 2 ([3]; [4] e [5]; [6]), são reapresentadas a seguir.

$$\mu = \frac{\sum_{i=1}^n z_i}{n} \qquad \gamma(j) = \frac{\sum_{i=1+j}^n (z_i - \mu)(z_{i-j} - \mu)}{n}$$

Uma vez determinados os resultados da média e da autocovariância de diversas ordens, o processo é classificado como “fracamente estacionário” se a sua média e a sua função de autocovariância são independentes do tempo.

Em síntese, as opções de classificação da estacionariedade do processo são:

- Não estacionário: a média ou a autocovariância é dependente do tempo;
- Fracamente estacionário: a média e a autocovariância não dependem do tempo i ou do deslocamento j ;
- Estritamente estacionário: todas as distribuições de probabilidade não são alteradas por translações no tempo.

Estacionariedade por períodos

Adicionalmente, uma série pode ser estacionária por períodos, flutuando ao redor de um nível ou de uma inclinação de curva por um certo período, em seguida alterando para um novo nível ou inclinação. A Figura 11 ilustra este caso na cotação do dólar durante o ano de 2015.

Este tipo de não estacionariedade é chamado homogêneo e, desde que não apresente comportamento explosivo, pode ser transformado em uma série estacionária através da diferenciação sucessiva da série original.

As equações base deste procedimento são apresentadas em [11]; [12].

$$\Delta z_i = z_i - z_{i-1} \qquad \Delta^p z_i = \Delta(\Delta^{p-1} z_i) \qquad [11]; [12]$$

Onde:

Δz_i diferenciação do termo z_i

$\Delta^p z_i$ p -ésima diferenciação de z_i

Após identificada a quantidade de diferenciações p necessária para tornar a série estacionária, define-se a série como “estacionária de ordem p ”⁵. Em geral, é suficiente calcular apenas uma

⁵ A ordem da estacionariedade não tem qualquer paralelo com a ordem da covariância!

ou duas diferenças para eliminar os períodos de estacionariedade. A diferenciação de ordem 1 da cotação do dólar utilizada no exemplo anterior é ilustrada na Figura 12.

Figura 12. Diferenciação da cotação do dólar

Fonte: Autor.

3.4 CLASSES DE PROCESSOS ESTOCÁSTICOS E PROPRIEDADES

Uma vez definida uma série temporal de interesse e sua caracterização como um processo estocástico, a continuidade de sua análise consiste na procura por um modelo para sua representação.

Este modelo deve ter como base uma equação com componentes aleatórias que aproxime o processo estocástico em questão da melhor maneira possível e a partir de apenas uma realização, ou seja, com base na série temporal que foi registrada. Obtido o modelo da série estacionária, é possível realizar a predição estacionária da saída e a geração de cenários através de séries sintéticas.

Para tanto, são utilizadas classes de processos estocásticos, cujas propriedades são bem conhecidas e que permitem a representação da série temporal através de um modelo. As classes mais importantes para este tipo de modelagem são o ruído gaussiano branco e os processos autorregressivo (AR – *Auto Regressive*), de média móvel (MA – *Moving Average*), autorregressivo de média móvel (ARMA – *Auto Regressive Moving Average*) e autorregressivo integrado de média móvel (ARIMA – *Auto Regressive Integrated Moving Average*), descritos a seguir.

Ruído branco

O conceito de ruído branco tem origem econométrica, como em (GUJARATI e PORTER, 2011) e (BUENO, 2011), e é muito presente no estudo de séries temporais, especialmente as estocásticas discretas. O termo se aplica a uma sequência de erros aleatórios, sem qualquer autocorrelação, com média e variância constantes. Por conveniência em sua aplicação, muitos autores impõem média igual a zero, porém esta condição não é de fato obrigatória.

A Figura 13 ilustra a simulação de 100 amostras de um ruído branco com média nula e variância unitária, onde se pode observar as características típicas esperadas, como a estacionariedade e a homogeneidade e independência temporal.

Destaca-se que esta classe de processo, mesmo quando não aplicada isoladamente para algum modelo, tem suas propriedades frequentemente utilizadas na composição de outras classes mais sofisticadas.

AR - Autorregressivo

Segundo (EHLERS, 2009), “processos AR podem ser usados como modelos se for razoável assumir que o valor atual de uma série temporal depende do seu passado imediato mais um erro aleatório”. Em outras palavras, em um processo autorregressivo a sua saída no instante i depende de seus valores nos instantes anteriores $i-1$, $i-2$, ..., $i-k$ e de um ruído branco, como na equação [13].

$$z_i = a_1 z_{i-1} + a_2 z_{i-2} + \dots + a_k z_{i-k} + e_i \quad [13]$$

Onde:

z_i processo autorregressivo de ordem k [ou AR(k)]

e_i ruído branco do instante i

a_i parâmetros do processo AR

A quantidade de coeficientes k necessária para descrever o processo indica com quantos valores anteriores ele tem dependência e é conhecida como ordem do processo. Uma simulação do processo AR de ordem 2, dado pela Equação [14], é ilustrado na Figura 14.

$$z_i = 0,5 z_{i-1} + 0,25 z_{i-2} + e_i \quad [14]$$

Como esperado, observa-se que os valores do processo oscilam em torno do zero e estão compreendidos dentro de um intervalo limitado. Esta característica ocorre pois o ruído branco possui média nula e pode assumir valores positivos e negativos.

MA - Média móvel

O processo MA possui equação semelhante ao AR, porém sua dependência temporal é relacionada apenas aos ruídos passados e não aos valores realizados. Ou seja, sua saída no instante i é uma média móvel do ruído branco que afeta o processo nos instantes anteriores $i-1$, $i-2$, ..., $i-k$ adicionada do erro do instante i , como na Equação [15].

$$z_i = b_1 e_{i-1} + b_2 e_{i-2} + \dots + b_r e_{i-r} + e_i \quad [15]$$

Onde:

z_i processo média móvel de ordem r [ou MA(r)]

b_i parâmetros do processo MA

De forma similar ao processo AR, a quantidade de coeficientes r necessária para descrevê-lo indica com quantos ruídos anteriores ele tem dependência e é conhecida como ordem do

processo. Uma simulação do processo MA de ordem 2, dado pela Equação [16], é ilustrado na Figura 15.

$$z_i = 0,35 e_{i-1} - 0,5 e_{i-2} + e_i \quad [16]$$

O processo MA apresenta características semelhantes ao processo AR simulado anteriormente, porém com maior estabilidade em torno do zero e intervalo de oscilação mais limitado, como é de se esperar para a resposta de média móvel sem influência dos valores passados.

ARMA - Autorregressivo com média móvel

O processo ARMA é a combinação dos processos AR e MA, ou seja, sua saída no instante i depende tanto de seus valores como da média móvel do ruído branco nos instantes anteriores $i-1, i-2, \dots, i-k$. Como sua principal vantagem, (EHLERS, 2009) cita que “pode-se obter uma representação adequada com um número menor de parâmetros” sendo que os processos “ARMA formam uma classe de modelos muito úteis e parcimoniosos para descrever dados de séries temporais”.

A Equação [17] representa um processo do tipo ARMA.

$$z_i = a_1 z_{i-1} + a_2 z_{i-2} + \dots + a_k z_{i-k} + b_1 e_{i-1} + b_2 e_{i-2} + \dots + b_r e_{i-r} + e_i \quad [17]$$

Onde:

z_i processo autorregressivo com média móvel de ordem (k,r) [ou ARMA(k, r)]

Neste caso, a ordem do processo ARMA é dada por um par (k,r) , onde k e r são, respectivamente, o número de parâmetros da parte AR e da parte MA. Um exemplo do processo

ARMA de ordem (2,2), dado pela combinação das Equações [14] e [16], é ilustrado na Figura 16.

$$z_i = 0,5 z_{i-1} + 0,25 z_{i-2} + 0,35 e_{i-1} - 0,5 e_{i-2} + e_i \quad [18]$$

A composição dos processos em equação única provocou a suavização da característica oscilatória dos processos AR e MA quando independentes, mantendo a oscilação em torno do zero e a limitação dos valores dentro de uma faixa controlada. Entretanto, convém ressaltar que os exemplos citados são apenas para ilustração das propriedades de suas classes, não representando o resultado da modelagem de uma série específica e não permitindo comparações mais elaboradas.

Vale destacar a observação de (BOX, JENKINS e REINSEL, 2008) quanto às funções de autocorrelação, que “ficam consideravelmente mais complicadas em processos ARMA. De um modo geral, para um processo ARMA(k,r) estacionário a função de autocorrelação tem um decaimento exponencial ou oscilatório após a defasagem r ”, e a complementação de (EHLERS, 2009), onde “em princípio este resultado pode ser utilizado para auxiliar na determinação da ordem (k,r) do processo, mas na prática pode ser bastante difícil distinguir entre decaimentos exponenciais e oscilatórios através das estimativas destas funções”.

ARIMA - Autorregressivo Integrado com Média Móvel

Os processos estocásticos do tipo ruído branco, AR, MA e ARMA apresentados anteriormente são importantes para a análise de séries temporais por conta de sua estacionariedade. Assim, caso a série temporal analisada seja fracamente estacionária, então as estruturas AR, MA e ARMA são adequadas para representá-la.

Consequentemente, pode-se obter um modelo com alto grau de confiabilidade para gerar estimativas em instantes futuros ou para a criação de séries sintéticas pois, como sua média e função de covariância não dependem do tempo, o processo conserva importantes propriedades estatísticas, característica altamente desejada para se realizar estimativas de longo prazo.

Porém, caso a série temporal registrada não seja estacionária, essas estruturas não são adequadas para descrevê-la diretamente. Neste caso, frequentemente se recorre à classe ARIMA, que é um processo construído com base no ARMA e integrado com técnicas de diferenciação da série temporal original.

Assim como apresentado anteriormente na Seção 3.3, o processo não estacionário é diferenciado sucessivas vezes, modificando a série original a fim de se obter uma série estacionária, na qual é então aplicado o modelo ARMA.

A estrutura dos processos é semelhante, tendo apenas a substituição da variável do processo z por sua diferencial. A equação [19] representa a estrutura de processo ARIMA.

$$z_i = a_1 \Delta^p z_{i-1} + a_2 \Delta^p z_{i-2} + \dots + a_k \Delta^p z_{i-k} + b_1 e_{i-1} + b_2 e_{i-2} + \dots + b_r e_{i-r} + e_i \quad [19]$$

Onde Δz_i é obtido a partir da diferenciação das Equações [11]; [12].

$$\Delta z_i = z_i - z_{i-1} \qquad \Delta^p z_i = \Delta(\Delta^{p-1} z_i)$$

A ordem de um processo ARIMA é dada por uma trinca (k,r,p) , onde k e r são as quantidades de parâmetros ARMA e p é o número de diferenciações.

Por fim, para se obter a saída do processo ARIMA é necessário integrar a saída tantas vezes quanto a série temporal original fora diferenciada para se obter a série temporal estacionária.

3.5 MÉTODO DE BOX-JENKINS

O método de Box-Jenkins é um recurso bastante empregado nas áreas de econometria (ENDERS, 2014), modelagem de sistemas e controle (LJUNG, 1998), finanças e meteorologia para a obtenção de modelos de séries temporais. A versatilidade de aplicações deste método provém de seu caráter iterativo para a obtenção dos modelos, ou seja, a partir do registro da série temporal, uma sequência de instruções é executada e ao final desta sequência um modelo

da série temporal é obtido. O objetivo do método de Box-Jenkins é obter um modelo parcimonioso que represente adequadamente a série temporal em questão com um número mínimo de parâmetros.

Teoricamente, qualquer uma das estruturas apresentadas na Seção 3.4 pode representar adequadamente um processo estacionário com um número infinito de parâmetros. Porém, na prática, mesmo um modelo que represente perfeitamente a série com infinitos parâmetros perde a sua utilidade, pois o esforço computacional requerido é muito grande, dificultando a realização de previsões ou a geração de séries sintéticas. Portanto, torna-se ideal buscar um modelo com um número finito – e reduzido – de parâmetros baseado no princípio da parcimônia.

Este princípio afirma que um modelo com muitos parâmetros pode ser desnecessário para representar uma série temporal, mas que um modelo com poucos parâmetros pode levar a pobres estimativas da saída da série e, por consequência, gerar séries sintéticas que não condizem com a realidade. O objetivo então é obter modelos parcimoniosos, com um número finito e adequado de parâmetros, que se mostre adequado para representar a série temporal após a fase de validação.

Assim, (BOX, JENKINS e REINSEL, 2008) propõe um método iterativo para a obtenção de modelos estocásticos para as séries temporais a partir de estruturas de modelos pré-definidas, que segue o fluxograma representado na Figura 17.

Figura 17. Fluxograma representativo do método de Box-Jenkins

Fonte: Autor.

O processo é iniciado com a escolha de uma das classes de modelo apresentadas anteriormente, começando com o menor número de parâmetros possível. Estes parâmetros são então estimados a partir de métodos como mínimos quadrados (MMQ), variáveis instrumentais (IV) e máxima verossimilhança (MMV). A principal diferença entre estes métodos de estimativa de parâmetros

encontra-se na formulação da função objetivo a ser otimizada (minimizada ou maximizada, conforme o caso).

Após a etapa de estimativa dos parâmetros, o modelo é validado através de um processo de verificação de suas propriedades estatísticas, que devem corresponder às da série temporal registrada. Além disso, o modelo obtido pode ser simulado a fim de verificar se os cenários sintéticos por ele gerados são consistentes com a realidade.

Caso o modelo seja considerado apto, o processo chega ao fim. Caso contrário, aumenta-se o número de parâmetros da estrutura escolhida e repete-se o processo até que um modelo válido seja obtido.

Este é um processo iterativo caracterizado por tentativa e erro. Após a obtenção de um modelo a partir de uma das estruturas pré-definidas, convém testar outras estruturas para verificar a quantidade de parâmetros utilizada e sua aderência à representação desejada.

3.6 GERAÇÃO DE CENÁRIOS DE RADIAÇÃO SOLAR

A aplicação da geração de cenários para análise da radiação solar com base em modelos como os citados é um tema bastante estudado, porém sua aplicação (seja do insumo ou da geração elétrica) com viés probabilístico é menos comum e tem aplicações específicas. A seguir são apresentados os principais trabalhos encontrados sobre o tema, identificando os avanços e tendências na obtenção de modelos que possam contribuir com suas experiências e desenvolvimentos.

Em nível nacional, uma das poucas publicações existentes é a tese de doutorado de (SIQUEIRA, 2008), que estuda a “geração de séries temporais sintéticas da irradiação solar diária, que reproduzam as principais características estatísticas das séries históricas, viabilizando a simulação e a avaliação de desempenho dos sistemas solares submetidos a regime de longo prazo”. No estudo, o autor avaliou o modelo ARMA e a aplicação de redes neurais para a geração de séries sintéticas, realizando interpolações temporais e espaciais em localidades do Estado de Pernambuco.

Convém destacar que a motivação do trabalho foi de “aprimorar e disponibilizar a base de dados sobre o recurso solar” ampliando “o conhecimento referente à extensão temporal e à distribuição espacial da radiação”. Ou seja, a proposta foi a geração sintética de dados para localidades carentes de informações solarimétricas, diferenciando-se desta tese, que propõe a geração de cenários sintéticos de radiação, para localidades com dados pré-existentes e com foco na previsão e análise estocástica da geração futura para aplicações no planejamento e operação energética.

Em outro trabalho desenvolvido no Brasil, (NOBRE, 2015), com objetivo de estudar previsões de curto prazo para antecipação de tempestades e com estudo de caso em Singapura, testou diferentes métodos, como persistência (repetição do passado), ARIMA e redes neurais, simulando a performance de sistemas fotovoltaicos em diversas localidades do país. Para a comparação dos resultados, foram utilizados testes com MMQ e contra persistência (razão entre o MMQ do registro e o MMQ da persistência). Como diferencial em suas análises, o trabalho usa variáveis meteorológicas, como umidade e pressão do ar, registradas em diversas estações espalhadas pela região.

O autor ainda destaca a ausência de estudos na região entre trópicos, motivada pela baixa representatividade de usinas instaladas, e que apresenta clima bastante dinâmico e com variações bruscas, tornando difícil a previsão da radiação.

Em estudos internacionais observa-se a preocupação com a predição da radiação há algumas décadas, como em (GOH e TAN, 1977), que apresenta um modelo ARMA para modelagem do nível de radiação. O trabalho foi inovador no sentido de trazer ferramentas estatísticas para a predição da radiação solar, até então realizada através de gráficos e equações matemáticas baseadas em médias de longo prazo.

Em seguida, (MUSTACCHI, CENA e ROCCHI, 1979) apresentam outro modelo ARMA, porém combinado com uma matriz de transição de Markov, melhorando as predições realizadas. Em (GORDON e REDDY, 1988), além de analisar diferentes modelos baseados em séries temporais, são também discutidas técnicas de validação baseadas em ferramentas estatísticas, especificamente os intervalos de confiança.

O método de Box-Jenkins da Seção 3.5 é aplicado em (SULAIMAN, *et al.*, 1997) para a obtenção de modelos do tipo AR e ARMA, com base diária e para localidades na Malásia sujeitas a alta variabilidade do recurso solar. Este trabalho também avalia técnicas para a validação de modelos de séries temporais a partir da análise de resíduos.

Um estudo mais recente (FORTUNA, NUNNARI e GALLO, 2014) apresenta metodologia para a obtenção da parte determinística da radiação solar, como procedimento anterior à aplicação do método de Box-Jenkins, preparando dados para a fase de identificação do modelo.

Com enfoque elétrico, (INMAN, PEDRO e COIMBRA, 2013) apresenta diversos métodos para a predição de recursos solares e da potência de saída de usinas baseados em séries temporais e em redes neurais artificiais. Neste trabalho são apresentadas as aplicações de modelos AR, MA, ARMA e ARIMA para a modelagem de séries de radiação solar. Além de incluir diversos métodos de modelos lineares e não lineares, o trabalho também apresenta critérios utilizados na literatura para a avaliação e validação de modelos.

Ainda neste enfoque, (LI, SU e SHU, 2014) apresenta um modelo ARMAX (*Auto Regressive Moving Average with exogenous inputs*) para prever a potência de saída de uma fazenda solar. O ARMAX é muito aplicado na identificação de sistemas como alternativa aos modelos de séries temporais do tipo ARIMA. Nestes modelos com entradas exógenas, outras variáveis que influenciam na variável de saída podem ser utilizadas como entradas. Se desejável, este

conceito pode ser extrapolado para as séries de radiação para verificar como outras variáveis medidas pela estação estão correlacionadas com a radiação em si.

Por fim, à margem dos modelos convencionais, em trabalho realizado na Itália, com base em quatro estações meteorológicas e considerando restrições da disponibilidade de dados, (AMATO, *et al.*, 1986) propôs a modelagem da radiação através de séries de Fourier e concluíram que as variabilidades da média e da variância são explicadas, respectivamente, em 95 % e 65 % por séries de primeira e segunda ordem. Essa pesquisa, apesar de possuir resultados ainda válidos, pode ser reavaliada e aprimorada com o atual volume de dados disponíveis e com a melhoria da capacidade de processamento.

4 ESTIMATIVA DE PARÂMETROS DOS MODELOS

A estimativa dos modelos pelo método de Box-Jenkins utiliza uma metodologia iterativa para a obtenção do conjunto de coeficientes e da variância do ruído branco presentes na estrutura de modelo autorregressiva através de um método de estimativa de parâmetros.

Essa estimativa é composta basicamente de duas etapas: o pré-processamento dos dados e a estimativa dos coeficientes e da variância do ruído presentes na equação da estrutura selecionada.

A etapa de pré-processamento dos dados consiste na retirada das tendências e médias do conjunto de dados. Isto se faz necessário pois as estruturas utilizadas (AR, ARMA, etc.) são específicas para representação de processos estocásticos que tenham média nula. No processo de geração das séries sintéticas as tendências e médias são reinseridas nos dados gerados para que a série seja devidamente representada.

A etapa de estimativa dos coeficientes e da variância do ruído branco presentes nas equações [13], [15] e [17] consiste na obtenção do modelo em si. Para isto, um método de estimativa de parâmetros deve ser aplicado. Diversos métodos foram propostos na literatura como o método de Yule-Walker (BOX e JENKINS, 1976) (MADSEN, 2007) e método dos mínimos quadrados (LJUNG, 1998).

O método de Yule-Walker possui como principal característica a simplicidade de implementação, pois é baseado em equações recursivas simples e propriedades estatísticas elementares, como a função de autocorrelação e a função de autocovariância.

Por outro lado, o método dos mínimos quadrados é bastante versátil, sendo baseado em um problema de otimização que visa minimizar a somatória do erro quadrático entre os valores medidos e estimados da série. Sua solução e implementação possui um grau de dificuldade superior ao método de Yule-Walker, porém diversos pacotes computacionais estatísticos possuem as funções já implementadas.

Este capítulo apresenta inicialmente, em detalhe, o processo completo de estimativa de modelos autorregressivos para a geração de séries sintéticas de radiação solar para a periodicidade diária. Em seguida são apresentados os resultados complementares da aplicação do mesmo processo para os períodos horário e mensal.

4.1 ANÁLISE DOS DADOS

Considerações

Os dados de radiação solar utilizados provêm de sistemas de estimativa por satélites. Estes dados são estimados por algoritmos meteorológicos e possuem resolução de três quilômetros. Todos os dados de radiação apresentados a seguir são reais e foram adquiridos em sua forma original da empresa espanhola Vortex para duas localidades do nordeste brasileiro.

Destaca-se, nesta fase da pesquisa, de geração de séries de radiação para diferentes períodos e para aplicações de longo prazo, a ausência de estudos equivalentes na literatura, que é restrita a aplicações com horizonte de poucas horas ou, no máximo, poucos dias, em trabalhos como (GOH e TAN, 1977), (GORDON e REDDY, 1988) e (SULAIMAN, *et al.*, 1997). A presente pesquisa pretende, entre outras contribuições, completar esta lacuna de conhecimento, estudando a geração de séries para prazos e períodos mais longos.

Análise preliminar

O conjunto de dados consiste de uma série de radiação horária contínua e com duração de 20 anos. Para aplicações com base diária, por exemplo, isto representa uma série com 7.300 dados. Entretanto, a utilização desta série única para a modelagem temporal pode reduzir, ou até mesmo eliminar, informações relevantes relacionadas à sazonalidade, fenômenos climáticos típicos, entre outros. Logo propôs-se a obtenção de uma série para a estimativa da energia diária total incidente para cada mês, semelhante ao realizado para as séries hidrológicas (KELMAN, 1987), (MACEIRA, KELMAN e DAMÁZIO, 1987) e (DAMÁZIO, *et al.*, 1983) e também para as séries temporais utilizadas na predição da geração eólica (MACEIRA, *et al.*). Esta tentativa é inovadora em duas frentes: no tratamento dos dados com periodicidade diária, e não mensal, como no caso dos trabalhos citados; e na remoção da sazonalidade anual de modelos diários pela separação da série original por séries mensais.

Uma vez levantada a possibilidade da existência de sazonalidades, foi realizada análise inicial com a verificação das principais características dos conjuntos de dados diários, organizados por mês: valores médio, máximo, mínimo e desvio padrão. O resultado desta análise preliminar é ilustrado na Figura 18.

Figura 18. Análise preliminar de média, valor máximo, média dos mínimos e desvio padrão para cada mês.

Fonte: Autor.

A figura permite verificar a diferença entre os valores de energia incidente em cada mês e a variabilidade destes dados: os meses de julho e dezembro possuem, respectivamente, a menor e a maior variância do conjunto de dados. Além disso, pode-se visualizar a diferença entre os valores máximo e mínimo registrados para cada mês. Meses como janeiro e dezembro apresentam grandes diferenças entre os valores extremos enquanto que os meses de junho e julho apresentaram as menores diferenças.

A oposição dos meses de julho e dezembro no desvio padrão é detalhada na Figura 19 e na Figura 20, onde pode-se visualizar dez séries registradas de radiação diária. Observa-se, como esperado, que as séries do mês de julho são mais comportadas, provavelmente por conta do período de estiagem típico da época, em oposição ao período de chuvas do final do ano.

Figura 19. Séries de energia diária incidente para o mês de julho.

Fonte: Autor.

Figura 20. Séries de energia diária incidente para o mês de dezembro.

Fonte: Autor.

Os mesmos dados também podem ser apresentados ordenados pela energia diária total, como na Figura 21 e na Figura 22. Neste caso, além de se observar a variação da radiação mais acentuada no mês de dezembro, também é possível notar a distribuição da quantidade de dias abaixo e acima da média bastante estável em julho e variável em dezembro.

Figura 21. Verificação da distribuição da energia total incidente para o mês de julho.

Figura 22. Verificação da distribuição da energia total incidente para o mês de dezembro.

A análise de distribuição por este método, apesar de não quantificada numericamente, permite rápida avaliação do comportamento das séries sintéticas a serem geradas e serão utilizadas oportunamente de forma complementar.

4.2 PRÉ-PROCESSAMENTO DOS DADOS

O pré-processamento dos dados é a primeira etapa necessária para a estimativa dos modelos de uma série temporal baseados na estrutura autorregressiva. Essa necessidade existe pois a estrutura AR representa apenas processos estocásticos com média nula (MADSEN, 2007) e, como pode-se notar na Figura 18, os valores de radiação para todos os meses oscilam em torno de uma média diferente de zero.

Para contornar essa característica da série, este trabalho realizou uma avaliação adicional de três formas distintas de pré-processamento dos dados para a retirada da tendência: pela média diária, pelo máximo diário e pela média mensal. Os três métodos são apresentados a seguir.

Método 1: retirada de tendência pela média diária

O primeiro método aplicado assume como tendência a média registrada nas 20 séries para cada dia. Escrevendo matematicamente, a tendência foi obtida da equação seguinte.

$$t_{ij} = \frac{1}{N_y} \sum_{k=1}^{N_y} x(i, j, k) \quad [20]$$

Onde:

t_{ij}	é a tendência do dia i para o mês j
i	dia do mês, variando entre 1 e 28, 30 ou 31
j	mês, variando entre 1 e 12
N_y	número total de séries (no caso deste estudo, N_y tem valor fixo igual a 20 e representa o número de séries temporais obtidas pelo sistema de satélites)
$x(i, j, k)$	valor de energia diária total incidente para o dia i do mês j no ano k .

Para a retirada da tendência e média simultaneamente, foi aplicado o procedimento descrito pela equação [21], sugerido por (GORDON e REDDY, 1988).

$$\tilde{x}(i, j, k) = \frac{x(i, j, k)}{t_{ij}} - 1 \quad [21]$$

Nesta equação a remoção de média e tendência da energia diária total incidente o dia i do mês j no ano k resulta na variável $\tilde{x}(i, j, k)$.

Este novo conjunto de dados após a remoção de tendência corresponde a uma nova variável, para a qual foi dado o nome de índice de limpidez do céu (GORDON e REDDY, 1988). Este índice busca representar o quanto a energia diária total incidente aproximou-se da média diária de um dado mês.

Índices iguais a zero significam que o valor medido é igual à média diária para o período enquanto que índices positivos significam que os valores registrados são maiores do que a média diária e índices negativos significam que os valores registrados são menores do que a média diária calculada pela equação [20].

Os dados originais, as tendências e os dados processados pela equação [21] são ilustrados a seguir, para os meses de julho e dezembro, respectivamente na Figura 23 e na Figura 24.

Figura 23. Tendência e remoção de tendência para o mês de julho

Fonte: Autor.

Nota-se que os dados após a remoção de tendência passam a oscilar em torno de zero. Ao se calcular a média após o pré-processamento, para os dois meses, de fato obtém-se média nula. A mesma análise se faz valer para a Figura 24, que representa o processo de extração de média e tendência para o mês de dezembro.

Figura 24. Tendência e remoção de tendência para o mês de dezembro

Fonte: Autor.

As tendências e os dados processados pelo método 1 (média diária) para todos os meses do ano são apresentados no Anexo I.

Método 2: retirada da tendência pelo máximo diário

Uma proposta alternativa à equação [20] implementada nesta etapa é considerar como tendência o valor máximo da energia diária total incidente. Para isso, a tendência é obtida através da equação [22].

$$t_{ij} = \max[x(i, j, k)] \quad [22]$$

Assim como no método anterior, a tendência é então retirada do conjunto de dados utilizando a equação [21].

Como ilustração, os dados dos meses de julho e dezembro são apresentados nas figuras a seguir. As tendências para todos os meses estão disponíveis em Anexo II.

Figura 25. Retirada de tendência pelo máximo diário, mês de julho

Fonte: Autor.

Figura 26. Retirada de tendência pelo máximo diário, mês de dezembro

Fonte: Autor.

Nota-se que os dados retirados de tendência pela aplicação do segundo método não oscilam em torno de zero. Assim, faz-se necessária uma etapa adicional, onde o conjunto de dados é subtraído de sua média. O resultado da etapa adicional encontra-se na Figura 27.

Método 3: retirada da tendência pela média mensal

Por fim, o terceiro método aplicado consiste em utilizar as médias mensais expostas na Figura 18 como a tendência do conjunto de dados. Neste caso, após a retirada da tendência pela aplicação da equação [21], obtém-se os dados ilustrados na Figura 28.

Nota-se na Figura 28 que a retirada de tendência pela média mensal faz com que o conjunto de dados resultantes oscile em torno do zero, portanto não se faz necessária uma etapa adicional como no método 2.

As tendências calculadas para todos os meses são apresentadas no Anexo III.

Comentários

Após a aplicação de um dos três métodos descritos, o conjunto de dados de índice de limpidez do céu está pronto para ser processado e modelado por séries temporais. Destaca-se, entretanto, que os modelos obtidos representam o processo estocástico do índice de limpidez e não da radiação incidente. Para tal resultado, deve ser percorrido o caminho inverso, reintroduzindo-se a tendência característica dos dados.

4.3 GERAÇÃO DE SÉRIES SINTÉTICAS PELO MÉTODO DE YULE-WALKER

Método

O método de Yule-Walker consiste na obtenção dos parâmetros do modelo autorregressivo apresentado pela equação [13] através de estimativas das funções de autocovariância e autocorrelação do processo estocástico dadas pelas equações [9] e [10], respectivamente reproduzidas abaixo:

$$z_i = a_1 z_{i-1} + a_2 z_{i-2} + \dots + a_k z_{i-k} + e_i$$

$$\gamma(j) = \frac{\sum_{i=1+j}^n (z_i - \mu)(z_{i-j} - \mu)}{n}$$

$$\rho(j) = \frac{\gamma(j)}{\sigma^2}$$

Inicialmente, a primeira equação é multiplicada por z_{i-j} em ambos os lados, obtendo:

$$z_i z_{i-j} = a_1 z_{i-1} z_{i-j} + a_2 z_{i-2} z_{i-j} + \dots + a_k z_{i-k} z_{i-j} + e_i z_{i-j} \quad [23]$$

Em seguida, ao se retirar o valor esperado e dividir ambos os lados por $\gamma(0)$, são obtidas as equações [24] e [25], conhecidas por equações de Yule-Walker (BOX e JENKINS, 1976) (MADSEN, 2007).

$$\rho_j = a_1\rho_{j-1} + a_2\rho_{j-2} + \dots + a_n\rho_{j-n}, \text{ para } j > 0; \quad [24]$$

$$\rho_0 = a_1\rho_1 + a_2\rho_2 + \dots + a_n\rho_n + \sigma_e^2, \text{ para } j = 0. \quad [25]$$

A estrutura representada pela equação [24] possui ordem n , assim ela é escrita n vezes para que se obtenha um sistema linear com n equações e n incógnitas, sendo estas incógnitas os parâmetros a_i . Resolvido o sistema e estimados os parâmetros do processo, a equação [25] é resolvida para se obter a variância do ruído, σ_e^2 .

Estimativa dos parâmetros dos modelos

A análise do método de Yule-Walker foi realizada através da estimativa do modelo AR para os três métodos de remoção da tendência descritos na seção 4.2 e para cada mês do ano. As estimativas foram realizadas para o modelo de ordem 1.

A escolha da ordem foi realizada após avaliação da influência do tamanho do conjunto de dados na estimativa dos parâmetros, onde se observou que os coeficientes de ordem 2 ou superior são pouco significativos na modelagem. A avaliação citada é apresentada no Anexo IV.

A partir da conclusão de que o coeficiente a_1 é o único significativo, a equação do modelo autorregressivo a ser estimada para cada um dos métodos descritos é dada pela equação [26].

$$z_i = a_1 z_{i-1} + e_i \quad [26]$$

Nas tabelas a seguir estão sumarizados os parâmetros dos modelos autorregressivos de primeira ordem obtidos, por mês, para os três métodos de retirada de tendência discutidos anteriormente: média diária, máxima diária e média mensal.

Tabela 2. Modelos AR, retirada de tendência pela média diária

Método 1		
Mês	a_1	Variância
Janeiro	-0,5922	0,0147
Fevereiro	-0,3980	0,0173
Março	-0,5311	0,0146
Abril	-0,5533	0,0160
Maio	-0,5336	0,0176
Junho	-0,3655	0,0234
Julho	-0,3610	0,0170
Agosto	-0,4039	0,0121
Setembro	-0,5023	0,0092
Outubro	-0,5414	0,0100
Novembro	-0,5456	0,0167
Dezembro	-0,6206	0,0162

Tabela 3. Modelos AR, retirada de tendência pela máxima diária

Método 2		
Mês	a_1	Variância
Janeiro	-0,5904	0,0105
Fevereiro	-0,3958	0,0126
Março	-0,5131	0,0106
Abril	-0,5569	0,0108
Maio	-0,5148	0,0118
Junho	-0,3486	0,0152
Julho	-0,3489	0,0117
Agosto	-0,4181	0,0091
Setembro	-0,4932	0,0075
Outubro	-0,5334	0,0079
Novembro	-0,5396	0,0120
Dezembro	-0,6142	0,0115

Tabela 4. Modelos AR, retirada de tendência pela média mensal

Método 3		
Mês	a_1	Variância
Janeiro	-0,6038	0,0150
Fevereiro	-0,4152	0,0179
Março	-0,5461	0,0150
Abril	-0,5747	0,0167
Maio	-0,5419	0,0186
Junho	-0,3579	0,0243
Julho	-0,3961	0,0182
Agosto	-0,4903	0,0138
Setembro	-0,5320	0,0102
Outubro	-0,5344	0,0102
Novembro	-0,5429	0,0170
Dezembro	-0,6137	0,0166

Destaca-se que, apesar dos valores dos coeficientes encontrados para os modelos de cada mês serem semelhantes, sua comparação direta não possui significado, pois as tendências foram retiradas por métodos diferentes.

Nos próximos itens serão apresentadas as séries sintéticas geradas para a energia diária total pelos modelos autorregressivos estimados, a verificação da distribuição das séries sintéticas e sua consistência com a distribuição das séries registradas.

Análise da distribuição das séries

A análise da distribuição foi realizada por procedimento semelhante ao utilizado na Figura 22, com a apresentação comparativa dos valores das séries históricas e das séries sintéticas geradas em cada método testado.

Para todos os métodos partiu-se de um valor inicial tal que o correspondente índice de limpidez do céu seja zero. Ou seja, todas as séries discutidas neste tópico partiram do mesmo valor inicial. Por conta deste procedimento, um dado a mais foi gerado por série sintética e o valor inicial foi omitido dos gráficos.

As comparações dos métodos de retirada de tendência para a energia total diária, para todos os meses, estão registradas nas figuras 29 a 40. Nessas figuras está plotada à esquerda uma série sintética gerada em cada um dos métodos propostos, a média diária registrada dos 20 anos e os valores máximo e mínimo do histórico. À direita encontram-se todos os 20 conjuntos de dados históricos em curvas amarelas e uma série sintética gerada em cada método avaliado (todos ordenados de forma crescente) assim como os valores máximo e mínimo do histórico para verificação da consistência dos dados.

Como exemplo, na Figura 29 (a) pode-se notar que as séries ultrapassam algumas vezes o limite máximo do histórico. Essa característica é esperada, apesar de indesejável, e deve ocorrer um número mínimo de vezes em um modelo que represente bem a série original.

Na Figura 29 (b), à direita, pode-se verificar a distribuição das séries sintéticas geradas em conjunto às registradas, sendo que, assim como no caso anterior, as séries geradas devem preferencialmente ter comportamento semelhante ao histórico, com um número mínimo de transgressões.

Analisando-se todos as séries calculadas, o método 2 se destacou negativamente por gerar frequentemente valores próximos ao máximo, não representando bem toda a faixa de valores esperados. Nas avaliações seguintes este método será descartado.

Figura 29. Métodos de geração de séries sintéticas da radiação diária (janeiro)

Fonte: Autor.

Figura 30. Métodos de geração de séries sintéticas da radiação diária (fevereiro)

Fonte: Autor.

Figura 31. Métodos de geração de séries sintéticas da radiação diária (março)

Fonte: Autor.

Figura 32. Métodos de geração de séries sintéticas da radiação diária (abril)

Figura 33. Métodos de geração de séries sintéticas da radiação diária (maio)

Figura 34. Métodos de geração de séries sintéticas da radiação diária (junho)

Figura 35. Métodos de geração de séries sintéticas da radiação diária (julho)

Fonte: Autor.

Figura 36. Métodos de geração de séries sintéticas da radiação diária (agosto)

Fonte: Autor.

Figura 37. Métodos de geração de séries sintéticas da radiação diária (setembro)

Fonte: Autor.

Figura 38. Métodos de geração de séries sintéticas da radiação diária (outubro)

Figura 39. Métodos de geração de séries sintéticas da radiação diária (novembro)

Figura 40. Métodos de geração de séries sintéticas da radiação diária (dezembro)

4.4 GERAÇÃO DE SÉRIES SINTÉTICAS PELO MÉTODO DOS MÍNIMOS QUADRADOS

Método

O método dos mínimos quadrados (MMQ) consiste na obtenção dos parâmetros de estruturas autorregressivas através da minimização de uma função objetivo que representa o erro quadrático entre os dados medidos e os estimados pela série temporal (BOX e JENKINS, 1976).

A função objetivo a ser minimizada pode ser escrita como:

$$J = \sum_{i=1}^N (y_{i, med} - y_{i, est})^2 \quad [27]$$

Onde:

J função objetivo a ser minimizada

N quantidade total de registros

$y_{i, med}$ valor medido de radiação solar diária

$y_{i, est}$ valor estimado por uma das estruturas descritas pelas equações [13], [15], [17] e [19].

Estimativa dos modelos

O MMQ foi utilizado para obter modelos para as séries de cada mês para cada uma das quatro estruturas expostas no capítulo 3. Assim, com os dois métodos para pré-filtragem dos dados podem ser obtidos oito modelos distintos para cada mês do ano, resultando em 96 modelos de séries temporais.

Assim como na avaliação do método de Yule-Walker, foram adotados apenas modelos de primeira ordem. Os resultados dos coeficientes e da variância obtidos em cada estimativa são apresentados nas tabelas 5 a 8.

Tabela 5. Modelos para geração de séries sintéticas com estrutura AR

	Modelos AR			
	Método 3		Método 1	
	a1	Variância	a1	Variância
Janeiro	-0,60	643630,00	-0,59	0,0147
Fevereiro	-0,42	755240,00	-0,40	0,0178
Março	-0,55	582690,00	-0,53	0,0146
Abril	-0,57	548050,00	-0,56	0,0161
Maio	-0,54	454020,00	-0,53	0,0177
Junho	-0,38	504640,00	-0,39	0,0231
Julho	-0,40	431280,00	-0,36	0,0170
Agosto	-0,49	444060,00	-0,40	0,0121
Setembro	-0,56	419880,00	-0,54	0,0088
Outubro	-0,53	483650,00	-0,54	0,0100
Novembro	-0,56	752340,00	-0,56	0,0166
Dezembro	-0,61	713140,00	-0,62	0,0162

Tabela 6. Modelos para geração de séries sintéticas com estrutura ARMA

	Modelos ARMA					
	Método 3			Método 1		
	a1	c1	Variância	a1	c1	Variância
Janeiro	-0,66	-0,09	644350,00	-0,68	-0,13	0,0147
Fevereiro	-0,35	0,08	763780,00	-0,31	0,10	0,0180
Março	-0,63	-0,11	582160,00	-0,60	-0,10	0,0146
Abril	-0,72	-0,23	547000,00	-0,73	-0,26	0,0160
Maio	-0,58	-0,06	455170,00	-0,56	-0,04	0,0177
Junho	-0,36	0,02	506820,00	-0,39	0,00	0,0232
Julho	-0,44	-0,05	433280,00	-0,38	-0,02	0,0170
Agosto	-0,56	-0,10	444910,00	-0,43	-0,03	0,0122
Setembro	-0,57	-0,02	421250,00	-0,52	0,02	0,0088
Outubro	-0,58	-0,07	485850,00	-0,59	-0,07	0,0101
Novembro	-0,77	-0,32	738180,00	-0,79	-0,36	0,0162
Dezembro	-0,69	-0,12	713150,00	-0,70	-0,13	0,0161

Tabela 7. Modelos para geração de séries sintéticas com estrutura ARI

	Modelos ARI			
	Método 3		Método 1	
	a1	Variância	a1	Variância
Janeiro	0,24	759000,00	0,25	0,0173
Fevereiro	0,27	975190,00	0,28	0,0232
Março	0,28	689440,00	0,28	0,0174
Abril	0,30	636930,00	0,30	0,0188
Maio	0,26	548060,00	0,25	0,0215
Junho	0,31	662400,00	0,30	0,0302
Julho	0,32	556540,00	0,33	0,0223
Agosto	0,29	545780,00	0,31	0,0156
Setembro	0,23	511370,00	0,23	0,0109
Outubro	0,26	583510,00	0,26	0,0120
Novembro	0,32	869750,00	0,32	0,0191
Dezembro	0,26	825800,00	0,26	0,0186

Tabela 8. Modelos para geração de séries sintéticas com estrutura ARIMA

	Modelos ARIMA					
	Método 3			Método 1		
	a1	c1	Variância	a1	c1	Variância
Janeiro	-0,57	-0,98	644940,00	-0,54	-0,97	0,0147
Fevereiro	-0,42	-0,95	801280,00	-0,39	-0,95	0,0188
Março	-0,52	-0,98	584380,00	-0,50	-0,98	0,0146
Abril	-0,49	-0,93	557350,00	-0,45	-0,92	0,0162
Maiο	-0,54	-0,99	461130,00	-0,54	-1,00	0,0178
Junho	-0,37	-0,98	513170,00	-0,38	-0,98	0,0236
Julho	-0,40	-0,97	447640,00	-0,36	-0,98	0,0175
Agosto	-0,50	-0,99	453300,00	-0,40	-0,97	0,0125
Setembro	-0,58	-1,00	422320,00	-0,54	-0,99	0,0089
Outubro	-0,54	-0,98	499190,00	-0,55	-0,97	0,0104
Novembro	-0,48	-0,96	752780,00	-0,48	-0,95	0,0167
Dezembro	-0,60	-0,97	729890,00	-0,61	-0,98	0,0165

Análise da distribuição das séries

A análise das séries modeladas através do método dos mínimos quadrados seguiu o mesmo procedimento realizado no método de Yule-Walker, com comparações entre as séries geradas com remoção de tendência pela média diária e pela média mensal, para cada uma das quatro estruturas apresentadas nas tabelas.

As figuras 41 a 44 apresentam os gráficos do mês de janeiro para as quatro estruturas. O conjunto completo das séries é apresentado no Anexo V.

Figura 41. Métodos de geração de séries sintéticas de estrutura AR (MMQ), para a energia total diária (janeiro)

Fonte: Autor.

Figura 42. Métodos de geração de séries sintéticas de estrutura ARMA (MMQ) para a energia total diária (janeiro)

Fonte: Autor.

Figura 43. Métodos de geração de séries sintéticas de estrutura ARI (MMQ) para a energia total diária (janeiro)

Fonte: Autor.

Figura 44. Métodos de geração de séries sintéticas de estrutura ARIMA (MMQ) para a energia total diária (janeiro)

Fonte: Autor.

Comentários

A análise dos resultados completos dos modelos de primeira ordem estimados com o método dos mínimos quadrados permite observar:

- Estrutura AR - o modelo é adequado para representar todos os meses do ano com os dois processos de pré-filtragem;
- Estrutura ARMA - o modelo é adequado para representar todos os meses do ano com os dois processos de pré-filtragem, porém verifica-se maior oscilação nas séries se comparadas com as da estrutura AR. Isso ocorre pois em um modelo com estrutura ARMA a dependência do ruído para gerar o valor sintético é maior, dado que se necessita do valor atual do ruído e também de seu valor anterior;
- Estrutura ARI - o modelo é adequado para representar todos os meses do ano com os dois processos de pré-filtragem. Neste caso ocorre o efeito inverso, onde as oscilações entre valores consecutivos diminuem se comparadas às séries de estrutura AR. Essa característica é devida ao termo integrador aos modelos; e
- Estrutura ARIMA – com exceção ao mês de janeiro (Figura 44), os modelos podem ser utilizados para a geração de séries sintéticas do total de energia diário. Neste mês houve a geração de pelo menos quatro valores acima do máximo registrado em ambos os métodos. Para o mês de outubro a série gerada a partir do método 3 também apresentou inconsistências.

4.5 MODELO PARA GERAÇÃO DE SÉRIES DE RADIAÇÃO MENSAL

Os métodos para a geração de séries de radiação diária discutidos anteriormente foram reaplicados a fim de se obter uma única equação responsável por gerar a radiação com periodicidade mensal.

As simulações foram realizadas com as estruturas AR, ARI, ARMA e ARIMA e para os métodos de remoção de tendência 1 e 3, estimando oito equações cujos parâmetros são apresentados na Tabela 9.

Tabela 9. Parâmetros dos modelos anuais

Estrutura	Método 1			Método 3		
	a_1	c_1	variância	a_1	c_1	variância
AR	-0,260	-	0,00297	-0,655	-	369220000
ARI	0,461	-	0,00369	-0,005	-	439410000
ARMA	-0,59	-0,35	0,00299	-0,58	0,14	372940000
ARIMA	-0,22	-0,92	0,00317	-0,70	-0,97	397960000

O resultado mais expressivo é a variância da pré-filtragem pelo método 3, que para qualquer das estruturas tem valor bastante elevado. A causa é justamente devida ao uso da técnica de remoção da tendência, que utiliza a média mensal e acaba por não considerar qualquer dos efeitos da sazonalidade, incorporando-os na própria variância. No caso do método 1, a remoção de tendência absorve as tipicidades da curva e o modelo de estimativa se torna mais limpo e permite estimativas mais precisas.

As figuras 45 a 48 apresentam uma série sintética gerada em cada estrutura e método de pré-filtragem, comparadas à curva média do histórico. Essa simples análise gráfica também evidencia a inadequabilidade do método para a modelagem da radiação com periodicidade mensal.

Figura 46. Séries sintéticas da radiação mensal a partir da estrutura ARI

Fonte: Autor.

Figura 47. Séries sintéticas da radiação mensal a partir da estrutura ARMA

Fonte: Autor.

Figura 48. Séries sintéticas da radiação mensal a partir da estrutura ARIMA

Fonte: Autor.

4.6 MODELO PARA GERAÇÃO DE SÉRIES DE RADIAÇÃO HORÁRIA

De forma análoga ao item anterior, as estruturas autorregressivas foram modeladas para a série de radiação com intervalo horário. Neste caso o pré-tratamento dos dados foi realizado apenas com os métodos 1 e 2. O método 3 não foi considerado por incompatibilidade dos períodos de medição.

Os oito modelos (quatro estruturas com duas tendências cada) foram estimados para cada mês, totalizando 96 equações para a geração de séries sintéticas. As tabelas 10 e 11 apresentam os parâmetros e as variâncias calculadas.

Tabela 10. Parâmetros dos modelos horários com estrutura AR e ARI

Mês	AR				ARI			
	Método 1		Método 2		Método 1		Método 2	
	a_1	variância	a_1	variância	a_1	variância	a_1	variância
Jan	-0,699	0,024	-0,753	0,013	0,274	0,026	0,061	0,015
Fev	-0,682	0,022	-0,682	0,017	0,257	0,025	0,104	0,020
Mar	-0,746	0,017	-0,703	0,013	0,189	0,018	0,191	0,014
Abr	-0,686	0,023	-0,595	0,019	0,291	0,025	0,306	0,022
Mai	-0,716	0,023	-0,633	0,017	0,215	0,025	0,207	0,020
Jun	-0,717	0,023	-0,645	0,018	0,222	0,026	0,225	0,021
Jul	-0,709	0,019	-0,578	0,017	0,183	0,022	0,296	0,020
Ago	-0,705	0,016	-0,660	0,014	0,239	0,017	0,146	0,016
Set	-0,597	0,024	-0,706	0,016	0,301	0,027	-0,044	0,018
Out	-0,704	0,016	-0,701	0,014	0,190	0,018	0,082	0,016
Nov	-0,759	0,019	-0,759	0,013	0,168	0,021	-0,003	0,015
Dez	-0,714	0,024	-0,754	0,015	0,269	0,026	0,059	0,016

Tabela 11. Parâmetros dos modelos horários com estrutura ARMA e ARIMA

Mês	ARMA						ARIMA					
	Método 1			Método 2			Método 1			Método 2		
	a_1	c_1	variância	a_1	c_1	variância	a_1	c_1	variância	a_1	c_1	variância
Jan	-0,833	-0,277	0,023	-0,715	0,088	0,013	-0,495	-0,922	0,023	-0,692	-0,975	0,013
Fev	-0,795	-0,219	0,022	-0,620	0,114	0,017	-0,527	-0,923	0,022	-0,645	-0,975	0,017
Mar	-0,803	-0,131	0,016	-0,741	-0,076	0,013	-0,581	-0,925	0,016	-0,598	-0,955	0,012
Abr	-0,842	-0,315	0,022	-0,739	-0,227	0,019	-0,465	-0,908	0,022	-0,450	-0,943	0,018
Mai	-0,787	-0,150	0,023	-0,665	-0,053	0,017	-0,576	-0,939	0,022	-0,540	-0,961	0,017
Jun	-0,793	-0,161	0,023	-0,701	-0,095	0,018	-0,656	-0,972	0,023	-0,602	-0,980	0,018
Jul	-0,749	-0,079	0,019	-0,694	-0,176	0,017	-0,612	-0,948	0,019	-0,499	-0,958	0,017
Ago	-0,803	-0,200	0,015	-0,636	0,043	0,014	-0,548	-0,927	0,015	-0,612	-0,976	0,014
Set	-0,777	-0,294	0,023	-0,559	0,311	0,015	-0,422	-0,924	0,023	-0,683	-0,986	0,016
Out	-0,748	-0,087	0,016	-0,646	0,110	0,014	-0,580	-0,946	0,016	-0,663	-0,978	0,014
Nov	-0,796	-0,088	0,019	-0,678	0,196	0,013	-0,617	-0,948	0,018	-0,693	-0,974	0,013
Dez	-0,838	-0,267	0,023	-0,715	0,093	0,014	-0,496	-0,907	0,023	-0,695	-0,972	0,014

A partir das estimativas apresentadas nas tabelas, foram geradas séries sintéticas para todos os modelos. Em seguida, foram comparadas graficamente as séries obtidas com os dados históricos, analisando-se independentemente cada método, estrutura e mês.

A Figura 49 ilustra um exemplo dessa comparação, onde são apresentados os resultados da estrutura AR para o período de janeiro a junho. Nesse conjunto de figuras, as curvas do histórico estão plotadas em amarelo, com seus valores médio e máximo indicados por linhas pretas, e as

curvas dos métodos 1 e 2, respectivamente, em azul e rosa. O resultado completo é apresentado no Anexo VI.

Como principais observações dos resultados, destaca-se que as séries geradas pelas estruturas AR e ARI mostraram maior suavidade nas transições entre valores se comparadas às séries das estruturas ARMA e ARIMA. Além disso, as duas últimas também apresentaram maior número de valores excedentes aos máximos registrados. Assim, apenas pela análise gráfica, já se destaca a melhor adequação das estruturas AR e ARI na geração de séries de radiação horária. Essa avaliação é, entretanto, reavaliada matematicamente na próxima seção.

Fonte: Autor.

4.7 RESULTADOS

Os modelos estimados e apresentados anteriormente foram também avaliados matematicamente, comparando-se o erro quadrático médio de 2.000 séries sintéticas em relação aos dados amostrados. Essa avaliação foi realizada independentemente para cada um dos modelos gerados e para duas fontes de dados reais e distintas, localizadas no nordeste brasileiro: Petrolina/PE e Pombal/PB. O uso de duas localidades se deu para permitir uma primeira verificação da identificação de uma estrutura de modelo como única e válida para qualquer fonte de dados.

A seguir são apresentados e discutidos os resultados obtidos para a aderência das séries, organizados conforme a periodicidade dos dados (diária, mensal e horária) e analisados através de histogramas da energia incidente por período.

Modelos para geração de séries sintéticas de radiação diária

A avaliação dos modelos de radiação diária foi realizada com equações mensais, para as quatro estruturas e dois métodos de tendência, totalizando 96 conjuntos de 2000 séries sintéticas para cada município. Como exemplo, são apresentados os histogramas de Petrolina, para os meses de julho e dezembro, respectivamente na Figura 50 e na Figura 51.

Os meses escolhidos apresentam contrastes da variância típica dos dados, com consequências observadas nas séries geradas, onde o mês de julho, de menor variância, apresenta histogramas com distribuições semelhantes dos dados sintéticos e originais, enquanto que em dezembro há significativa discrepância.

Em seguida, são apresentados os erros quadráticos médios calculados para todos os modelos deste período para Petrolina e Pombal, respectivamente na Tabela 12 e na Tabela 13, com os menores valores de erro para cada mês destacados em verde.

Duas observações importantes são obtidas dos resultados das tabelas:

- Não há uma estrutura única que se destaque como a melhor para a geração de séries sintéticas de radiação para todos os meses; e
- O modelo com o menor erro para Petrolina não é o mesmo para Pombal, não permitindo a indicação clara de um modelo unânime para localidades distintas. Esta indefinição,

entretanto, não significa a impossibilidade de modelagem, mas sim a existência de diversos modelos similares, todos adequados para a representação desejada.

Figura 50. Histogramas normalizados, 2 mil séries de radiação diária, mês de julho

Fonte: Autor.

Figura 51. Histogramas normalizados, 2 mil séries de radiação diária, mês de dezembro

Fonte: Autor.

Tabela 12. Erro quadrático médio dos modelos diários para Petrolina

PETROLINA								
	AR - Met1	AR - Met3	ARI - Met1	ARI - Met3	ARMA - Met1	ARMA - Met3	ARIMA - Met1	ARIMA - Met3
Jan	0,043	0,061	0,068	0,076	0,005	0,015	0,116	0,128
Fev	0,021	0,015	0,013	0,019	0,024	0,020	0,140	0,142
Mar	0,025	0,040	0,015	0,048	0,016	0,005	0,139	0,160
Abr	0,031	0,027	0,037	0,004	0,113	0,124	0,148	0,157
Mai	0,005	0,022	0,014	0,018	0,008	0,019	0,013	0,014
Jun	0,030	0,037	0,006	0,023	0,052	0,029	0,085	0,077
Jul	0,059	0,059	0,062	0,003	0,037	0,049	0,021	0,008
Ago	0,004	0,020	0,012	0,004	0,004	0,017	0,132	0,145
Set	0,023	0,019	0,030	0,029	0,031	0,014	0,104	0,143
Out	0,039	0,040	0,047	0,042	0,035	0,033	0,068	0,078
Nov	0,038	0,028	0,047	0,046	0,141	0,091	0,102	0,109
Dez	0,039	0,047	0,073	0,071	0,016	0,013	0,145	0,133

Tabela 13. Erro quadrático médio dos modelos diários para Pombal

POMBAL								
	AR - Met1	AR - Met3	ARI - Met1	ARI - Met3	ARMA - Met1	ARMA - Met3	ARIMA - Met1	ARIMA - Met3
Jan	0,016	0,023	0,036	0,042	0,030	0,012	0,143	0,153
Fev	0,009	0,009	0,015	0,003	0,005	0,003	0,113	0,111
Mar	0,000	0,015	0,001	0,002	0,075	0,063	0,115	0,119
Abr	0,024	0,014	0,019	0,021	0,261	0,183	0,090	0,101
Mai	0,011	0,002	0,015	0,008	0,056	0,058	0,060	0,059
Jun	0,034	0,039	0,018	0,040	0,035	0,025	0,002	0,004
Jul	0,033	0,014	0,019	0,010	0,020	0,013	0,029	0,012
Ago	0,023	0,034	0,028	0,013	0,025	0,010	0,095	0,138
Set	0,017	0,018	0,017	0,017	0,086	0,070	0,008	0,002
Out	0,026	0,024	0,026	0,026	0,027	0,027	0,047	0,042
Nov	0,027	0,028	0,029	0,026	0,008	0,013	0,041	0,018
Dez	0,029	0,031	0,036	0,036	0,025	0,020	0,082	0,068

Modelos para geração de séries sintéticas de radiação mensal

Os modelos de radiação mensal foram também comparados para as quatro estruturas e dois métodos de remoção de tendência, avaliando-se as séries geradas para os 8 conjuntos de cada município. Os histogramas dos dados reais e sintéticos de Petrolina são ilustrados na Figura 52.

Neste caso se observa claramente que o método 3 (remoção de tendência pela média mensal) não consegue reproduzir a forma do histograma dos dados reais, enquanto que o método 1 tem boa aderência. A comparação entre as estruturas apenas pela análise visual não tem a mesma obviedade, porém pode se identificar certa vantagem para o modelo AR.

Em seguida, os resultados numéricos dos erros quadráticos, para os dois locais, são apresentados na Tabela 14. Os valores obtidos confirmam as interpretações visuais anteriores para Petrolina, onde o método 1 é sempre melhor que o método 3 e a estrutura AR tem a melhor aderência. Adicionalmente, os mesmos aspectos são verificados no outro conjunto de dados, fortalecendo a conclusão que o modelo AR com remoção de tendência pela média é o mais indicado quando se trata de análises da radiação mensal.

Figura 52. Histogramas normalizados, 2 mil séries sintéticas de radiação mensal

Fonte: Autor.

Tabela 14. Erro quadrático médio dos modelos mensais

Localidade	Modelo							
	AR-Met1	AR-Met3	ARI-Met1	ARI-Met3	ARMA-Met1	ARMA-Met3	ARIMA-Met1	ARIMA-Met3
Petrolina	0,056	0,098	0,058	0,130	0,059	0,139	0,063	0,148
Pombal	0,057	0,115	0,069	0,166	0,066	0,173	0,075	0,014

Modelos para geração de séries sintéticas de radiação horária

A avaliação dos modelos de radiação horária foi realizada da mesma forma que para a radiação diária, ou seja, utilizando equações mensais, para as quatro estruturas e dois métodos de tendência, para cada município. Os histogramas ilustrativos dos resultados também foram escolhidos para os meses de julho e dezembro e são apresentados, respectivamente, na Figura 53 e na Figura 54.

Em ambos os meses e para todas as estruturas se observa que os dados sintéticos horários gerados pelo método 1 se aproximam melhor dos dados reais, com vantagem para as estruturas AR e ARI. Porém, apesar de comparativamente melhores, há bastante geração de valores superiores aos máximos históricos registrados.

O resultado completo dos erros quadráticos médios, calculados para todos os conjuntos analisados e para as duas localidades, é apresentado na Tabela 15 e na Tabela 16, com o melhor resultado para cada mês destacado em verde.

Neste caso, o método 1 manteve seu destaque como melhor estratégia para a retirada da tendência e a estrutura ARI dominou as recomendações como melhor modelo para a geração de séries. A estrutura AR também apresentou bons resultados, inclusive sendo a melhor para alguns meses, enquanto que ARMA e ARIMA se mostraram bastante inferiores. Os resultados são válidos para as duas cidades, porém mais estudos podem ser conduzidos para se confirmar a identificação do ARI como definitiva.

Figura 53. Histogramas normalizados, 2 mil séries de radiação horária, mês de julho.

Fonte: Autor.

Figura 54. Histogramas normalizados, 2 mil séries de radiação horária, mês de dezembro.

Fonte: Autor.

Tabela 15. Erro quadrático médio dos modelos horários para Petrolina

PETROLINA								
	AR - Met1	AR - Met2	ARI - Met1	ARI - Met2	ARMA - Met1	ARMA - Met2	ARIMA - Met1	ARIMA - Met2
Jan	0,057	0,103	0,045	0,109	0,082	0,106	0,075	0,114
Fev	0,053	0,094	0,037	0,099	0,078	0,101	0,083	0,109
Mar	0,055	0,103	0,034	0,104	0,071	0,109	0,086	0,126
Abr	0,057	0,089	0,046	0,086	0,092	0,103	0,077	0,106
Mai	0,029	0,098	0,033	0,101	0,047	0,101	0,066	0,118
Jun	0,030	0,074	0,034	0,077	0,046	0,079	0,075	0,105
Jul	0,023	0,102	0,036	0,104	0,031	0,104	0,066	0,117
Ago	0,049	0,079	0,040	0,085	0,069	0,081	0,074	0,103
Set	0,082	0,141	0,075	0,145	0,113	0,150	0,105	0,153
Out	0,074	0,116	0,056	0,118	0,083	0,116	0,107	0,132
Nov	0,061	0,106	0,051	0,110	0,075	0,104	0,092	0,111
Dez	0,071	0,109	0,063	0,116	0,095	0,110	0,090	0,119

Tabela 16. Erro quadrático médio dos modelos horários para Pombal

POMBAL								
	AR - Met1	AR - Met2	ARI - Met1	ARI - Met2	ARMA - Met1	ARMA - Met2	ARIMA - Met1	ARIMA - Met2
Jan	0,066	0,114	0,055	0,111	0,103	0,116	0,092	0,123
Fev	0,052	0,128	0,041	0,129	0,089	0,127	0,084	0,144
Mar	0,051	0,089	0,044	0,087	0,095	0,090	0,082	0,111
Abr	0,058	0,121	0,051	0,121	0,100	0,124	0,089	0,144
Mai	0,071	0,093	0,076	0,097	0,141	0,095	0,111	0,130
Jun	0,076	0,096	0,073	0,097	0,119	0,091	0,121	0,140
Jul	0,056	0,105	0,042	0,107	0,095	0,111	0,109	0,144
Ago	0,052	0,108	0,052	0,109	0,079	0,106	0,088	0,134
Set	0,065	0,133	0,059	0,133	0,123	0,132	0,090	0,151
Out	0,068	0,117	0,057	0,117	0,100	0,114	0,100	0,141
Nov	0,066	0,131	0,059	0,130	0,085	0,130	0,097	0,146
Dez	0,066	0,117	0,061	0,118	0,109	0,117	0,091	0,139

4.8 CONCLUSÕES PARCIAIS

Neste capítulo foram avaliados distintos modelos de radiação solar com periodicidade horária, diária e mensal, analisando-se o método de pré-filtragem dos dados, a estrutura do modelo e o algoritmo utilizado para a estimativa dos parâmetros.

Para a radiação horária foram comparados 96 modelos e pôde-se notar que, de forma geral, apenas o pré-processamento pelo método Yule-Walker é recomendado para as duas cidades analisadas. O método de retirada da tendência pela média diária se mostrou como o melhor em todas as simulações realizadas. Quanto à estrutura, AR e ARI apresentaram resultados semelhantes, com ligeira vantagem para a ARI.

O período diário foi avaliado com a proposição de 11 modelos de energia total incidente para cada mês, resultando em 132 modelos distintos. Séries sintéticas foram geradas para cada um deles e a análise comparativa com dados reais mostra que não se pode propor uma única estrutura de modelo para todos os meses ou até mesmo para as duas cidades utilizadas como estudo de caso.

Para estudos de longo prazo, com base na energia mensal incidente, foram propostos oito modelos e concluiu-se que pode ser recomendado o modelo AR com remoção de tendência pelo valor médio para as duas cidades analisadas. Sugere-se também a avaliação do uso do modelo PAR, variação convencionalmente utilizada no setor nacional em aplicações hidrológicas e que permite diferenciação mensal do modelo.

Quanto ao método de estimativa dos parâmetros, é perceptível o frequente destaque do Yule-Walker em relação aos mínimos quadrados. Este método, além de apresentar os melhores resultados, tem implementação fácil e é indicado como o mais adequado para a continuidade do estudo.

Assim, no total, 236 modelos para geração de séries sintéticas foram desenvolvidos e avaliados nesta etapa. Destaca-se que não se pode apontar um único modelo como o melhor para qualquer período de análise e que, para cada aplicação, deve ser avaliado o mais adequado.

5 *APRIMORAMENTO DO MODELO*

Os capítulos anteriores desta tese estudaram modelos de séries temporais, especialmente aplicadas para avaliação do comportamento da radiação solar incidente em parques de geração de energia elétrica. As avaliações foram bastante amplas e com boa fundamentação metodológica nos conceitos de modelagem de séries, permitindo evolução em diversas frentes de estudo, sejam conceituais ou práticas.

Em paralelo, observando-se o cenário nacional, há oportunidades de aplicação dos modelos estudados tanto para previsão da geração imediata (ex. operação da distribuição e do sistema interligado), como para projeto de sistemas isolados (ex. dimensionamento de armazenamento) e avaliações energéticas de longo prazo (ex. risco de contratos e suprimento).

Adicionalmente, a evolução do mercado de geração solar no Brasil impacta nas três frentes citadas, com: grande crescimento de unidades de microgeração em baixa tensão, implementação constante de sistemas em localidades isoladas e a realização cada vez mais frequente de leilões de usinas fotovoltaicas centralizadas de grande porte.

Dentre essas evoluções, a instalação de grandes usinas gera maiores impactos, tanto econômicos como técnicos, merecendo destaque por sua importância (em volume de energia, de potência e financeiro) e potenciais consequências, como o não atendimento da demanda e de contratos.

Partindo deste cenário, esta etapa da pesquisa se propôs a aprimorar as modelagens realizadas com foco no suprimento energético, aplicável tanto do ponto de vista do gerador como dos operadores, permitindo melhor avaliação das incertezas de geração e, consequentemente, do planejamento do sistema e da contratação da energia.

Para tanto, o modelo autorregressivo – AR, identificado anteriormente como melhor alternativa para a análise de séries de longo prazo, foi implementado e aprimorado com sua variação periódica PAR, assim como a metodologia já consolidada e utilizada nos estudos do sistema hidrelétrico, porém inovadora para a modelagem da radiação solar.

Em seguida, após a estimativa das duas estruturas para modelos de primeira, segunda e terceira ordem, foi estudada sua dependência temporal, assim como a aderência a dados reais de radiação de Petrolina.

A seguir são apresentados o equacionamento dos modelos autorregressivos convencional e periódico, a distribuição dos resíduos encontrados, sua verificação por diferentes métodos de cálculo de erro e recomendações para aplicações práticas.

Destaca-se que o equacionamento supracitado é uma adaptação do apresentado nos capítulos 2 e 3, complementado por (COSE - UNICAMP, 2010), reapresentando os conceitos dos modelos e com exemplo de implementação para melhor compreensão.

5.1 ESTIMATIVAS DO MODELO AR

Um processo autorregressivo de ordem p , $AR(p)$, é definido como um modelo onde o valor atual do processo z_t é expresso como uma combinação linear dos valores p passados e um ruído branco (BOX e JENKINS, 1976), onde a_t é uma série de variáveis aleatórias com distribuição normal, média zero e variância σ_a^2 , ou seja, $N(0, \sigma_a^2)$. O modelo $AR(p)$ pode ser escrito como:

$$z_t = a_t + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} \quad [28]$$

A partir dos dados da série estacionária z_t , as funções autocovariância e autocorrelação são calculadas através da estimativa de seus coeficientes, respectivamente, por [29] e [30].

$$c_k = \frac{1}{N-k} \sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z}), \quad k = 0, 1, \dots, K \quad [29]$$

$$r_k = \frac{c_k}{c_0} \quad [30]$$

Onde K corresponde ao intervalo em que as correlações de dados são analisadas e

$$\bar{z} = \frac{1}{N} \sum_{t=1}^N z_t \quad [31]$$

Uma vez que o cálculo das funções de autocorrelação é realizado, o modelo é ajustado para diferentes ordens através da estimativa dos coeficientes de autocorrelação parcial (ϕ_{kk}).

Como exemplo, um modelo de ordem 3 - AR (3) - tem a seguinte forma:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \phi_3 z_{t-3} + a_t \quad [32]$$

Sendo $\phi_3 = \hat{\phi}_{33}$, o coeficiente ϕ_1 é dado pela estimação do coeficiente de autocorrelação $\hat{\phi}_{31}$ e o coeficiente ϕ_2 é obtido por $\hat{\phi}_{32}$. Estes coeficientes são estimados por meio das equações de Yule-Walker [33].

$$r_j = \hat{\phi}_{31} r_{j-1} + \hat{\phi}_{32} r_{j-2} + \hat{\phi}_{33} r_{j-3} \quad [33]$$

Onde $j = 1, 2, 3$ e $r_j = r_{-j}$.

Para o modelo $AR(3)$ a equação [33] expressa-se da seguinte forma:

$$\begin{cases} r_1 = \hat{\phi}_{31} r_0 + \hat{\phi}_{32} r_{-1} + \hat{\phi}_{33} r_{-2} \\ r_2 = \hat{\phi}_{31} r_1 + \hat{\phi}_{32} r_0 + \hat{\phi}_{33} r_{-1} \\ r_3 = \hat{\phi}_{31} r_2 + \hat{\phi}_{32} r_1 + \hat{\phi}_{33} r_0 \end{cases} \quad [34]$$

E o seu residual a_t é calculado por [35] para todos os dados disponíveis, verificando que $a_t \sim N(0, \sigma_a^2)$.

$$a_t = z_t - \phi_1 z_{t-1} - \phi_2 z_{t-2} - \phi_3 z_{t-3} \quad [35]$$

O desenvolvimento do estudo utilizou a mesma série histórica de radiação dos capítulos anteriores, com medição para o período 1996-2015, e o pré-tratamento para tornar a série estacionária foi realizado através da remoção da tendência pelos valores médios.

Na estimativa dos modelos foram considerados 18 dos 20 anos da série para o cálculo dos coeficientes, permitindo posterior comparação dos dados observados para os dois anos seguintes com as séries sintéticas geradas.

A Tabela 17 apresenta os valores da série z_t após o tratamento, com o cálculo do desvio e da média.

Tabela 17. Série mensal de radiação, pré-tratada para remoção de tendência.

	Janeiro	Fevereiro	Março	Abril	Maio	Junho	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
1996	0.1422	1.3010	-0.2744	-1.5745	0.1436	-0.5767	0.5959	-0.6962	-0.6087	-0.2319	-1.9652	-0.0457
1997	-1.1465	-0.7432	-2.5641	-1.1323	-1.2227	0.7043	0.3487	0.5074	2.1703	0.2079	0.8474	0.6745
1998	-0.5231	1.4989	0.8862	1.6388	1.3071	-0.5805	-0.0690	1.0946	0.9808	1.5976	-0.3222	0.2314
1999	0.7008	0.2953	-0.0152	0.5938	-0.1383	0.8060	0.9052	-0.8349	-0.7306	-1.1237	-0.6450	-1.3983
2000	-1.1056	-0.3585	-0.3215	-0.7254	0.4292	-0.3692	0.3538	-0.1319	-1.9373	0.9628	-0.8856	-0.8169
2001	0.8203	0.2678	0.1704	0.7934	1.1148	-0.3528	-1.1545	-0.7971	-0.3910	0.2588	1.3882	-0.2285
2002	-1.7707	0.0549	0.5175	0.1881	0.5257	-0.3911	1.1276	1.0588	0.1477	0.9255	1.2082	0.0790
2003	0.6159	0.0204	0.5637	0.6058	-0.3970	0.6077	-0.4207	0.2138	-0.3943	0.0056	0.2653	1.4575
2004	-1.8709	-0.6029	-0.2098	0.3927	0.9371	-0.0301	-0.2414	0.3278	-0.0993	0.5402	0.6374	1.2129
2005	0.0981	-0.2820	-0.2448	-0.2099	-1.4266	-1.0164	-0.8401	-0.1041	1.0445	0.9773	0.5207	-0.4221
2006	1.1340	0.0880	-0.9053	-1.1711	-0.7198	-1.6905	-0.0405	1.0900	-0.8781	-0.9529	0.0451	1.1965
2007	1.6109	-2.0377	-0.0112	0.4597	0.3246	0.5446	1.0784	-0.8881	-0.5311	0.4404	0.9806	0.6259
2008	1.0799	0.2915	-0.9425	-0.7489	-0.5842	0.4142	-0.3447	-0.4773	0.8499	-0.0394	0.9974	-0.5470
2009	0.2089	-0.7580	0.1498	-1.1322	-1.5744	-0.8968	-1.0901	-2.0018	0.6355	-1.8133	0.9057	-0.5994
2010	0.0885	-0.1886	-0.1242	-0.8022	1.4844	-1.4250	-2.6411	0.5678	-1.3446	-0.5769	-1.3300	-1.2393
2011	-0.2243	-1.4237	-0.1869	0.6973	-1.1828	0.8359	0.1938	2.2569	0.2693	-1.9399	-1.0232	0.4752
2012	0.7118	0.4015	2.3862	1.9654	1.3432	2.2879	1.2584	-0.8101	1.0589	1.1596	-0.6299	1.3422
2013	-0.5703	2.1751	1.1260	0.1615	-0.3640	1.1286	0.9800	-0.3754	-0.2419	-0.3978	-0.9950	-1.9975
Desvio	1	1	1	1	1	1	1	1	1	1	1	1
Média	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

A partir dos dados da série tratada foram calculadas as funções de autocovariância e autocorrelação das equações [29] e [30], obtendo-se os seguintes resultados.

$$c_0 = 0,944 \quad r_0 = 1,000$$

$$c_1 = 0,225 \quad r_1 = 0,239$$

$$c_2 = 0,170 \quad r_2 = 0,180$$

$$c_3 = 0,159 \quad r_3 = 0,169$$

$$c_4 = 0,083 \quad r_4 = 0,088$$

$$c_5 = 0,121 \quad r_5 = 0,128$$

$$c_6 = 0,050 \quad r_6 = 0,053$$

Em seguida, através das equações de Yule-Walker [33] são obtidos os parâmetros ϕ_1 a ϕ_n para cada ordem desejada do modelo. Como exemplo, as equações para o modelo AR de terceira ordem dadas por [34] formam os seguintes sistemas.

$$\begin{bmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0,239 & 0,180 \\ 0,239 & 1 & 0,239 \\ 0,180 & 0,239 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} = \begin{bmatrix} 0,239 \\ 0,180 \\ 0,169 \end{bmatrix}$$

A solução do sistema obtém os parâmetros ϕ_1 , ϕ_2 e ϕ_3 do modelo AR(3):

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} = \begin{bmatrix} 0,193 \\ 0,108 \\ 0,108 \end{bmatrix}$$

Assim, o modelo AR (3) para a série de dados pode ser escrito pela equação [36].

$$z_t = 0.193 z_{t-1} + 0.108 z_{t-2} + 0.108 z_{t-3} + a_t \quad [36]$$

Onde $t = 1, \dots, N$.

Após a obtenção das equações para as diferentes ordens desejadas, os resíduos a_t são calculados utilizando-se a equação [35]. A Figura 55 ilustra a distribuição dos resíduos calculados para as ordens 1, 2 e 3 do modelo AR, confirmando a característica da distribuição normal em todos os casos.

Fonte: Autor.

Uma vez estimados os modelos é possível criar séries sintéticas, com qualquer quantidade e duração, através da geração de resíduos aleatórios que estejam dentro do intervalo dos resíduos calculados e possuam igual probabilidade de ocorrência.

A Figura 56 ilustra exemplo de um cenário sintético, em conjunto aos dados realmente observados, para os 24 meses seguintes ao período utilizado na calibração do modelo AR(3).

Figura 56. Radiação mensal observada e prevista, modelo AR(3)

Fonte: Autor.

5.2 MODELO AUTORREGRESSIVO PERIÓDICO - PAR E ESTIMATIVAS

Os modelos periódicos autorregressivos PAR foram propostos por (JONES e BRELSFORD, 1967), e depois estudados por (VECCHIA, 1985) e (ANDERSON e VECCHIA, 1993). A estimativa dos parâmetros deste modelo é efetuada através do método de máxima verossimilhança (VECCHIA, 1985) cuja função é escrita pelas seguintes equações.

$$L_m(\Phi_m, \tau_m | Z_m) \propto \tau_m^{n_m/2} \exp \left\{ -\frac{\tau_m}{2} (Z_m - X_m \Phi_m)' (Z_m - X_m \Phi_m) \right\} \quad [37]$$

Onde:

$$n_m = \begin{cases} n-1, & m = 1, \dots, p_m \\ n, & m = p_m + s, s = 1, \dots, 12 \end{cases} \quad [38]$$

$$Z_m = \begin{pmatrix} Z_{t(1,m)} \\ Z_{t(2,m)} \\ \vdots \\ Z_{t(n,m)} \end{pmatrix}_{n \times 1} \quad [39]$$

$$\Phi_m = \begin{pmatrix} \phi_{1,m} \\ \phi_{2,m} \\ \vdots \\ \phi_{p_m,m} \end{pmatrix}_{p_m \times 1} \quad [40]$$

$$X_m = \begin{pmatrix} Z_{t(1,m)-1} & \dots & Z_{t(1,m)-p_m} \\ Z_{t(2,m)-1} & \dots & Z_{t(2,m)-p_m} \\ \vdots & \ddots & \dots \\ Z_{t(n,m)-1} & \dots & Z_{t(n,m)-p_m} \end{pmatrix}_{n \times p_m} \quad [41]$$

E os estimadores de máxima verossimilhança são dados por [42] e [43].

$$\hat{\Phi}_m = (X'_m X_m)^{-1} (X'_m Z_m) \quad [42]$$

$$\hat{\tau}_m^{-1} = \frac{1}{n_m} (Z_m - X_m \hat{\Phi}_m)' (Z_m - X_m \hat{\Phi}_m) \quad [43]$$

O modelo PAR (3) está definido como em [32], porém com coeficientes diferenciados para cada intervalo pré-estabelecido m (no caso mensal, $m = 1, 2, \dots, 12$), conforme a equação [44].

$$z_t = \phi_{1,m} z_{t-1} + \phi_{2,m} z_{t-2} + \phi_{3,m} z_{t-3} + a_t \quad [44]$$

Na estimativa da estrutura PAR de ordem 3 do exemplo mensal, são calculados os três coeficientes para cada intervalo de análise, totalizando 12 equações e 36 parâmetros. Assim, para cada previsão a ser realizada, deve se considerar a equação específica do mês desejado. A tabela seguinte apresenta os resultados obtidos para o modelo PAR(3).

Tabela 18. Parâmetros do modelo PAR(3) mensal

	Janeiro	Fevereiro	Março	Abril	Maior	Junho	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
ϕ_{1m}	-0.0031609	-0.1052534	0.3707063	0.7923203	0.3148325	-0.2896871	0.6567496	0.122884	0.0975242	0.1581622	0.121295392	0.1779883
ϕ_{2m}	0.0718592	0.2373029	0.1360879	-0.2124388	0.2079257	0.6854297	-0.0245297	-0.2149732	-0.1402646	0.0199317	0.386600177	0.1642139
ϕ_{3m}	0.0725182	0.0805206	0.1327309	-0.0545637	0.1525813	0.0128936	-0.0669072	0.0355052	0.4179472	0.1960682	-0.199650993	0.1779181

Por exemplo, a equação do modelo PAR (3) para o mês de janeiro pode ser escrita como em [45].

$$z_t = -0.0032 z_{t-1} + 0.0719 z_{t-2} + 0.0725 z_{t-3} + a_t \quad [45]$$

Após a estimativa dos parâmetros, o resíduo a_t é calculado de forma semelhante ao modelo AR, verificando novamente que $a_t \sim N(0, \sigma_a^2)$, utilizando a equação [46].

$$a_t = z_t - \phi_{1,m} z_{t-1,m} - \phi_{2,m} z_{t-2,m} - \phi_{3,m} z_{t-3,m} \quad [46]$$

O cálculo dos resíduos em modelos periódicos é realizado para cada período considerado e para cada ordem do modelo. Ou seja, para os modelos mensais de ordem 1, 2 e 3 são obtidas 36

diferentes distribuições de resíduo. Como exemplo, a Figura 57 ilustra o desvio padrão das distribuições de resíduo mensais do modelo PAR(3). Em seguida, a Figura 58 apresenta o histograma dos resíduos dos meses de janeiro e abril, respectivamente os de maior e menor variação.

Figura 57. Desvio padrão dos resíduos mensais do modelo PAR(3)
Desvio padrão dos resíduos PAR(3)

Fonte: Autor.

Figura 58. Distribuição dos resíduos PAR(3) em janeiro e abril

Fonte: Autor.

De forma análoga ao realizado para o modelo AR(3), a Figura 59 ilustra exemplo de um cenário sintético, em conjunto aos dados realmente observados, para os 24 meses seguintes ao período utilizado na calibração do modelo PAR(3).

Figura 59. Radiação mensal observada e prevista, modelo PAR(3)

Fonte: Autor.

5.3 VERIFICAÇÃO

A verificação dos modelos foi realizada de forma comparativa entre as previsões geradas e os dados observados, utilizando como índices de comparação os erros quadrático médio, absoluto médio e relativo percentual médio.

Para tanto, após o ajuste dos modelos AR e PAR de ordens 1, 2 e 3 com os dados de 18 anos, foram produzidas 1.000 previsões por mês para dois anos à frente, com valores aleatórios dos resíduos. As previsões foram então aplicadas nas fórmulas de cálculo dos erros:

- A raiz do erro quadrático médio (REQM) indica a média do afastamento de todos os valores previstos e o seu real valor, penalizando as maiores diferenças pelo uso do efeito quadrático;

$$REQM = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad [47]$$

- O erro absoluto médio (EAM) indica a média do afastamento de todos os valores previstos e o seu real valor, sem penalizações; e

$$EAM = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad [48]$$

- O erro relativo percentual médio (EAM) indica a média do quociente entre o erro absoluto e o valor observado, ponderando os erros individuais pela sua representatividade em relação ao valor original.

$$EPM = \frac{100}{n} \sum_{j=1}^n \frac{|y_j - \hat{y}_j|}{y_j} \quad [49]$$

Em todas as equações, $n = 1, \dots, 24$, y_j é o valor real e $\hat{y}_j$ é o valor previsto.

Os erros foram calculados para as 1000 séries de cada estrutura e ordem, totalizando 18 erros médios. A fim de permitir a comparação entre os diferentes modelos e erros, as médias foram ajustadas para valores por unidade, com base no menor valor dos erros calculados. Os resultados são apresentados na Tabela 19 e na Figura 60.

Tabela 19. Erros médios dos modelos

Modelo	Erro (pu)		
	REQM	EAM	EPM
AR (1)	1,034	1,023	1,016
AR (2)	1,037	1,026	1,019
AR (3)	1,059	1,047	1,038
PAR (1)	1,000	1,000	1,000
PAR (2)	1,011	1,013	1,013
PAR (3)	1,017	1,022	1,022

Figura 60. Erros dos modelos (pu)

Fonte: Autor.

5.4 COMENTÁRIOS

Este capítulo aprimorou o estudo realizado anteriormente para a estrutura autorregressiva AR com intervalo mensal, testando-a com maior profundidade e realizando comparações com sua variação periódica PAR. Além da modelagem em si, foram avaliadas as distribuições de resíduo, a dependência temporal e a aderência aos dados reais.

A análise da distribuição dos resíduos foi realizada visualmente, onde se observou características da distribuição normal em todas as simulações, especialmente nos modelos de menor ordem. Nenhum modelo ou ordem foi descartado e não se considerou necessário evoluir nesta análise.

A dependência temporal foi inicialmente avaliada pela análise dos coeficientes de cada ordem dos modelos, porém todos apresentaram dependências próximas e sem variações significativas, não permitindo conclusões diretas. Como análise complementar, a comparação dos erros encontrados permite identificar que os modelos de primeira ordem são sempre melhores que os demais, independentemente do tipo do modelo e do método de cálculo do erro.

A observação dos erros também permite concluir que o modelo periódico sempre apresenta melhores resultados que o autorregressivo convencional, independentemente da ordem e do método de cálculo do erro.

Como principal conclusão deste capítulo, destaca-se o modelo PAR (1) como o mais adequado para a representação da radiação com periodicidade mensal, capaz de gerar série sintéticas com boa precisão e de fácil implementação.

6 APLICAÇÕES PRÁTICAS EM LEILÕES E EM RISCOS DE GERAÇÃO

Este capítulo apresenta uma aplicação prática da modelagem de séries sintéticas de radiação solar, baseada nos resultados e conclusões dos estudos realizados, avaliando a produção de energia de uma planta fotovoltaica e as consequências da variação da radiação nos contratos de comercialização, na confiabilidade do suprimento e na receita esperada.

As análises partiram da modelagem de um contrato típico de compra de energia no mercado regulado, que faz a contabilidade da geração totalizada mensalmente. Para a sua simulação foram então geradas séries sintéticas com o modelo autorregressivo periódico de primeira ordem PAR(1), indicado anteriormente como o melhor para estudos de séries mensais.

A aplicação dos cenários gerados no modelo com as regras contratuais permitiu análises diversas, como a estimativa da receita esperada, a proposição de estratégia para leilões, a avaliação de riscos de geração e a análise das probabilidades de perda de receita.

Os itens seguintes apresentam o estudo realizado e os resultados encontrados.

6.1 REGRAS DE LEILÕES DE ENERGIA

O estudo de caso realizado segue as regras de acordos típicos dos leilões de energia brasileira para fontes fotovoltaicas (ANEEL, 2016), representadas formalmente por documento conhecido por “Power Purchase Agreement” - PPA. No entanto, o processo de análise pode ser utilizado com qualquer regra de comercialização ou para prever o risco de desempenho da geração esperada.

O acordo padrão para os leilões acima mencionados dura 20 anos, nos quais o vendedor concorda em entregar um valor fixo de energia média E_{agr} , único para todo o período, com contabilidade mensal e controle de desempenho anual.

Este valor exclusivo é limitado pela geração média esperada E_{exp} da usina, calculada com base no histórico de radiação de longo prazo da localidade.

No entanto, uma vez que a geração elétrica depende do recurso solar e apresenta variações ao longo dos diferentes anos, existem regras para ajustar a diferença entre os valores esperados e os valores reais, respectivamente, E_{exp} e E_g .

Essas variações de geração, embora esperadas, são penalizadas reduzindo-se a geração atual E_g para a geração contabilizada E_{ac} . As regras de contabilização são diferenciadas em quatro faixas de geração anual E_g em relação à energia contratada E_{agr} , de acordo com as regras descritas a seguir e ilustradas na Figura 61:

- (A) 115+ % - Em caso de geração anual superior a 115 % da energia contratada, a energia referente a 115 % é remunerada ao preço do contrato e o excedente é valorado a apenas 30 %, ou seja, há uma penalização de 70 % no preço da energia excedente a 115 %. A equação ilustra este cálculo.

$$E_{ac}^A = 115 \% * E_{agr} + 30 \% * \left(\frac{E_g}{E_{agr}} - 115 \% \right) * E_{agr} \quad [50]$$

- (B) 100 - 115 % - A geração real dentro da faixa de 100 % a 115 % do previsto em contrato não implica em penalizações e o vendedor é remunerado pela energia entregue e pelo valor acordado.

$$E_{ac}^B = E_g \quad [51]$$

- (C) 90 – 100 % - A geração inferior à prevista, porém limitada a 10 % de redução, é penalizada conforme a equação [52]. Neste caso, além de não receber pela energia não gerada, o vendedor paga uma penalidade adicional de 6% do preço do contrato.

$$E_{ac}^C = E_g - 6\% * \left(100 \% - \frac{E_g}{E_{agr}} \right) * E_{agr} \quad [52]$$

- (D) 0 – 90 % - A geração inferior a 90 % da contratada implica em duas penalidades adicionais, conforme ilustra a equação [53]. A parcela referente aos primeiros 10 % devidos é penalizada como no item anterior, ou seja, o vendedor não recebe pela energia e paga um adicional de 6 % do valor do contrato (sobre 10 % da energia contratada apenas). Além disso, a parcela não entregue e inferior a 90 % é penalizada com 15 % do valor do contrato.

$$E_{ac}^D = E_g - 15 \% * \left(90 \% - \frac{E_g}{E_{agr}} \right) * E_{agr} - 6 \% * 10 \% * E_{agr} \quad [53]$$

Figura 61. Contabilização da energia nos contratos de compra de energia (PPA)

6.2 GERAÇÃO DE SÉRIES SINTÉTICAS

A geração de séries sintéticas adotou o modelo *PAR* (1), uma vez que foi o que apresentou os melhores resultados para a geração de séries mensais de radiação.

A calibração do modelo utilizou dados amostrados mensais por um período ininterrupto de 20 anos, que representam a totalidade dos dados disponíveis, pois não será necessário realizar validações adicionais. Uma vez calibrado, foram geradas 1.000 séries sintéticas mensais, com duração de um ano cada.

Vale ressaltar que as séries geradas por este método são diferentes das séries históricas, mas igualmente prováveis. A Figura 62 apresenta um exemplo de 20 séries geradas e uma amostra real do mesmo período.

Figura 62. Exemplo de séries sintéticas

Fonte: Autor.

A energia incidente das séries tem, além da dependência temporal, influência das distribuições mensais dos resíduos calculados anteriormente e, quando acumuladas anualmente e analisadas em conjunto, apresentam uma nova variação característica da distribuição anual. A Figura 63 exemplifica essa distribuição.

Figura 63. Distribuição das séries de energia incidente
Distribuição das Séries de Energia Incidente

Fonte: Autor

O histograma desses valores é apresentado na Figura 64, onde, como esperado, há semelhança com a distribuição normal. As pequenas diferenças são causadas por conta do dado original ter periodicidade mensal, onde cada mês apresenta distribuição normal perfeita, e o acumulado anual depender dos sorteios realizados na geração das séries.

Figura 64. Histograma das séries de energia incidente

6.3 PREMISSAS

A estratégia de participação nos leilões foi avaliada modelando-se as regras de penalidade do PPA e calculando a redução de receita esperada, para cada cenário de radiação gerado e para diferentes quantidades de energia contratada.

Os valores de energia contratada são proporcionais à energia esperada e variam na faixa de 85-100%, com intervalos de 1%.

Os cenários de radiação foram obtidos pelo cálculo de 1.000 séries de 12 meses, em seguida agregadas em 1.000 séries anuais. Isso permitiu melhor precisão e, ao mesmo tempo, levar em consideração o período contábil do PPA.

Os valores de energia gerada e as receitas apresentadas neste relatório estão em valores relativos - por unidade ou porcentagem - para permitir comparações adequadas. O processo de cálculo, no entanto, usa as equações mencionadas anteriormente com dados absolutos e reais para o município de Petrolina. Também deve notar-se que os valores relativos apresentados são válidos para qualquer tamanho de planta de geração.

6.4 RECEITA ESPERADA

Como primeira análise, a geração de energia considerada no PPA foi calculada para o primeiro ano de operação, para cada possibilidade de energia contratada. A Figura 65 mostra (em barras

azuis) a energia do acordo, a energia extra contabilizada e a energia total. A curva amarela indica a perda de receita esperada para cada valor contratado.

No entanto, a energia fotovoltaica é caracterizada por uma perda de eficiência nas células ao longo do tempo, reduzindo naturalmente sua capacidade de geração. Ou seja, para cada ano de operação a estratégia de venda pode ser diferente. Como exemplo, a Figura 66 ilustra a mesma simulação para a operação da planta em seu 20º ano.

Conforme observado no gráfico do ano 1, há uma pequena diferença na receita de contratos na faixa de 90 – 100 %, possivelmente não justificando estudos elaborados para definir a estratégia de vendas. Neste exemplo, a perda de receita é de 0,128 % em um contrato completo (100 %) e 0,075 % na melhor estratégia (93 %). A diferença de 0,053 % representa cerca de 23.000,00 reais por ano em uma planta de 100 MW.

Porém, na simulação do ano 20 é possível notar um comportamento bastante diferente em relação à redução de receita, uma vez que as penalidades por não cumprimento da geração prometida tornam-se significativas nos maiores contratos. A melhor opção para o ano 20 é o contrato de 85 %, que apresenta uma perda de receita esperada de 15,9 %, contra 17,2 % do contrato completo. Essa redução é superior a 500 mil reais por ano, valor bastante considerável.

6.5 DEFINIÇÃO DO MELHOR LANCE DE ENERGIA

A análise realizada no item anterior pode ser aplicada para qualquer ano de geração da planta, que, quando avaliados em conjunto, permitem definir estratégias de venda da energia em diversos tipos de contrato. Como exemplo, com referência no contrato padrão do mercado regulado, foram simulados os primeiros 20 anos de operação da planta. A curva de redução de receita de cada ano, analisado individualmente, é apresentada na Figura 67, onde o ponto máximo de cada curva indica a melhor estratégia a ser adotada para cada ano.

Este resultado é importante para negociações anuais, como contratos no mercado livre ou com compradores diferentes ao longo do tempo. No caso do estudo aplicado ao PPA do mercado regulado, no entanto, o contrato é definido com um único valor energético prometido durante toda a duração. Desta forma, é necessário encontrar o valor que fornece a menor perda de receita acumulada ao longo do contrato.

Para isso, as regras do contrato foram simuladas para todos os cenários, considerando as perdas acumuladas no período total de contratos de 1 a 20 anos. A Figura 68 ilustra as curvas de redução de receita calculadas e os melhores resultados.

Os resultados desta simulação mostram que o melhor lance de energia em um leilão com PPA de 20 anos, no local estudado, é de 88% da geração esperada.

Esta opção representa uma perda de receita média de 8,17%, em comparação com 8,72% do contrato completo. Essa diferença equivale a cerca de 5 milhões de reais ao longo do período do contrato.

6.6 PERDA DE RECEITA

Embora seja possível identificar a alternativa ótima considerando o período total, sua receita anual não será necessariamente a melhor em todos os anos. Essa diferença, mesmo que pequena, pode existir em alguns casos. A Figura 69 ilustra essa perda de receita acumulada ao longo dos anos para cada alternativa.

Observa-se que a perda de receita ao longo do tempo tem evoluções diferentes para cada tipo de contrato, tornando a escolha do lance de energia uma questão particular e dependente da duração desejada, do custo do dinheiro e do interesse de antecipação do retorno por parte do empreendedor. As diferenças entre as receitas de cada contrato podem ser observadas em detalhe na Figura 70.

Neste gráfico a curva de 1 pu é utilizada como referência e sua receita não é apresentada. As demais curvas representam a receita incremental acumulada ao longo do contrato. Como se pode observar, nos primeiros passos de redução do valor do lance a receita aumenta, indicando menor ocorrência de cenários com penalidade por falta de entrega de energia. Nos passos finais, apesar de não haver falta de entrega, há penalidade por excesso, impactando significativamente os primeiros anos.

Comparando, por exemplo, as curvas de 1,00 e de 0,88 pu, respectivamente o contrato padrão e o de maior receita acumulada, é possível destacar a importância dos interesses paralelos do investidor, que ao optar por majorar o seu retorno terá indiretamente que postergar parte das parcelas iniciais.

6.7 PROBABILIDADES DE GERAÇÃO ANUAL

A aplicação das séries sintéticas implementadas permite também avaliar as probabilidades de geração de uma usina, ou, do ponto de vista do comprador⁶, analisar os riscos de suprimento e atendimento à demanda.

⁶ “ponto de vista do comprador”, neste caso, se refere aos agentes que dependem da energia que será gerada, como consumidores, operador, regulador e distribuidores.

Esta avaliação adotou as mesmas premissas dos casos anteriores, ou seja, do ponto de vista do gerador e com correções complementares da energia equivalente produzida por conta de penalidades contratuais.

Na Figura 71 é apresentado um exemplo da distribuição da variação da receita para os contratos de 0,85 a 0,89 pu, além do referencial de 1 pu. Esses contratos foram escolhidos por conta da facilidade de visualização de seus comportamentos. Os cenários são apenas para o primeiro ano e estão ordenados em sentido de energia crescente.

Figura 71. Distribuição de receita, por contrato
Distribuição de receita por tipo de contrato - Ano 1

Fonte: Autor.

Nos cenários com menor radiação o contrato de 1 pu recebe mais penalidades por não cumprir a entrega de energia prometida, porém, por serem pequenas, a diferença de receita para os outros contratos é pouco significativa. Por outro lado, nos melhores cenários os contratos com menor comprometimento de energia entregam muitas vezes excessos superiores a 15 % e têm redução da receita, tornando-os pouco interessantes como estratégia. Esse aspecto confirma e complementa os resultados médios apresentados na Figura 65, permitindo outra perspectiva da análise com a avaliação dos riscos da geração.

A comparação entre as curvas do gráfico anterior pode também ser realizada de forma comparativa, como na Figura 72. Neste caso são apresentadas todas as opções de contrato na faixa 0,85-0,99 pu, com valores relativos ao de 1 pu. Como se pode observar, os contratos superiores a 0,91 pu têm comportamento muito semelhante na maior parte dos cenários, diferindo ligeiramente nas séries iniciais. Os demais contratos apresentam prejuízo nos cenários bons e não têm vantagem aparente quando a radiação é menor.

Figura 72. Distribuição de receita, relativa, por contrato

Distribuição da Receita Relativa por tipo de contrato - Ano 1

Fonte: Autor.

6.8 VARIAÇÃO DA RECEITA ESPERADA

Os resultados apresentados no item anterior são exemplos dos riscos do primeiro ano de operação de uma planta. Em caso de contratos mais longos, é necessário realizar a avaliação dos riscos de todo o período contratual. Para isso, como exemplo, foram geradas as distribuições de probabilidade de ocorrência da geração por 20 anos, por valor de energia comprometida no contrato. O resultado é apresentado no gráfico da Figura 73.

Figura 73. Distribuição da redução de receita, acumulada no contrato de 20 anos

Distribuição da redução de receita, acumulada 20 anos, por lance

Fonte: Autor.

Esta simulação ilustra um caso particular de contratos muito longos, onde pode acontecer de um mesmo cenário favorecer um contrato nos primeiros anos e outro no final do período, compensando as perdas por penalidades e tornando a comparação mais complexa. Neste caso, apesar de algumas curvas se destacarem, não é clara a identificação da melhor alternativa.

A fim de facilitar a comparação entre os contratos, a Figura 74 apresenta as distribuições da redução de receita relativas ao contrato de 1 pu, também em valores acumulados para o período de 20 anos.

As influências das penalidades já foram deduzidas anteriormente e são coerentes com este gráfico, onde as principais diferenças ocorrem nas penalizações dos maiores contratos nos piores cenários e dos menores contratos nos melhores cenários.

Este gráfico tem por objetivo, além de visualizar estes comportamentos, permitir a análise da potencialidade de realização de cada contrato. Ou seja, verificar a frequência em que um lance almejado realmente representará a melhor decisão do vendedor da energia. Por exemplo, o contrato de 0,88 pu (curva em azul claro) anteriormente destacado como a maior entrega de energia, pouco se destaca como a melhor opção absoluta dos cenários.

Outra forma de análise é a visualização do histograma comparativo entre duas opções de contrato. Como exemplo, a Figura 75 e a Figura 76 apresentam, respectivamente, as relações entre as receitas para os comprometimentos de 0,88/1,00 e 0,88/0,87 pu.

Figura 75. Histograma da receita relativa entre os contratos de 0,88 e 1,00 pu

Fonte: Autor.

Figura 76. Histograma da receita relativa entre os contratos de 0,88 e 0,87 pu

Fonte: Autor.

Na comparação com o contrato de 1,00 pu (Figura 75) a opção pelo de 0,88 é bem clara, pois em um número ínfimo de cenários a decisão não seria acertada. Entretanto, no caso da Figura 76 percebe-se que em cerca de 60 % das séries o contrato de 0,87 pu é ligeiramente melhor, enquanto que no restante das séries há vantagens para o de 0,88, sendo que em algumas a vantagem tem valor mais expressivo.

7 CONCLUSÕES

Esta tese apresentou o desenvolvimento de pesquisa sobre a geração de séries sintéticas de radiação solar, realizada através do uso de modelos estocásticos e com fins de subsidiar estudos elétricos e energéticos. Como resultado principal, foi implementado e validado um modelo para a geração de cenários de radiação com base mensal do tipo PAR (1), posteriormente aplicado em caso prático na definição de estratégias para um leilão típico do mercado regulado, permitindo avaliação probabilística da receita esperada para o contrato.

Ao longo do trabalho foram obtidas diversas conclusões sobre os processos de modelagem, o tratamento de dados e a periodicidade de análise, entre outros. Esses resultados são bastante relevantes para o tema tratado e são indicados a seguir ao longo dos itens 7.1 e 7.2, onde são apresentadas, respectivamente, as considerações finais e principais contribuições originadas pela pesquisa. Em seguida são destacadas as publicações realizadas e temas para ulterior desenvolvimento.

7.1 CONSIDERAÇÕES FINAIS

A modelagem de um processo estocástico pode ser extremamente difícil do ponto de vista matemático, uma vez que há de se especificar, testar e validar um modelo ideal – e ainda inexistente – para uma representação aderente à realidade e com resíduos desprezíveis. Para tanto, este trabalho utilizou o método de Box-Jenkins, baseado em processos estocásticos conhecidos, para identificar o modelo ideal de representação da radiação solar. Esta escolha se deu por conta da metodologia se tratar da mais consagrada na literatura e ser utilizada em diversos campos do conhecimento. Em seu processo, modelos com estruturas pré-definidas foram testados e adequados aos dados registrados, a fim de produzir um modelo simples para a aplicação da geração de séries sintéticas e a consequente simulação de cenários energéticos.

Inicialmente, as estruturas utilizadas foram a autorregressiva AR, autorregressiva integrada ARI, autorregressiva com média móvel ARMA e autorregressiva integrada com média móvel ARIMA. Adicionalmente, ainda durante o desenvolvimento do método de Box-Jenkins, foram realizados estudos complementares para o tratamento de tendências e sazonalidade e para a estimativa de coeficientes pelos métodos de Yule-Walker e dos mínimos quadrados. As análises

utilizaram dados reais de localidades do nordeste brasileiro e consideraram as periodicidades horária, diária e mensal.

Na avaliação do pré-tratamento dos dados, a retirada da tendência e da sazonalidade com base nos valores médios se mostrou a melhor estratégia. Quanto ao método de estimativa dos parâmetros, foi perceptível o frequente destaque do Yule-Walker em relação aos mínimos quadrados, sendo em alguns casos o único adequado.

A estrutura ideal para os modelos não foi identificada de forma única, apresentando resultados diferentes conforme o período de análise. Para dados horários os modelos AR e ARI se destacaram, com ligeira vantagem para o segundo. Já no período diário não houve indicação clara do melhor modelo, sendo que os resultados foram diferentes tanto entre os meses de análise como entre as fontes de dados utilizadas nas simulações. A modelagem das séries mensais apresentou as conclusões mais evidentes, com a recomendação da estrutura AR.

Partindo das conclusões obtidas no estudo de estimativa de modelos e do interesse em aplicações práticas na área energética, a pesquisa se estendeu com o aprofundamento do estudo do modelo autorregressivo mensal, identificado como o mais adequado para estudos de longo prazo, e de sua variação periódica PAR, adotada usualmente em estudos hidrológicos. Para tanto, os modelos foram implementados e analisados, para diferentes ordens de atraso, através de testes comparativos do erro encontrado em cada modelo em relação aos dados reais observados e não utilizados nas calibrações.

O melhor resultado foi obtido pelo modelo autorregressivo periódico de ordem 1 – PAR (1), capaz de representar a radiação com boa precisão e de fácil implementação. Em um pequeno número de cenários, alguns meses apresentaram resultados ligeiramente melhores com o modelo de ordem 2, mas esse ganho e sua representatividade não justificam o aumento da complexidade da implementação para aplicações práticas. Os resultados do modelo autorregressivo convencional foram sempre piores.

Uma vez identificado o melhor modelo, e sua tendência e incertezas definidas, um gerador de séries sintéticas de radiação foi implementado. Essas séries foram então processadas por um modelo de contratos de compra de energia e suas penalidades, permitindo a definição de estratégias de lances para leilões de energia, a avaliação dos riscos de suprimento e a análise das incertezas das receitas futuras, entre outros.

Como resultado principal, destaca-se que a melhor estratégia de lance de energia em leilões regulados para o local estudado é o contrato de 88 % da geração esperada. Esta estratégia estima uma receita 0,6 % maior do que a oferta total durante o período do contrato, equivalente a cerca de cinco milhões de reais em uma usina de 100 MW.

O processo desenvolvido também permitiu avaliar melhor as perdas de receita devido à variação da radiação ao longo do tempo e suas consequentes penalidades. No caso estudado a perda de receita de um contrato padrão foi avaliada em 8,72 % do total esperado.

Complementarmente, as simulações realizadas estimaram a evolução da receita acumulada ao longo do contrato, para diferentes comprometimentos e de formas absoluta e relativa, permitindo melhor avaliação do fluxo de caixa conforme a alternativa escolhida.

Esta análise foi concluída com o cálculo das probabilidades de geração, tanto para anos independentes como para períodos de contratos padrão do mercado regulado, estimando os riscos de suprimento e consequentes penalidades contratuais. Com esses dados um tomador de decisão pode avaliar melhor a potencialidade de realização da alternativa escolhida como a de maior retorno esperado frente às demais opções.

Alternativamente, esta avaliação foi realizada de forma comparativa entre possíveis contratos, onde é possível, através de histogramas, a análise da probabilidade (e amplitude) de um determinado contrato apresentar melhor performance que outro. Por exemplo, na comparação entre o contrato de 88 % citado e seu principal concorrente (87 %), em cerca de 60 % das séries o segundo é ligeiramente melhor, apesar de ter vantagem pouco expressiva. Em outras palavras, o “pior” contrato (em valor médio esperado) possui maior probabilidade de ser a melhor opção para o investidor.

7.2 CONTRIBUIÇÕES

O trabalho realizado proporcionou, além dos resultados específicos apresentados no item anterior, importantes contribuições, aplicáveis tanto no setor elétrico como em questões acadêmicas mais amplas.

A primeira contribuição relevante é a conceituação da modelagem estocástica da radiação solar com fins energéticos, foco normalmente não utilizado neste tipo de estudo. Os modelos foram avaliados considerando, principalmente, os objetivos de: prever a radiação e a energia gerada

de forma probabilística; a conseguinte geração de cenários sintéticos; não necessitar de interpolações ou variáveis climáticas adicionais; e a comparação com a metodologia consolidada no setor para estudos da fonte hidrelétrica.

O trabalho também contribui na avaliação e comparação de diversos métodos para a modelagem. A comparação inicial, realizada na etapa de pré-tratamento dos dados com métodos para a remoção da tendência, não é usualmente realizada, ou ao menos apresentada e discutida, nas publicações sobre o tema. Esta pesquisa avaliou e registrou a remoção da tendência por três métodos de pré-tratamento e, adicionalmente, pela utilização de modelos periódicos, capazes de remover a tendência de forma independente em cada um de seus períodos, além de permitir melhor captura do comportamento das séries.

O estudo contribui, de forma inovadora para aplicações nacionais, identificando os modelos mais adequados para a previsão da geração com periodicidade horária e mensal. Para o período diário, apesar de não existir indicação explícita, os resultados permitiram a conclusão de que mais que uma estrutura pode ser utilizada para o período, sem grandes benefícios pela definição de alguma em particular.

A dependência temporal do modelo de previsão mensal também foi avaliada de forma original, indicando a ordem de modelo mais adequada para aplicações nos locais estudados.

Por fim, destaca-se a realização de estudo de caso, com dados reais e diretamente aplicável ao mercado, tratando de forma inovadora a expectativa de geração solar com viés probabilístico, permitindo outra perspectiva para a tomada de decisões de investimentos e de comercialização de energia.

7.3 TÓPICOS PARA ULTERIOR DESENVOLVIMENTO

O uso de cenários em aplicações que envolvam a intermitência ou a variabilidade do recurso solar, como nas aplicações práticas realizadas, é bastante relevante, porém não se limita às avaliações energéticas mensais. Os resultados desta pesquisa podem ser utilizados em diversas aplicações, com qualquer periodicidade de análise, como, por exemplo:

- Período horário: avaliação de impactos na rede elétrica (seja por conta de geração distribuída ou no sistema de transmissão), na variação de tensão, sentido de fluxo,

configuração de proteções e variação do fator de potência; no projeto e na operação de sistemas;

- Período diário: projetos de sistemas de energia não conectados à rede, que possuem vida útil altamente dependente do dimensionamento do armazenamento, cujo ciclo de carga e descarga é diário e normalmente calculado de forma determinística e simplificada; e
- Período mensal: planejamento energético, avaliação de custo/benefício da inserção da energia solar na matriz, comercialização de energia, riscos de suprimento, entre outros.

Além dos estudos aplicados, se sugere também a realização de pesquisas na modelagem dos dados, como: a simulação de outras estruturas de modelos estocásticos; a utilização de periodicidades não usuais, como quinzenal ou semanal; e a sincronização dos períodos dos modelos com as datas de solstício⁷ e de zênite solar⁸ no local de medição.

7.4 PUBLICAÇÕES REALIZADAS

Durante o período de desenvolvimento desta pesquisa foram realizadas as seguintes publicações:

- Capítulo no livro **“A Energia Solar – Tecnologia e Regulação”** (TAHAN, *et al.*, 2013);
- Artigo **“Energia Solar na Matriz Energética – Impactos Técnicos e Econômicos no Sistema Interligado Nacional”**, apresentado no *V Congresso Brasileiro de Energia Solar* (GEMIGNANI, KAGAN e CASTRO, 2014), posteriormente republicado com adaptações na *Revista Brasileira de Energia Solar* (2014);
- Artigo **“Time Series Modelling for Solar Irradiance Estimation in Northeast Brazil”**, apresentado no *6th IEEE International Conference on Renewable Energy Research and Applications* (2017) (LOURENÇO, *et al.*, 2017), premiado com o **“Best Paper Award”**;

⁷ O solstício é o momento em que o sol atinge a maior declinação em latitude, coincidente com os trópicos, e caracteriza o início do verão e do inverno.

⁸ O zênite solar se dá pela incidência dos raios solares de forma perpendicular à superfície terrestre. Esta condição ocorre duas vezes por ano para localizações entre os trópicos e nenhuma vez para latitudes superiores.

- Artigo “**Solar Radiation Synthetic Sequences and Bidding Strategies for Energy Auctions**”, apresentado no *6th IEEE International Conference on Renewable Energy Research and Applications* (2017) (GEMIGNANI, ROSTEGUI e KAGAN, 2017); e
- Artigo “**Solar Radiation Synthetic Series for Power Purchase Agreements**”, convidado para submissão ao periódico *Environmental Science and Pollution Research* (JCR 2,74), em revisão (2018).

REFERÊNCIAS⁹

- AMATO, U. et al. Markov processes and Fourier analysis as a tool to describe and simulate daily solar irradiance. **Solar Energy**, n. 37, p. 179-194, 1986.
- ANDERSON, P.; VECCHIA, A. Asymptotic results for periodic autoregressive moving average processes. **Journal of Time Series Analysis**, 1993. 1-18.
- ANEEL - AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Projeto estratégico: Modelo de otimização do despacho hidrotérmico (chamada n. 001/2008)**. Brasília, p. 18. 2008.
- ANEEL. **Edital de leilão nº 04/2016: Anexo II**. Brasília. 2016.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Energia solar fotovoltaica - Terminologia (NBR 10899)**. Rio de Janeiro: ABNT, 2013. 11 p.
- BARRETO, L. A. L.; GUERRERO-SALAZAR, P. Modelagem hidrológica para estudos de operação a longo prazo de sistemas hidrelétricos. **Revista Brasileira de Energia Elétrica**, Rio de Janeiro, v. 30, Janeiro/Junho 1975.
- BOX, G. E. P.; JENKINS, G. M. **Time Series Analysis, Forecasting and Control**. São Francisco: Holden-Day, 1976. 575 p.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis: forecasting and control**. 4^a. ed. New Jersey: John Wiley & Sons, 2008.
- BUENO, R. L. S. **Econometria de Séries Temporais**. 2^a. ed. [S.l.]: Cengage, 2011.
- COSE - UNICAMP. **Modelos de séries temporais para previsão de vazões médias mensais**. UNICAMP. Campinas, p. 49. 2010.
- DAMÁZIO, J. M. et al. **Seleção de métodos para estimação de vazões com tempos de retorno elevados**. Simpósio Brasileiro de Hidrologia. Blumenau: [s.n.]. 1983.
- EHLERS, R. S. **Análise de séries temporais**. 5^a. ed. [S.l.]: [s.n.], 2009.
- ELETROBRAS; CEPEL. **Modelo NEWAVE - Manual de referência**. [S.l.], p. 103.

⁹ De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

ENDERS, W. **Applied Econometric Time Series**. 4^a. ed. [S.l.]: John Willey & Sons, 2014.

FEIJÓ, C. Conceitos Básicos de Séries Temporais, 2010. Disponível em: <http://www.icad.puc-rio.br/cfeijo/pdf/revis%C3%A3o%20b%C3%A1sica%20s%C3%A9ries%20temporais_material%20de%20apoio_curso%20teoria%20macroeconomica_PPGE%20UFF.pdf>.

FORTUNA, L.; NUNNARI, S.; GALLO, A. **A Typical Day Based Approach To Detrend Solar Radiation Time Series**. Proceedings of the 3rd ACM International Workshop on Multimedia Analysis for Ecological Data. [S.l.]: ACM. 2014. p. 25-30.

GEMIGNANI, M. M. F.; KAGAN, N.; CASTRO, G. M. **Energia solar na matriz energética - impactos técnicos e econômicos no sistema interligado nacional**. V Congresso Brasileiro de Energia Solar. Recife: [s.n.]. 2014. p. 10.

GEMIGNANI, M. M. F.; ROSTEGUI, G. J.; KAGAN, N. **Solar Radiation Synthetic Sequences and Bidding Strategies for Energy Auctions**. 6th IEEE International Conference on Renewable Energy Research and Applications. San Diego: [s.n.]. 2017.

GOH, T. N.; TAN, K. J. Stochastic modeling and forecasting of solar radiation data. **Solar Energy**, v. 19, n. 6, p. 755-757, 1977.

GORDON, J. M.; REDDY, T. A. Time series analysis of daily horizontal solar radiation. **Solar Energy**, v. 41, n. 3, p. 215-226, 1988.

GUERRERO-SALAZAR, P. et al. **Geração de séries hidrológicas mensais para estudos energéticos**. IV SNPTEE - Seminário Nacional de Produção e Transmissão de Energia Elétrica. Rio de Janeiro: [s.n.]. 1977.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. 5^a. ed. [S.l.]: Mcgraw Hill, 2011.

INMAN, R. H.; PEDRO, H. T. C.; COIMBRA, C. F. M. Solar forecasting methods for renewable energy integration. **Progress in Energy and Combustion Science**, v. 39, n. 6, p. 535-576, Dezembro 2013.

JONES, R.; BRELSFORD, W. Time series with periodic structure. **Biometrika**, 1967. 403-408.

KELMAN, J. **A stochastic model for daily precipitation**. World congress on water resources. Mexico: [s.n.]. 1979. p. 12.

KELMAN, J. A stochastic model for daily streamflow. **Journal of Hydrology**, Amsterdam, Ed. 47 1980. 235-249.

KELMAN, J. **Daily synthetic streamflow sequences and the evaluation of dam safety**. Fall annual meeting of the american geophysical union. San Francisco: [s.n.]. 1984.

KELMAN, J. **Uso de séries sintéticas no planejamento e operação de sistemas hidrotérmicos**. Seminário latinoamericano de aproveitamento dos recursos hídricos. Medellin: [s.n.]. 1987.

KELMAN, J. et al. **Revisão do conceito de energia firme através do uso de séries hidrológicas sintéticas**. V SNPTEE - Seminário nacional de produção e transmissão de energia elétrica. Recife: [s.n.]. 1979.

KELMAN, J.; DAMAZIO, J. M. **Synthetic hydrology and spillway design**. 20th Congress of the international association for hydraulic research. Moscou: [s.n.]. 1983.

KELMAN, J.; DAMÁZIO, J. M.; COSTA, J. P. **A multivariate synthetic daily streamflow generator**. 4th international hydrology symposium. Fort Collins: [s.n.]. 1985.

KELMAN, J.; OLIVEIRA, G. C.; PEREIRA, M. V. F. **Geração de séries hidrológicas - método de desagregação**. V SNPTEE - Seminário nacional de produção e transmissão de energia elétrica. Recife: [s.n.]. 1979.

LI, Y.; SU, Y.; SHU, L. An ARMAX model for forecasting the power output of a grid connected photovoltaic system. **Renewable Energy**, v. 66, p. 78-89, Junho 2014.

LJUNG, L. **System identification**. Boston: Birkhäuser, v. 1, 1998.

LOURENÇO, L. F. N. et al. **Time Series Modelling for Solar Irradiance Estimation in Northeast Brazil**. 6th IEEE International Conference on Renewable Energy Research and Applications. San Diego: [s.n.]. 2017.

MACEIRA, M. E. P. et al. **Modelagem estatística de ventos para a utilização nos modelos de planejamento e operação**. XXI SNPTEE - Seminário nacional de produção e transmissão de energia elétrica. Florianópolis: [s.n.]. 2011.

MACEIRA, M. E. P.; KELMAN, J.; DAMÁZIO, J. M. **Utilização de modelos PAR para simulação e previsão de séries hidrológicas mensais**. VII Simpósio brasileiro de hidrologia e recursos hídricos. Salvador: [s.n.]. 1987.

MACEIRA, M. E. P.; PENNA, D. D. J.; DAMÁZIO, J. M. Geração de cenários sintéticos de energia e vazão para o planejamento da operação energética. **Cadernos do IME - Série estatística**, Rio de Janeiro, v. 21, p. 11-35, 2006.

MACEIRA, M. et al. **Modelagem estatística de ventos para a utilização nos modelos de planejamento e operação**. XXI SNPTTE - Seminário nacional de produção e transmissão de energia elétrica. Florianópolis: [s.n.].

MADSEN, H. **Time series analysis**. 1ª. ed. [S.l.]: Chapman & Hall, 2007.

MINISTÉRIO DE MINAS E ENERGIA; EMPRESA DE PESQUISA ENERGÉTICA. **Acompanhamento das Medições Anemométricas**. Disponível em: <<http://sistemas.epe.gov.br/ama>>. Acesso em: 21 Junho 2016.

MINISTÉRIO DE MINAS E ENERGIA; EMPRESA DE PESQUISA ENERGÉTICA. **Plano Decenal de Expansão de Energia 2024**. Brasília, p. 466. 2015.

MINISTÉRIO DE MINAS E ENERGIA; EMPRESA DE PESQUISA ENERGÉTICA. **Plano Decenal de Expansão de Energia 2026**. Brasília, p. 271. 2017.

MUSTACCHI, C.; CENA, V.; ROCCHI, M. Stochastic simulation of hourly global radiation sequences. **Solar Energy**, v. 23, n. 1, p. 47-51, 1979.

NASCIMENTO, C. E. S.; KELMAN, J. Um modelo para geração estocástica de chuvas diárias. **RBE - Revista Brasileira de Engenharia**, v. 13, n. 2, p. 19-36, dezembro 1995.

NOBRE, A. M. **Short-term Solar Irradiance Forecasting for Photovoltaic Applications in Tropical Regions**. Universidade Federal de Santa Catarina. Florianópolis. 2015.

PEREIRA, M. V. F. et al. Stochastic streamflow models for hydroelectric systems. **Water Resources Research**, Março 1984. 379-390.

PROMON. **Conceptos básicos de planeamiento operativo**. Projeto Parseica. [S.l.]. 1992.

SIQUEIRA, A. N. **Geração sintética da irradiação solar diária no Brasil para aplicações energéticas**. Tese de doutorado - Universidade Federal de Pernambuco. Recife, p. 126. 2008.

SULAIMAN, M. Y. et al. Analysis of residuals in daily solar radiation time series. **Renewable Energy**, v. 11, n. 1, p. 97-105, Maio 1997.

TAHAN, C. M. V. et al. **A Energia Solar: Tecnologia e Regulação**. São Paulo: Cia das Palavras, 2013.

VECCHIA, A. Maximum likelihood estimation for periodic autoregressive moving average models. **Technometrics**, 1985. 375-384.

ANEXO I – REMOÇÃO DA TENDÊNCIA PELA MÉDIA DIÁRIA

Figura 77. Tendências mensais

Fonte: Autor.

Figura 78. Dados de 2015 retirados de tendência

Fonte: Autor.

ANEXO II – REMOÇÃO DA TENDÊNCIA PELO MÁXIMO DIÁRIO

Figura 79. Tendências para todos os meses de acordo com o método 2

Fonte: Autor.

ANEXO III – REMOÇÃO DA TENDÊNCIA PELA MÉDIA MENSAL

Figura 80. Retirada de tendência pela média mensal

Fonte: Autor.

ANEXO IV – INFLUÊNCIA DO TAMANHO DO CONJUNTO DE DADOS NA ESTIMATIVA DOS PARÂMETROS DO MODELO AR

Na modelagem estocástica, após a escolha de uma estrutura de modelo, parte-se para a etapa de estimativa dos parâmetros, onde um conjunto de dados é processado pelo método Yule-Walker para que os coeficientes e a variância do ruído do modelo sejam obtidos.

Este conjunto de dados, discutido em trabalhos como (LJUNG, 1998), deve ser bem escolhido, pois se não houver informações suficientes sobre o processo podem ser geradas estimativas ruins dos parâmetros do modelo.

Com esta questão em mente, foi realizado estudo complementar para verificação da influência do tamanho do conjunto de dados na estimativa dos parâmetros do modelo, onde aplicou-se o seguinte procedimento: primeiramente, apenas um único conjunto de dados foi utilizado, correspondendo à série de radiação registrada para o ano de 2015 e, a cada passo, um conjunto de dados de um ano foi adicionado, correspondendo às séries registradas para o ano de 2014 e assim por diante.

O resultado é exibido nas figuras seguintes, por mês e por coeficiente.

A análise dos gráficos permite notar que o aumento do número de conjunto de dados leva à diminuição do módulo dos coeficientes a_2 , a_3 , a_4 e a_5 , enquanto que o coeficiente a_1 tende a uma estabilização em torno de um valor diferente de zero. Isto significa a interpretação, pelo modelo, de que apenas o dia imediatamente anterior ao desejado tem influência na previsão.

Pode-se notar também que os meses outubro e dezembro necessitam de apenas quatro conjuntos de dados para os coeficientes do modelo se estabilizarem em valores próximos aos finais. Os meses que apresentam o segundo melhor desempenho neste quesito são junho, agosto e novembro, com seis conjuntos de dados. Já os demais meses necessitaram de 10 a 15 conjuntos para a estabilização no valor final.

Além disso, como se busca o modelo com o menor número de coeficientes possível, pode-se notar nas mesmas figuras que o valor do coeficiente a_1 é mais significativo que os demais. Assim, para todos os meses é suficiente o modelo autorregressivo de primeira ordem.

Figura 81. Influência do tamanho do conjunto de dados na estimativa do modelo, JAN-MAR

Fonte: Autor.

Figura 82. Influência do tamanho do conjunto de dados na estimativa do modelo, ABR-JUN

Fonte: Autor.

Figura 83. Influência do tamanho do conjunto de dados na estimativa do modelo, JUL-SET

Fonte: Autor.

Figura 84. Influência do tamanho do conjunto de dados na estimativa do modelo, OUT-DEZ

Fonte: Autor.

ANEXO V – ANÁLISE DAS SÉRIES GERADAS PELO MÉTODO DOS MÍNIMOS QUADRADOS

A análise das séries modeladas através do método dos mínimos quadrados foi realizada através de comparações entre as séries de radiação diária geradas com remoção de tendência pela média diária e pela média mensal, para cada uma das quatro estruturas (AR, ARI, ARMA e ARIMA) e para cada mês.

A seguir são apresentados todos os gráficos mensais. Nessas figuras está plotada à esquerda uma série sintética gerada por cada um dos métodos propostos, a média diária registrada dos 20 anos e os valores máximo e mínimo do histórico. À direita encontram-se todos os 20 conjuntos de dados históricos em curvas amarelas e uma série sintética gerada por cada método avaliado, assim como os valores máximo e mínimo do histórico para verificação da consistência dos dados.

Análise das séries – Estrutura AR

Figura 86. Métodos de geração de séries horárias com MMQ, estrutura AR (fevereiro)

Fonte: Autor.

Figura 87. Métodos de geração de séries horárias com MMQ, estrutura AR (março)

Fonte: Autor.

Figura 88. Métodos de geração de séries horárias com MMQ, estrutura AR (abril)

Fonte: Autor.

Figura 89. Métodos de geração de séries horárias com MMQ, estrutura AR (maio)

Fonte: Autor.

Figura 90. Métodos de geração de séries horárias com MMQ, estrutura AR (junho)

Fonte: Autor.

Figura 91. Métodos de geração de séries horárias com MMQ, estrutura AR (julho)

Fonte: Autor.

Figura 92. Métodos de geração de séries horárias com MMQ, estrutura AR (agosto)

Fonte: Autor.

Figura 93. Métodos de geração de séries horárias com MMQ, estrutura AR (setembro)

Fonte: Autor.

Figura 94. Métodos de geração de séries horárias com MMQ, estrutura AR (outubro)

Fonte: Autor.

Figura 95. Métodos de geração de séries horárias com MMQ, estrutura AR (novembro)

Fonte: Autor.

Figura 96. Métodos de geração de séries horárias com MMQ, estrutura AR (dezembro)

Fonte: Autor.

Estrutura ARMA

Figura 97. Métodos de geração de séries horárias com MMQ, estrutura ARMA (janeiro)

Fonte: Autor.

Figura 98. Métodos de geração de séries horárias com MMQ, estrutura ARMA (fevereiro)

Fonte: Autor.

Figura 99. Métodos de geração de séries horárias com MMQ, estrutura ARMA (março)

Fonte: Autor.

Figura 100. Métodos de geração de séries horárias com MMQ, estrutura ARMA (abril)

Fonte: Autor.

Figura 101. Métodos de geração de séries horárias com MMQ, estrutura ARMA (maio)

Fonte: Autor.

Figura 102. Métodos de geração de séries horárias com MMQ, estrutura ARMA (junho)

Fonte: Autor.

Figura 103. Métodos de geração de séries horárias com MMQ, estrutura ARMA (julho)

Fonte: Autor.

Figura 104. Métodos de geração de séries horárias com MMQ, estrutura ARMA (agosto)

Fonte: Autor.

Figura 105. Métodos de geração de séries horárias com MMQ, estrutura ARMA (set)

Fonte: Autor.

Figura 106. Métodos de geração de séries horárias com MMQ, estrutura ARMA (out)

Fonte: Autor.

Figura 107. Métodos de geração de séries horárias com MMQ, estrutura ARMA (nov)

Fonte: Autor.

Figura 108. Métodos de geração de séries horárias com MMQ, estrutura ARMA (dez)

Fonte: Autor.

Análise das séries – Estrutura ARI

Figura 109. Métodos de geração de séries horárias com MMQ, estrutura ARI (janeiro)

Fonte: Autor.

Figura 110. Métodos de geração de séries horárias com MMQ, estrutura ARI (fevereiro)

Fonte: Autor.

Figura 111. Métodos de geração de séries horárias com MMQ, estrutura ARI (março)

Fonte: Autor.

Figura 112. Métodos de geração de séries horárias com MMQ, estrutura ARI (abril)

Fonte: Autor.

Figura 113. Métodos de geração de séries horárias com MMQ, estrutura ARI (maio)

Fonte: Autor.

Figura 114. Métodos de geração de séries horárias com MMQ, estrutura ARI (junho)

Fonte: Autor.

Figura 115. Métodos de geração de séries horárias com MMQ, estrutura ARI (julho)

Fonte: Autor.

Figura 116. Métodos de geração de séries horárias com MMQ, estrutura ARI (agosto)

Fonte: Autor.

Figura 117. Métodos de geração de séries horárias com MMQ, estrutura ARI (set)

Fonte: Autor.

Figura 118. Métodos de geração de séries horárias com MMQ, estrutura ARI (out)

Figura 119. Métodos de geração de séries horárias com MMQ, estrutura ARI (nov)

Fonte: Autor.

Figura 120. Métodos de geração de séries horárias com MMQ, estrutura ARI (dez)

Fonte: Autor.

Análise das séries – Estrutura ARIMA

Figura 121. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (janeiro)

Fonte: Autor.

Figura 122. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (fev)

Fonte: Autor.

Figura 123. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (março)

Fonte: Autor.

Figura 124. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (abril)

Fonte: Autor.

Figura 125. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (maio)

Fonte: Autor.

Figura 126. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (junho)

Fonte: Autor.

Figura 127. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (julho)

Fonte: Autor.

Figura 128. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (agosto)

Fonte: Autor.

Figura 129. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (set)

Fonte: Autor.

Figura 130. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (out)

Fonte: Autor.

Figura 131. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (nov)

Fonte: Autor.

Figura 132. Métodos de geração de séries horárias com MMQ, estrutura ARIMA (dez)

Fonte: Autor.

ANEXO VI – ANÁLISE DAS SÉRIES DE RADIAÇÃO HORÁRIA (MMQ)

A partir das estimativas realizadas para as estruturas AR, ARI, ARMA e ARIMA, com remoção de tendência pela média e pela máxima diárias, para a radiação horária, foram desenvolvidos modelos mensais para cada combinação estrutura X tendência. As comparações das séries sintéticas geradas para cada modelo com os dados históricos são apresentadas nas figuras seguintes.

Nesse conjunto de figuras, as curvas do histórico estão plotadas em amarelo, com seus valores médio e máximo indicados por linhas pretas, e as curvas dos métodos 1 e 2, respectivamente, em azul e rosa.

Figura 133. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura AR

Fonte: Autor.

Figura 134. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura AR

Fonte: Autor.

Figura 135. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARI

Fonte: Autor.

Figura 136. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARI

Fonte: Autor.

Figura 137. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARMA

Fonte: Autor.

Figura 138. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARMA.

Fonte: Autor.

Figura 139. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARIMA

Fonte: Autor.

Figura 140. Comparação de métodos para a geração de séries sintéticas horárias de radiação solar com estrutura ARIMA

Fonte: Autor.