



Carolina Nascimento Nogueira Lima

**Estimação do impacto do El Niño/La Niña na
intensidade dos ventos do Nordeste brasileiro
utilizando os Modelos GAS**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio

Orientador : Prof. Cristiano Augusto Coelho Fernandes

Co-Orientador: Prof. Gutemberg Borges França

Rio de Janeiro
Julho de 2014



Carolina Nascimento Nogueira Lima

**Estimação do impacto do El Niño/La Niña na
intensidade dos ventos do Nordeste brasileiro
utilizando os Modelos GAS**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela comissão examinadora abaixo assinada.

Prof. Cristiano Augusto Coelho Fernandes

Orientador

Departamento de Engenharia Elétrica — PUC-Rio

Prof. Gutemberg Borges França

Co-Orientador

Departamento de Meteorologia - UFRJ

Prof. Adrian Heringer Pizzinga

Departamento de Estatística - UFF

Prof. Ana Maria Bueno Nunes

Departamento de Meteorologia - UFRJ

Prof. José Eugenio Leal

Coordenador do Centro Técnico Científico — PUC-Rio

Rio de Janeiro, 11 de Julho de 2014

Todos os direitos reservados. Proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Carolina Nascimento Nogueira Lima

Graduou-se em Estatística na Escola Nacional de Ciências Estatísticas - ENCE/IBGE em 2011.

Ficha Catalográfica

Lima, Carolina Nascimento Nogueira

Estimação do impacto do El Niño/La Niña na intensidade dos ventos do Nordeste brasileiro utilizando os Modelos GAS / Carolina Nascimento Nogueira Lima; orientador: Cristiano Augusto Coelho Fernandes; co-orientador: Gutemberg Borges França. — Rio de Janeiro : PUC-Rio, Departamento de Engenharia Elétrica, 2014.

v., 91 f: il. ; 29,7 cm

1. Dissertação (Mestrado em Engenharia Elétrica) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Inclui referências bibliográficas.

1. Engenharia Elétrica – Tese. 2. Energia Eólica. 3. Vento. 4. El Niño/La Niña. 5. Modelos GAS. 6. Nordeste do Brasil. I. Fernandes, Cristiano Augusto Coelho. II. França, Gutemberg Borges. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Agradecimentos

Primeiramente agradeço a Deus por ter me dado condições de estar concluindo mais essa etapa da minha vida e por me fazer acreditar que não devemos nunca desistir de nossos objetivos.

Ao meu grande amor, Rodrigo, por todo carinho, dedicação, por estar incondicionalmente ao meu lado e pela paciência que teve durante todo esse período. À minha família, em especial meus pais, que mesmo distantes, me apoiaram em todos os momentos nessa caminhada e se sacrificaram para que eu estivesse realizando meus sonhos. À Lucia por todo auxílio e conselhos. Amo muito vocês.

Ao meu orientador Cristiano Fernandes, por todos os ensinamentos tanto em suas disciplinas quanto em minha dissertação, e por toda dedicação e incentivo ao longo desse tempo.

Ao meu co-orientador Gutemberg França, pela atenção e por agregar conhecimentos de Meteorologia ao meu trabalho, os quais foram imprescindíveis para a compreensão do tema.

Um agradecimento especial à Gilson Matos por toda ajuda e atenção que deu a mim durante o desenvolvimento dessa dissertação. Além da enorme contribuição para o meu trabalho.

Aos amigos Daiane Santos, Adriano Koshiyama, Luiz Aleixo, Caroline Ponce e Taíse Lyra pelas palavras de incentivo e companheirismo de sempre.

À Sandra Canton por ter me recomendado para o mestrado da PUC e por me estimular a seguir em frente.

À CAPES e à PUC-Rio, pelos auxílios concebidos, sem os quais esse trabalho não poderia ter sido realizado.

Este trabalho foi parcialmente desenvolvido com o apoio do projeto de P&D ANEEL PD-7625-0001/2013.

Resumo

Lima, Carolina Nascimento Nogueira; Fernandes, Cristiano Augusto Coelho; França, Gutemberg Borges. **Estimação do impacto do El Niño/La Niña na intensidade dos ventos do Nordeste brasileiro utilizando os Modelos GAS**. Rio de Janeiro, 2014. 91p. Dissertação de Mestrado — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

A energia eólica é hoje uma das mais promissoras fontes de energia do mundo por ser limpa e abundante. O estudo de fenômenos que estão relacionados com alterações na circulação atmosférica, como o El Niño, são de extrema importância pela possibilidade de afetarem a geração eólica. A fim de explorar o possível efeito de tais fenômenos nos ventos da região Nordeste do Brasil, a qual possui a maior capacidade eólica instalada, é realizada uma análise estatística para a quantificação desse efeito através dos modelos *Generalized Autoregressive Score* (GAS). Estes permitem a modelagem de séries temporais para diferentes distribuições de probabilidade. Nesse estudo, o GAS é aplicado às séries de velocidade do vento a partir das distribuições Weibull, Gama e Beta. Os resultados mostram que a distribuição Beta produz o melhor ajuste quando se considera o efeito do El Niño através da Temperatura da Superfície do Mar (TSM), na velocidade do vento de uma usina eólica no Nordeste do Brasil. Em particular estes mostram que, o maior impacto do El Niño ocorre nos meses de abril e maio, onde para um aumento de 1°C da TSM é observada uma variação de aproximadamente 1.3% na velocidade do vento.

Palavras-chave

Energia Eólica ; Vento ; El Niño/La Niña ; Modelos GAS ; Nordeste do Brasil

Abstract

Lima, Carolina Nascimento Nogueira; Fernandes, Cristiano Augusto Coelho (advisor); França, Gutemberg Borges. **Estimation of the impact of the El Niño/La Niña in the intensity of the winds in Northeast Brazil using the GAS models.** Rio de Janeiro, 2014. 91p. Msc. Dissertation — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Wind energy is now one of the most promising energy sources of the world being both clean and abundant. The study of phenomena that are related to changes in atmospheric circulation, such as El Niño, are extremely important for its ability to affect wind generation. In order to explore the possible effect of such phenomena in the winds of the Northeast region of Brazil which has the largest installed wind capacity, a statistical analysis is performed to quantify this effect through the Generalized Autoregressive Score (GAS) models. These allows the modeling of time series for different probability distributions. In this study, GAS models are applied to the wind speed series using Weibull, Gamma and Beta distributions. The results shows that the Beta distribution produces the best fit when considering the effect of El Niño through Sea Surface Temperature (SST) in the wind speed of an wind farm in the Northeast of Brazil. In particular these show that the greatest impact of the El Niño occurs in the months of April and May, where for an increase of 1°C of SST is observed a variation of approximately 1.3% in wind speed.

Keywords

Wind Energy ; Wind ; El Niño/La Niña ; GAS Models ; Northeast of Brazil.

Sumário

1	Introdução	11
1.1	Objetivo	12
1.2	Estrutura da dissertação	13
2	Circulação Atmosférica e Regime dos Ventos no Nordeste brasileiro	14
2.1	Caracterização da área de estudo	14
2.2	Circulação Atmosférica no Nordeste	15
3	O fenômeno El Niño	17
3.1	Oscilação Sul	17
3.2	Métodos para a identificação dos períodos e quantificação de El Niño e La Niña	21
3.3	Episódios de El Niño e La Niña (1950 – 2012)	24
3.4	Relação entre os fenômenos El Niño e Dipolo do Atlântico	25
4	Revisão Bibliográfica	26
5	Metodologia	29
5.1	Modelo GAS	29
5.2	Estimação e Inferência	33
5.3	Parametrização	34
5.4	Modelo Univariado WEIBULL	35
5.5	Modelo Univariado GAMA	37
5.6	Modelo Univariado BETA	40
5.7	Critérios de seleção e medidas de aderência do modelo	42
5.8	Teste de Diebold-Mariano para comparação entre previsões	43
5.9	Resíduos do modelo	44
5.10	Mensuração do efeito das variáveis exógenas nos modelos GAS	45
6	Resultados	48
6.1	Análise descritiva	50
6.2	Modelos Univariados GAS	59
7	Conclusões e Extensões	77
A	Apêndice A	86
A.1	Modelo Gama - SARIMA com os parâmetros λ e α variantes no tempo	86

Lista de figuras

2.1	Área de estudo.	14
2.2	Curvas sazonais normalizadas de ventos de uma usina eólica e vazão de um posto natural.	15
3.1	Padrão de circulação atmosférica em anos neutros. Fonte: Adaptado de http://www.pmel.noaa.gov/tao/el_nino .	19
3.2	Padrão de circulação atmosférica em anos de El Niño. Fonte: Adaptado de http://www.pmel.noaa.gov/tao/el_nino .	20
3.3	Padrão de circulação atmosférica em anos de La Niña. Fonte: Adaptado de http://www.pmel.noaa.gov/tao/el_nino .	21
3.4	Região Niño no Oceano Pacífico Equatorial Fonte: www.cpc.ncep.noaa.gov	22
6.1	Fluxograma das etapas realizadas na modelagem.	49
6.2	Médias anuais de velocidade do vento para a usina do Nordeste para o período de 1982 a 2012.	51
6.3	Valores mensais médios para a velocidade do vento no período de 1982 a 2012.	51
6.4	Rosa dos ventos da usina eólica.	52
6.5	Série mensal e histograma da velocidade do vento na usina eólica no período de 01/1982 a 12/2012.	52
6.6	Box Plot das médias horárias por mês da velocidade do vento para a usina eólica entre 01/1982 e 12/2012.	53
6.7	Periodograma da série mensal da velocidade do vento entre 01/1982 e 12/2012.	54
6.8	(a) Box Plot das médias mensais de velocidade do vento; (b) Função de Autocorrelação; (c) Função de Autocorrelação Parcial; (d) Função de autocorrelação da série ao quadrado.	54
6.9	Série mensal de Temperatura da Superfície do Mar na região Niño 3.4.	55
6.10	(a) Box Plot das médias mensais de TSM da região Niño 3.4; (b) Função de Autocorrelação; (c) Função de Autocorrelação Parcial; (d) Função de autocorrelação da série ao quadrado.	55
6.11	Correlação cruzada entre velocidade do vento e TSM na região Niño 3.4.	56
6.12	Médias horárias da velocidade do vento por mês condicionada a fase de ENOS para o período de 01/1982 a 12/2012.	57
6.13	Boxplot das médias horárias da velocidade do vento por mês condicionada a fase de ENOS para o período de 01/1982 a 12/2012.	57
6.14	Ajuste da série de vento para período de estimação do Modelo 1 Weibull-SARIMA.	60
6.15	FAC e histograma dos resíduos quantílicos do Modelo 1 - Weibull-SARIMA.	61

6.16	FAC dos resíduos quantílicos do Modelo 1 - Weibull-SARIMA.	61
6.17	Ajuste da série de vento para período de estimação do Modelo 1 - Gama-SARIMA.	62
6.18	FAC e Histograma dos resíduos quantílicos do Modelo 1 - Gama-SARIMA.	64
6.19	FAC dos resíduos do Modelo 1 - Gama-SARIMA.	64
6.20	Ajuste da série de vento para período de estimação do Modelo 1 - Gama-CNO com o parâmetro de escala variante em t.	67
6.21	Componente sazonal e autoregressiva do ajuste da série de vento para período de estimação do Modelo 1 - Gama-CNO.	67
6.22	FAC e Histograma dos resíduos quantílicos do Modelo 1 - Gama-CNO.	68
6.23	Ajuste da série de vento para período de estimação do Modelo 1 - Beta-SARIMA.	69
6.24	Ajuste da série de vento para período de estimação do Modelo 2 - Beta-SARIMA.	70
6.25	FAC e Histograma dos resíduos quantílicos do Modelo 2 - Beta-SARIMA.	71
6.26	FAC dos resíduos quantílicos do Modelo 2 - Beta-SARIMA.	71
6.27	Efeito da variável exógena TSM na velocidade do vento para diferentes variações mensais de temperatura - Modelo Beta-SARIMA.	72
6.28	Ajuste da série de vento para período de estimação do Modelo 1 - Beta-CNO.	73
6.29	FAC e Histograma dos resíduos quantílicos do Modelo 2 - Beta-CNO.	74
6.30	Componente sazonal e autoregressiva do ajuste da série de vento para período de estimação do Modelo 1 - Beta-CNO.	74
6.31	Efeito estimado da variável exógena TSM na velocidade do vento para diferentes variações mensais de temperatura - Modelo Beta-CNO.	75
A.1	FAC dos resíduos quantílicos Gama-SARIMA com dois parâmetros variantes no tempo.	89

Lista de tabelas

3.1	Episódios El Niño e La Niña. Fonte: NOAA, 2014.	24
6.1	Tabela de Contingência ENOS x Vento.	58
6.2	Medidas de ajuste do modelo Weibull-SARIMA.	60
6.3	P-valor dos testes para os resíduos quantílicos do modelo Weibull-SARIMA.	61
6.4	Critérios de adequação dos modelos Gama-SARIMA.	63
6.5	Medidas de ajuste dos modelos Gama-SARIMA.	63
6.6	P-valor dos testes para os resíduos quantílicos dos modelos Gama-SARIMA.	64
6.7	Critérios de adequação dos modelos Gama-CNO.	66
6.8	Medidas de ajuste dos modelos Gama-CNO.	66
6.9	P-valor dos testes para os resíduos dos modelos Gama-CNO.	68
6.10	Critérios de adequação dos modelos Beta-SARIMA.	70
6.11	Medidas de ajuste dos modelos Beta-SARIMA.	70
6.12	P-valor dos testes para os resíduos dos modelos Beta-SARIMA.	71
6.13	Critérios de adequação dos modelos Beta-CNO.	73
6.14	Medidas de ajuste dos modelos Beta-CNO.	73
6.15	P-valor dos testes para os resíduos dos modelos Beta-CNO.	74
A.1	Estimativas dos parâmetros dos modelos Gama-SARIMA e erros padrão.	90
A.2	Estimativas dos parâmetros dos modelos Gama-CNO e erros padrão.	90
A.3	Estimativas dos parâmetros dos modelos Beta-SARIMA e erros padrão.	91
A.4	Estimativas dos parâmetros dos modelos Beta-CNO e erros padrão.	91

1

Introdução

A energia eólica é hoje uma das mais promissoras fontes de energia do mundo, por ser limpa e abundante. Esta sofreu grande avanço nos últimos anos juntamente com outras fontes alternativas devido à necessidade de se aumentar a oferta de energia e ao mesmo tempo reduzir os impactos ambientais oriundos de outras fontes.

O mercado global de energia eólica atual é liderado principalmente pela Europa, Estados Unidos e China. A maior capacidade de energia eólica instalada no mundo em 2012 encontrava-se na China (75.324 MW), seguida por Estados Unidos (60.007 MW) e Alemanha (31.308 MW). O Brasil encontra-se na 16ª posição e lidera o *ranking* na América Latina (GWEC, 2013).

Um dos países com maior capacidade disponível de geração de energia elétrica através da força dos ventos é o Brasil, principalmente pelas suas características climáticas e topográficas favoráveis à implantação de projetos eólicos. O mesmo possui atualmente 148 usinas eólicas com uma capacidade instalada atual de 3.600 MW, segundo ABEEólica (2014).

O fortalecimento da energia eólica no Brasil nos últimos anos se deve principalmente à implantação do Programa de Incentivos às Fontes Alternativas de Energia Elétrica (PROINFA) em 2004, o qual busca o aumento da participação de energias alternativas, em especial a eólica, na Matriz Energética Brasileira (MME, 2013). Com isso, o programa vem impulsionando o crescimento das fontes alternativas através de investimentos neste âmbito, com apoio financeiro do Banco Nacional de Desenvolvimento (BNDES), a fim de tornar a matriz energética mais limpa e aumentar a segurança do fornecimento de energia.

Em 2011, a participação da energia eólica na Matriz Energética Brasileira era de apenas 0,9% segundo o Balanço Energético Nacional (2013), mas esta sofreu um aumento significativo em 2013, onde a mesma alcançou 3% de participação (Boletim Mensal de Dados do setor Eólico, 2014). Além disso, através do Plano Decenal de Energia do Governo Brasileiro (PDE 2021) aprovado no início de 2013, o governo planeja alcançar uma capacidade instalada de 16.000 MW em 2021 (GWEC, 2013).

O maior potencial eólico brasileiro disponível está situado nas regiões Nordeste e Sul. A região Nordeste é a que possui maior capacidade instalada e se destaca devido aos ventos alísios constantes que sopram de leste e as brisas marítimas na costa litorânea. A mesma dispunha de 89 parques eólicos em 2013 e possui uma das melhores jazidas do mundo (Boletim Mensal de Dados do Setor Eólico, 2013). Além disso, a região apresenta ventos de velocidade acima de 7 m/s, definida como aproveitável para a produção de energia (Grubb & Meyer, 1993).

Com o aumento da importância das fontes limpas e sustentáveis de energia, decorrente dos avanços tecnológicos ocorridos nos últimos anos, vieram também o aumento da oferta de energia e a intensificação dos impactos ambientais, entre eles a emissão de gases poluentes como o dióxido de carbono (CO_2) na atmosfera devido à queima de combustíveis fósseis por fontes de energia não renováveis. Este fato vem contribuindo para o aquecimento global e estudos mostram que o mesmo pode afetar o potencial energético das fontes de energia renováveis (Schaeffer *et al.*, 2008). Com isso, as fontes de energia limpa, como a eólica vem ganhando seu espaço, uma vez que sua potência instalada no Brasil hoje, é capaz de reduzir a emissão de gases CO_2 em 3.095.574 T/ano (ABEEólica, 2014).

Essas e outras variações climáticas globais e regionais podem interferir no regime dos ventos, influenciando assim a geração de energia eólica em diferentes regiões.

Estudos de alguns fenômenos que estão relacionados com alterações na circulação atmosférica, como o El Niño e o La Niña, são de extrema importância pois eles afetam o padrão da mesma. O Nordeste brasileiro é uma região que sofre influência dessas anomalias, principalmente na precipitação, fato este já conhecido e abordado em muitos estudos (Sousa, 2003; Lopes *et al.*, 2007).

Assim, a motivação da nossa investigação dá-se em função da importância de se quantificar o possível efeito do El Niño e La Niña no regime de ventos e conseqüentemente na produção de energia eólica no Nordeste brasileiro, a qual pode ser influenciada positivamente ou negativamente em anos de ocorrência destes fenômenos, além de, como previamente mencionado, ser a região de maior capacidade instalada.

1.1

Objetivo

O objetivo dessa dissertação é a investigação, através de modelos estatísticos, dos efeitos de El Niño/La Niña, também conhecido como El Niño

Oscilação Sul (ENOS), na intensidade dos ventos na região nordeste do Brasil. A modelagem utilizada é a proposta por Creal *et al.* (2008), modelos *Generalized Autoregressive Score* (GAS), que serão aplicados às séries de vento, os quais permitem modelar séries univariadas ou multivariadas com características como: não-gaussianidade, não-linearidade e não-estacionariedade. A utilização dos modelos GAS nesse contexto é justificada, uma vez que é sabido que a série de velocidade do vento não possui distribuição Gaussiana, como abordado por Brown *et al.* (1984).

Assim, através dos modelos da classe GAS busca-se o melhor ajuste da velocidade do vento para determinadas regiões no nordeste. Isto foi realizado através da incorporação, no modelo, da variável exógena relacionada ao El Niño/La Niña, Temperatura da Superfície do Mar (TSM) do Oceano Pacífico. Esta metodologia é utilizada a fim de se obter um modelo que permita aprimorar o poder de explicação da velocidade do vento e mensurar o possível efeito dos fenômenos El Niño e La Niña nos mesmos, para que futuramente possa-se fornecer previsões mais acuradas condicionadas a estas ocorrências.

1.2

Estrutura da dissertação

Esta dissertação está estruturada em seis capítulos além desta introdução. O segundo capítulo é destinado ao detalhamento da circulação atmosférica e caracterização dos ventos na região de estudo. O terceiro capítulo especifica os fenômenos climáticos de interesse, El Niño e La Niña e suas conhecidas influências. O quarto capítulo faz uma revisão bibliográfica de trabalhos que abordam o tema proposto. O quinto capítulo se refere a metodologia dos modelos GAS. O sexto capítulo descreve os dados utilizados nessa dissertação, bem como os procedimentos realizados e os resultados encontrados. Por fim, o sétimo capítulo apresenta as conclusões e propostas de trabalhos futuros.

2

Circulação Atmosférica e Regime dos Ventos no Nordeste brasileiro

2.1

Caracterização da área de estudo

Segundo Rao *et al.* (1993), a região Nordeste do Brasil (NEB) está localizada aproximadamente entre as latitudes 1° S - 18° S e longitudes 35° O - 47° O, conforme Figura 2.1 e abrange nove estados brasileiros, a saber: Alagoas, Bahia, Ceará, Maranhão, Paraíba, Piauí, Pernambuco, Rio Grande do Norte e Sergipe. Numa área de $1.558.196 \text{ km}^2$ e que em 2010 possuía 53.081.950 habitantes (IBGE, 2011), a região é banhada pelo Oceano Atlântico ao norte e a leste. A temperatura média anual no Nordeste varia de 20° a 28°C e os três tipos de clima presentes nele são: equatorial úmido, clima litorâneo úmido; clima tropical e clima semi-árido (Cavalcanti *et al.*, 2009).

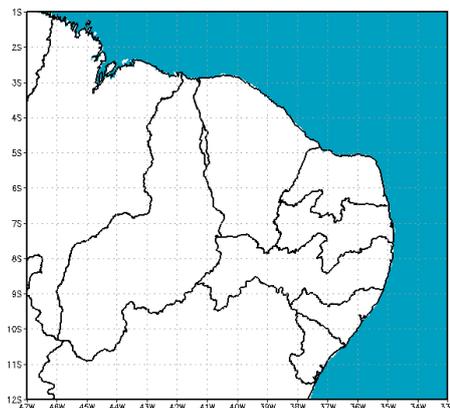


Figura 2.1: Área de estudo.

Segundo, Cavalcanti *et al.* (2009) o Nordeste apresenta uma variabilidade interanual pronunciada, em especial na precipitação, e é uma das regiões da América do Sul cuja variação intrassazonal é mais evidente. O clima da região é determinada pela influência da variação e posicionamento dos Anticiclones Subtropicais do Atlântico Sul (ASAS), Atlântico Norte (ASAN) e do cavado equatorial. E alguns fatores climáticos como: a vegetação, o relevo,

a posição geográfica e os sistemas de pressão atuantes na região determinam a distribuição dos elementos climáticos do Nordeste.

Na literatura, muito se encontra sobre a característica complementar entre o vento e a precipitação na região nordeste do Brasil. Diversos estudos sobre o potencial de energia do Nordeste brasileiro, como Amarante *et al.* (2001b), Marinho & Aquino (2009) e Pinto *et al.* (2013), exploram a complementaridade sazonal dos regimes hidrológico e eólico, o que hoje motiva um sistema de junção entre a energia por fonte de usinas eólica e hidrelétricas para a geração de energia.

Na Figura 2.2 são apresentadas as curvas sazonais normalizadas das séries históricas de vento de uma usina eólica no Ceará, disponibilizada pela empresa 3TIER, e de vazão, retirada do site do Operador Nacional do Sistema Elétrico - ONS [<http://www.ons.org.br>] para um posto natural pertencente ao Rio Parnaíba, o qual banha regiões próximas a esta usina.

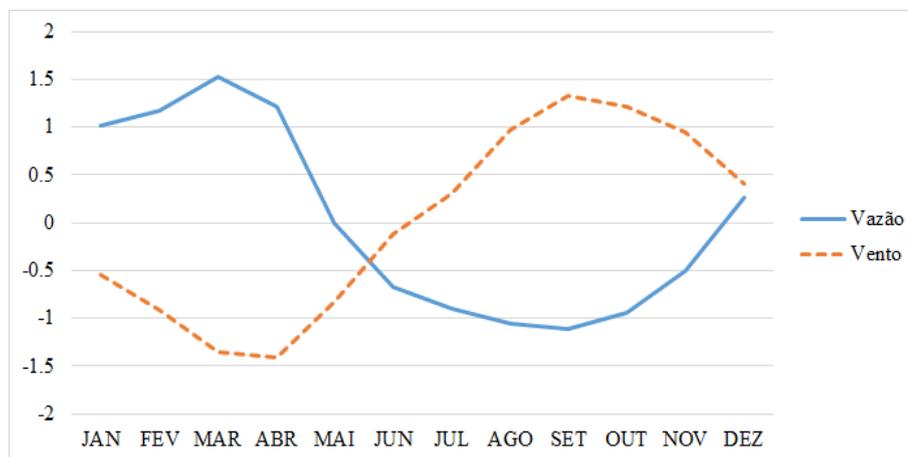


Figura 2.2: Curvas sazonais normalizadas de ventos de uma usina eólica e vazão de um posto natural.

Nota-se pela Figura 2.2, que o período de maior intensidade dos ventos no Nordeste coincide com os períodos de seca — ou baixa precipitação; e vice-versa. Motivada por tal característica, atualmente a energia eólica é utilizada como uma forma alternativa complementar para a geração de eletricidade.

2.2 Circulação Atmosférica no Nordeste

O globo terrestre recebe energia provinda do sol durante todo o ano. No entanto, em cada região do planeta esta energia é recebida com diferentes intensidades em cada estação do ano. Devido a essa variação da radiação solar, o ar se move em resposta as diferenças horizontais de pressão oriundas desta

variação. Assim, o vento pode ser definido como sendo o deslocamento das massas de ar, na tentativa de igualar estas diferenças de pressão.

O vento atua nas condições do tempo, sendo ele responsável pelo transporte de umidade e energia na atmosfera terrestre. O mesmo pode ser influenciado por alguns fatores como: relevo, altura e rugosidade do solo. Assim, o vento perde intensidade a medida que se afasta do litoral, uma vez que em direção ao continente o mesmo encontra obstáculos como os citados anteriormente. Com isso, o litoral do Nordeste é a região que possui o maior aproveitamento de energia pelo vento e onde estão situadas grande parte dos parques eólicos.

A região Nordeste é caracterizada por produzir ventos constantes e intensos, onde predominam os ventos alísios e as brisas. Os alísios são ventos constantes, predominantes em baixos níveis de altitude em regiões tropicais (Amarante *et al.*, 2001a). Estes se deslocam da parte alta pressão para a de baixa pressão, formando uma componente de leste para oeste no hemisfério sul. As brisas são ventos menos intensos que os alísios e podem ser divididas entre brisas terrestres e brisas marítimas. As brisas marítimas acontecem durante o dia, quando o vento se desloca do mar para o continente. Já a brisa terrestre é o inverso da brisa marítima, ocorrendo durante a noite ao se deslocar do continente para o oceano.

3

O fenômeno El Niño

A grande variabilidade da circulação atmosférica que ocorre em diversas escalas de tempo e espaço fundamenta o termo “teleconexão”, utilizado em meteorologia, que segundo Cavalcanti *et al.* (2009) justifica o fato de anomalias ocorridas em determinada região poderem estar relacionadas a anomalias ocorridas em outro tempo ou espaço. Preferencialmente, estas teleconexões são também conhecidas como eventos de baixa frequência, ainda que o padrão das mesmas possa perdurar-se em pequenas escalas de tempo.

Uma das primeiras teleconexões globais descobertas foi a Oscilação Sul, que será descrita no decorrer desse capítulo.

3.1

Oscilação Sul

A Oscilação Sul foi descoberta por Walker e Bliss e documentada em estudos realizados em 1932 e 1937 como Bliss & Walker (1932). Esta se refere à existência da diferença de pressão no Oceano Pacífico e no Oceano Índico. Foi observado que enquanto no Pacífico a pressão sofria um aumento, no Índico ocorria o inverso. Esse efeito “gangorra” das pressões atmosféricas entre os dois oceanos foi então definido por Walker como “Oscilação Sul” devido a essa flutuação anual.

Desde a década de 60 a Oscilação Sul e o El Niño vêm sendo tratados conjuntamente, sendo atribuído ao termo “El Niño Oscilação Sul” (ENOS). Assim, sempre que for citado nessa dissertação ENOS, estará se fazendo uma referência ao fenômeno El Niño/La Niña.

Segundo Oliveira (1999), o fenômeno El Niño é definido em grande parte da literatura como sendo um aquecimento anormal da temperatura das águas do Oceano Pacífico Equatorial. A origem deste nome deu-se em função do período de sua primeira observação, na época do Natal, próximo à costa norte do Peru sendo, portanto, associado ao menino Jesus, que em espanhol é referido como “El Niño”. Este também pode ser compreendido como uma combinação entre o aquecimento irregular do Oceano Pacífico e do enfraquecimento dos

ventos alísios na região equatorial do mesmo. Além disso, acredita-se que a razão do fenômeno ocorrer somente no Oceano Pacífico se deve a sua grande extensão embora exista influência do mesmo em muitos lugares do planeta.

Logo, segundo Cavalcanti *et al.* (2009), o El Niño é um fenômeno de interação entre o oceano e a atmosfera e principal fonte de variabilidade climática interanual global, sendo ele capaz de produzir alterações no clima em diversas regiões do mundo. Este por sua vez, pode alterar a pressão, a TSM, o padrão dos ventos, assim como acarreta mudança no regime de chuvas em regiões tropicais. Vale ressaltar que os efeitos do El Niño ocorrem de forma diferente em cada região, sendo este, portanto, responsável tanto por secas severas quanto por chuvas em abundância. O La Niña pode ser compreendido como a fase fria do fenômeno El Niño.

3.1.1

Padrão de circulação atmosférica em anos neutros

Segundo Ahrens (2008), em anos normais, isto é, sem a presença de El Niño ou La Niña, a maior pressão na região sudeste do Oceano Pacífico e a baixa pressão no lado oposto a este, próximo à Indonésia, induzem a formação de ventos alísios de leste ao longo do Equador.

As águas do Oceano Pacífico Equatorial sob a influência da radiação solar sofrem alterações através dos ventos alísios que sopram de leste para oeste entre as latitudes 0° e 30° . Estes transportam as águas quentes superficiais próximas à costa Pacífico Equatorial fazendo com que haja um desnível, que pode ser observado na Figura 3.1 na região da Termoclina, a qual por definição é a região de transição entre as águas mais quentes das mais frias. Com isso, tem-se um nível menor do mar próximo à superfície terrestre e um nível maior no Oceano Pacífico Equatorial Oeste. Assim, as águas superficiais próximas à costa se apresentam de forma mais fria devido a sua substituição pelas águas frias das partes profundas do oceano, e as águas do Pacífico Oeste mais quentes devido ao transporte das águas quentes pelos alísios (Oliveira, 1999).

Devido ao maior aquecimento das águas do Pacífico Oeste proveniente das trocas de águas citadas anteriormente, ocorre o processo de evaporação que dá origem ao surgimento de nuvens em regiões de baixos níveis de altitude, próximas a superfície. As nuvens se formam através dessa evaporação do ar pelo movimento ascendente do mesmo que normalmente encontra ar ascendente e chuvas intensas sobre o Pacífico Oeste e ar descendente e tempo seco em geral sobre o Pacífico Oriental.

O mecanismo da circulação atmosférica detalhado anteriormente forma a

chamada Circulação de Walker também conhecida como Célula de Walker, que atua no sentido zonal, sendo esta relacionada à variação da pressão atmosférica nas regiões leste e oeste do Oceano Pacífico. O desenho esquemático da mesma atuante sobre o Oceano Pacífico Equatorial em anos neutros é apresentado na Figura 3.1.

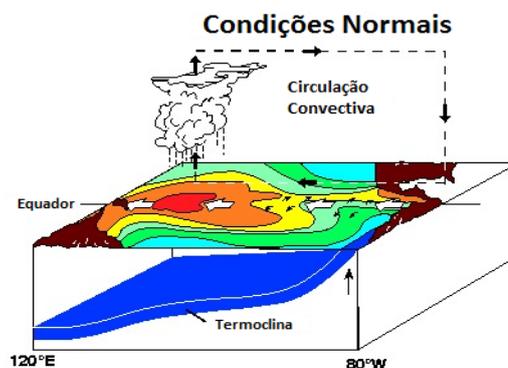


Figura 3.1: Padrão de circulação atmosférica em anos neutros. Fonte: Adaptado de <http://www.pmel.noaa.gov/tao/elniño>.

3.1.2

Padrão de circulação atmosférica em anos de El Niño

Em anos com a presença do fenômeno El Niño, os ventos alísios perdem sua força na região equatorial do Pacífico, de forma até a mudarem sua direção, indo de oeste para leste. Isto se deve ao fato de a pressão atmosférica sofrer uma redução no Pacífico Leste e um aumento no Pacífico Oeste.

Com o enfraquecimento dos alísios, as águas se apresentam de forma quente em todo o Pacífico Equatorial e, devido a esse aquecimento das águas em toda a extensão do Pacífico, tem-se a evaporação que por sua vez acarreta a formação de nuvens na região. Entretanto, neste caso a célula de Walker sofre um processo de bipartição oriundo do aumento da contracorrente que transporta água quente de oeste para leste, de forma que se possa observar a formação de nuvens no Pacífico Central e no Leste (Figura 3.2). A partir da bipartição da célula de Walker ocorre uma mudança da posição do ramo ascendente fazendo com que as águas do Pacífico Equatorial Leste se apresentem com temperaturas mais altas que o normal. Neste caso, têm-se dois ramos descendentes: um na região da Indonésia e outro no norte e leste da Amazônia e norte do Nordeste brasileiro. Entretanto, o encontro do ar descendente de altos níveis da troposfera com o ar ascendente é capaz de inibir a formação de nuvens, o que favorece, porém não determina a seca em determinadas regiões, como o Nordeste brasileiro (Oliveira, 1999).

A região da Termoclina se apresenta de forma mais profunda próxima à costa Oeste do Pacífico em condições de El Niño como pode ser observado na Figura 3.2.

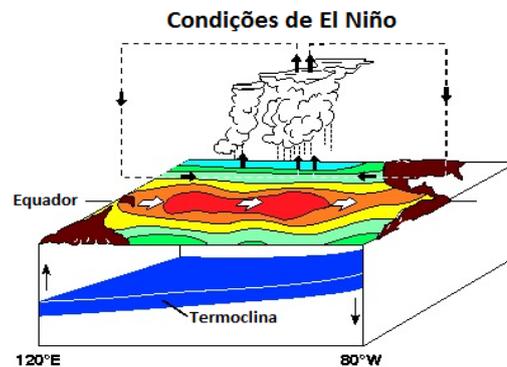


Figura 3.2: Padrão de circulação atmosférica em anos de El Niño. Fonte: Adaptado de <http://www.pmel.noaa.gov/tao/elnino>.

3.1.3

Padrão de circulação atmosférica em anos de La Niña

As águas do Pacífico Equatorial na presença do fenômeno La Niña, sofrem um resfriamento anormal. Nesta condição os ventos alísios se tornam mais intensos de forma que o desnível, citado anteriormente em anos normais, aumente (Figura 3.3). Com isso, ocorrerá um maior represamento das águas quentes no Pacífico Oeste que por sua vez irá gerar evaporação do ar, induzindo o mesmo a movimentos ascendentes, formando as nuvens de chuva. Todavia, na presença de La Niña, a célula de Walker se torna mais longa fazendo com que os movimentos descendentes do Pacífico Equatorial Oeste fiquem mais intensos, inibindo a formação de nuvens e consequentemente a precipitação (Oliveira, 1999).

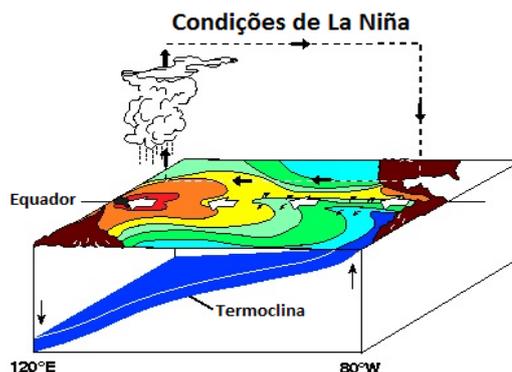


Figura 3.3: Padrão de circulação atmosférica em anos de La Niña. Fonte: Adaptado de <http://www.pmel.noaa.gov/tao/elnino>.

3.2

Métodos para a identificação dos períodos e quantificação de El Niño e La Niña

Para a identificação e medição da magnitude de El Niño e La Niña, pode-se utilizar os seguintes índices: *Southern Oscillation Index* (SOI), a anomalia da *sea surface temperature* (SST), o *Multivariate ENSO Index* (MEI) ou ainda o *Oceanic Niño Index* (ONI).

A seguir é apresentada a descrição de cada um dos índices e o cálculo dos mesmos.

Vale ressaltar que para a classificação anual dos fenômenos El Niño e La Niña, toma-se por base a ocorrência de no mínimo cinco meses consecutivos de anomalias extremas de TSM, no caso de índices mensais, de indicação da ocorrência do fenômeno em questão e para os demais índices, obtidos por média entre dois e três meses, toma-se por base três meses consecutivos.

3.2.1

Southern Oscillation Index (SOI)

O cálculo deste índice é baseado nas diferenças de anomalia de pressão ao nível do mar entre Taiti (Polinésia Francesa) e Darwin (Austrália), por estas regiões apresentarem um comportamento correspondente ao das mudanças de temperatura dos oceanos Pacífico Tropical e Oriental segundo CPC (2013). Esta descoberta foi desenvolvida por Gilbert Walker e publicada em artigos entre os anos de 1923 e 1937, em Bliss & Walker (1932).

Existe ainda, o Equatorial SOI, calculado a partir das anomalias padronizadas de pressão ao nível do mar entre a Indonésia e o Pacífico Equatorial Oriental.

A série mensal do SOI e do Equatorial SOI é disponibilizada para o período de 1951 a 2013 no site do Climate Prediction Center [<http://www.cpc.ncep.noaa.gov/data>]. A interpretação dos mesmos é realizada de forma que para valores negativos do índice tem-se caracterização dos fenômenos El Niño, enquanto que para valores positivos, La Niña para valores próximos de zero, não há indícios de ocorrência de nenhum dos fenômenos.

3.2.2 Sea Surface Temperature (SST)

A quantificação da intensidade ou classificação da ocorrência do El Niño/La Niña pode ser realizado através da anomalia da SST do Pacífico Equatorial, uma vez que estes fenômenos estão diretamente relacionados a alterações da temperatura da superfície do mar no Oceano Pacífico Equatorial. O cálculo para as anomalias de SST do Pacífico são realizadas em regiões Niño referenciadas pelos cientistas como: Niño 1+2 (0-10°S)(90°O-80°O), Niño 3 (5°N-5°S)(150°O-90°O), Niño 4 (5°N-5°S) (160°L-150°O) e Niño 3.4 (5°N-5°S)(170-120°O) (NOAA, 2014). A representação destas regiões pode ser observada na Figura 3.4. Para identificação da ocorrência do El Niño pelas anomalias de SST tem-se que para valores positivos de anomalia da SST no Pacífico têm-se indicações de que a temperatura do mar está acima da média climatológica, caracterizando assim a presença de El Niño. O inverso ocorre em episódios de La Niña. O valor base para tais conclusões é de +/- 0.5°C.

Tanto as séries mensais de temperaturas da superfície do mar no Pacífico quanto as anomalias das mesmas são disponibilizadas para as quatro regiões Niño durante o período de 1950 a 2014 no site do NOAA [<http://www.noaa.gov>].

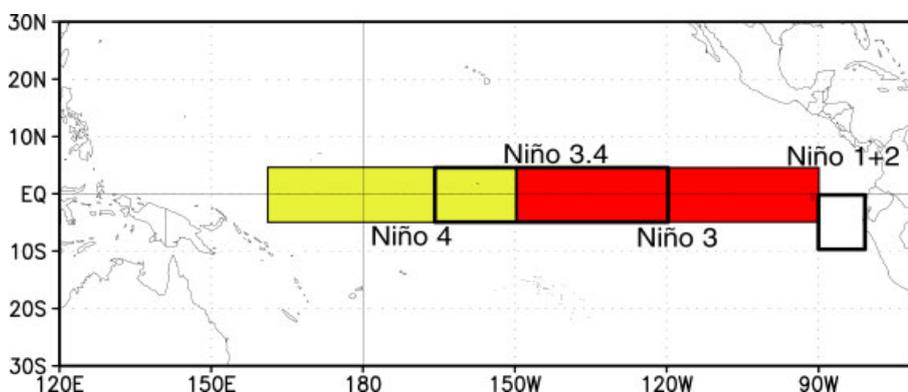


Figura 3.4: Região Niño no Oceano Pacífico Equatorial Fonte: www.cpc.ncep.noaa.gov

3.2.3

Multivariate ENOS Index (MEI)

Segundo Esrl (2013), o índice multivariado de ENOS é calculado pelo método estatístico de componentes principais (*Principal Components Analysis* (PCA)) utilizando seis variáveis observadas no Oceano Pacífico Equatorial: pressão do nível do mar, componentes zonal e meridional do vento à superfície, temperatura da superfície do mar, temperatura do ar na superfície e fração de nebulosidade total do céu. A utilização de um número maior de variáveis para o cálculo deste índice o torna mais complexo.

A série bimestral do MEI é disponibilizada no site do NOAA [<http://www.esrl.noaa.gov/psd/enso/mei/table.html>] para os períodos entre 1950 a 2014.

A interpretação do índice é dada de forma que os valores negativos representam a ocorrência do fenômeno La Niña, enquanto que para valores positivos do índice tem-se a representação da ocorrência do fenômeno El Niño.

3.2.4

Oceanic Niño Index (ONI)

Segundo Dahlman (2009), este índice é calculado através da temperatura média da superfície do mar para uma faixa definida no Oceano Pacífico Tropical. Esta é chamada de região Niño 3.4 e sua localização é compreendida entre as latitudes 5° N e 5° S e longitude 120° O e 170° O, como mostra a Figura 3.4.

O cálculo do índice se baseia inicialmente na SST mensal da região Niño 3.4. Posteriormente é realizada a média entre as médias do mês em questão e dos meses anteriores e sucessores a este, que posteriormente será comparada com as médias dos mesmos três meses de períodos anteriores. Mais detalhes sobre o cálculo deste índice pode ser encontrado no site do CPC (2013).

Para a classificação do El Niño através desse índice, deve-se ter valores de ONI maiores ou iguais a 0,5 (°C) para a caracterização da fase quente do fenômeno. Em relação a fase fria, La Niña, a condição seria valores de ONI menores ou iguais a -0,5. Quanto a intensidade de cada fenômeno, os mesmos podem ser classificados como: fraco (0,5 a 0,9), moderado (1,0 a 1,4) e forte ($\geq 1,5$). Para períodos com valores no intervalo entre -0,5 e 0,5 é caracterizado como ano neutro, isto é, sem a ocorrência dos fenômenos.

A série trimestral com defasagem de um mês do índice também são disponibilizadas no site do Centro de Previsão Climática do NOAA para o período de 1950 a 2014.

3.3

Episódios de El Niño e La Niña (1950 – 2012)

Na Tabela 3.1 estão listados os períodos na presença dos fenômenos tais como suas classificações quanto à intensidade dos mesmos. Estas realizadas através do limiar de +/-0.5 °C das anomalias de TSM da região Niño 3.4. Os episódios de El Niño e La Niña são observados com frequência de 2 a 7 anos, apesar de nos últimos anos o fenômeno La Niña ter ocorrido com menor frequência.

Tabela 3.1: Episódios El Niño e La Niña. Fonte: NOAA, 2014.

Fenômeno	Início	Fim	Duração
El Niño	1982	1983	Abr/Jul
La Niña	1983	1984	Ago/Dez
La Niña	1984	1985	Set/Ago
El Niño	1986	1988	Jul/Jan
La Niña	1988	1989	Abr/Abr
El Niño	1991	1992	Abr/Mai
El Niño	1994	1995	Ago/Fev
La Niña	1995	1996	Ago/Fev
El Niño	1997	1998	Abr/Mar
La Niña	1998	2001	Jun/fev
El Niño	2002	2003	Abr/Dez
El Niño	2004	2004	Jun/Dez
La Niña	2005	2006	Out/Fev
El Niño	2006	2006	Ago/Dez
La Niña	2007	2008	Jul/Mai
La Niña	2008	2009	Out/Fev
El Niño	2009	2010	Jun/Mar
La Niña	2010	2011	Jun/Mar
La Niña	2011	2012	Ago/Fev

3.4

Relação entre os fenômenos El Niño e Dipolo do Atlântico

Alguns estudos como os de Hastenrath *et al.* (1977) e Covey *et al.* (1978) acreditam que a relação entre o fenômeno ENOS e o clima na região nordeste do Brasil não é direta e sim por intermédio do Oceano Atlântico Tropical.

Markham *et al.* (1977), encontraram evidências de que a ocorrência do fenômeno El Niño pode estar associada a anomalias positivas de TSM do Atlântico Tropical Norte e anomalias negativas de TSM do Atlântico Sul.

Segundo Aragão *et al.* (1996), o fenômeno Dipolo do Atlântico pode ser interpretado como uma mudança anormal na temperatura da superfície do mar no Oceano Atlântico Tropical. Tal fenômeno é capaz de inibir ou não a formação de nuvens em alguns áreas, reduzindo ou aumentando a precipitação nestas regiões. O Dipolo Negativo é caracterizado pelo aquecimento anormal das águas do Atlântico Norte e resfriamento anormal das águas do Atlântico Tropical e Sul, inibindo a formação de nuvens numa região, desfavorecendo a precipitação, enquanto que para o Dipolo Positivo ocorre o processo inverso.

Assim, o Dipolo do Atlântico pode ser calculado através da diferença entre as anomalias normalizadas de TSM do Atlântico Norte e anomalias normalizadas de TSM do Atlântico Sul.

Aragão *et al.* (1996) ainda relaciona ambos os fenômenos, ENOS e Dipolo do Atlântico, de forma que para valores de anomalias de TSM Pacífico acima do normal (positivos), isto é, ocorrência do ENSO, tem-se Dipolo negativo e vice-versa. Logo, o autor associa o Dipolo negativo no Atlântico com anomalias positivas no Pacífico Equatorial, coincidindo com o período de secas no NEB e vice-versa.

4

Revisão Bibliográfica

Muito se sabe da influência do fenômeno El Niño na precipitação no mundo, como secas severas no Brasil e na Indonésia, no entanto, esta quando comparada à influência do mesmo no regime dos ventos os estudos são raros.

Nesse capítulo, é realizada uma revisão bibliográfica de trabalhos que abordam a relação entre o El Niño e o regime de ventos tanto em âmbito nacional quanto mundial, com foco em produção eólica. Os trabalhos utilizam diferentes técnicas estatísticas para estudar tal relação.

Neto *et al.* (2006) analisa as condições de vento sobre a região nordeste do Brasil com foco para o estado do Ceará entre os meses de setembro e novembro através de uma comparação dos anos de La Niña e El Niño forte, 1973 e 1982, respectivamente. Foi realizada uma comparação do efeito dos fenômenos El Niño e La Niña na velocidade e direção do vento na região a partir de duas fontes de dados. Foi utilizado o modelo numérico de previsão CSU-RAMS para a simulação do vento na região para o período de interesse e dados de reanálise *National Centers for Environmental Prediction* (NCEP). Os resultados mostraram que através dos dados de reanálise em grande escala, os ventos se pronunciavam de forma mais intensa quando a região estava sob ocorrência do El Niño e que os resultados oriundos da simulação do vento na região se mostraram de forma bem similar a estes. Além disso, notou-se que ventos mais intensos foram observados na parte norte do Nordeste brasileiro. Em relação a direção do vento na região, esta apresentou uma maior diferença entre o período de La Niña (1973) e El Niño (1982) sobre o oceano Atlântico, no litoral do Ceará do que o inverso, isto é, diferença entre os períodos de El Niño e La Niña, Para o litoral do estado, os ventos se pronunciaram de forma aproximadamente constante. Conclui-se que em anos de ocorrência do fenômeno El Niño, a região mostra melhores condições para a geração eólica.

Em Harper *et al.* (2007), é proposta a utilização de dois métodos estatísticos a fim de se identificar e quantificar os efeitos do ENSO na Energia Eólica. Isto é realizado a partir de dados observacionais de quatro aeroportos situados em diferentes localidades nos Estados Unidos. No estudo,

são utilizados dados horários de vento e nestes realizada uma partição da série em grupos de acordo com a classificação do fenômeno no ano correspondente. Para testar a significância local e global de algum possível efeito do El Niño na velocidade do vento, realiza-se o teste do sinal, um método estatístico não paramétrico. E por fim, para a construção de um intervalo de confiança é feita uma reamostragem por *bootstrap*. Com a utilização dos métodos citados acima, concluiu-se que a magnitude dos sinais dos efeitos do El Niño não é significativa ainda que os sinais deste efeito existam. Foi encontrada evidência de que, em geral, existe uma tendência de redução da velocidade média do vento em meses de ocorrência o fenômeno El Niño, sendo esta mais pronunciada entre os meses de janeiro a abril e setembro a outubro. Devido a estes indícios, os efeitos ainda que não significantes em magnitude, são responsáveis pela diminuição da energia média produzida, uma vez que esta é sensível à distribuição de probabilidade da velocidade do vento. Os períodos de ocorrência do fenômeno La Niña se apresentaram de forma similar ao da fase quente, El Niño.

Oliveira & Costa (2011), realizam uma análise do comportamento dos ventos do Nordeste nos períodos de El Niño forte (1982-1983) e La Niña forte (1973-1974) através de dados de reanálise e dados simulados para a velocidade e direção do vento a partir de um modelo atmosférico. A análise do efeito dos fenômenos climáticos foi realizada através da investigação do comportamento dos oceanos Atlântico Tropical e Pacífico Equatorial. A metodologia proposta baseia-se na correlação mensal entre as anomalias de vento e índices de Niño do Pacífico e Dipolo do Atlântico. Pela análise da correlação, notou-se que as mesmas apresentaram maior significância entre os meses de fevereiro e maio. Além disso, a partir dos dados simulados pelo modelo proposto, estimou-se a geração eólica para os períodos de interesse juntamente com os parâmetros da distribuição de Weibull para a velocidade do vento para cada mês. Os resultados encontrados mostraram que para o período de El Niño forte (1982-1983) o vento se pronuncia de forma mais intensa. Já no período que corresponde à fase fria do evento climático, o vento atua de forma menos intensa.

Pinto *et al.* (2011), faz um estudo das dinâmicas hidrológicas e eólicas do Brasil, mais especificamente em duas regiões, Nordeste e Sul do país. Os dados utilizados são oriundos da reanálise NCEP. O objetivo principal deste é a avaliação do impacto da inclusão das usinas de energia eólica na geração de energia no cenário brasileiro. O estudo leva em consideração a complementariedade do vento e da vazão, a qual é bem definida para a região Nordeste, bem como as mudanças climatológicas que afetam a

região. A análise da complementariedade é realizada através de médias móveis das correlações entre vento e vazão. A modelagem proposta se baseia na decomposição das séries de vento em componentes como: El Niño, La Niña, ciclo solar, sazonalidade. Isto foi realizado através da modelagem do espectro de cada uma das séries por um filtro digital. Posteriormente são gerados novos processos estocásticos das séries originais, onde estas são alimentadas pelos filtros digitais, sem a perda das características originais da série. Os autores ainda ressaltam a importância de se propor modelos capazes de captar a estrutura dinâmica das séries, bem como suas mudanças estruturais, para que não se perca o comportamento original da série e tornando ainda mais difícil análises mais específicas, principalmente no que diz respeito a mudanças climáticas.

Em Matos (2013), é apresentada uma abordagem estatística mais complexa através dos modelos de séries temporais GAS a fim de analisar cenários de vento, utilizando o fator de capacidade de uma usina de energia eólica localizada no Ceará, e de vazão do Rio Paraíba localizada no estado de Minas Gerais. Na modelagem do vento foi utilizado o fator de capacidade da usina e para a análise geral dos modelos todo arcabouço estatístico de análise de séries temporais. A motivação da utilização de modelos GAS se dá em função da busca por uma melhoria da captura da dinâmica da série temporal pelo mesmo. Para a análise das variáveis de interesse, foram desenvolvidos modelos com densidade preditiva beta e gama, respectivamente, inicialmente de forma univariada. Os modelos univariados propostos para ambas as análises foram desenvolvidos em dois tipos de estruturas de evolução: parâmetros variantes no tempo e tipo *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Por fim, o mesmo desenvolveu um modelo bivariado do fator de capacidade do vento com a vazão. Este, apresentou resultados semelhantes na previsão *in-sample* e *out-of-sample* dos modelos univariados. O autor resalta ainda a importância da complementariedade sazonal entre o vento e a vazão no processo de modelagem.

5 Metodologia

Nesse capítulo aborda-se a metodologia proposta para o desenvolvimento dessa dissertação, os *Generalized Autoregressive Score Models*, ou modelos GAS, também conhecidos como *Dynamic Score Models*.

5.1 Modelo GAS

Segundo Cox (1981), os modelos para séries temporais não-Gaussianas onde os parâmetros são variantes no tempo podem ser de dois tipos: modelos guiados por parâmetros (*parameter driven models*) e modelos guiados por observações (*observation driven models*). A principal diferença entre os modelos guiados por observações e os guiados por parâmetros é o fato de que nos primeiros, apesar de os parâmetros serem processos estocásticos, estes podem ser perfeitamente previstos usando-se a esperança condicional como função de previsão.

Os modelos GAS, propostos por Creal *et al.* (2008), fazem parte de uma nova classe de modelos guiados por observações cujos parâmetros variam ao longo do tempo. Esta é uma abordagem que permite a captura do comportamento dinâmico de séries não-Gaussianas, o que a torna mais flexível.

Segundo Creal *et al.* (2008), a formulação GAS para séries temporais, diferentemente dos demais modelos frequentemente propostos na literatura, introduz uma função que relaciona os parâmetros variantes no tempo da distribuição condicional assumida para a série, com valores defasados dos mesmos e valores defasados do *score* condicional, além de valores correntes e/ou defasados de variáveis exógenas. Esta função pode definir o mecanismo de atualização dos parâmetros através de duas abordagens: evolução do tipo SARIMA, a qual é construída com base na formulação dos *Autoregressive Moving Average Models* (ARMA), e a evolução por componentes não observáveis, fundamentada pelo arcabouço dos modelos estruturais de Harvey.

A utilização dos modelos GAS para a modelagem de séries temporais apresenta características favoráveis, como o fato de a avaliação da função de verossimilhança possuir forma fechada. Além disso, estes modelos são capazes de explorar a estrutura completa da função de densidade preditiva para a atualização dos parâmetros variantes no tempo. Estes são estimados por máxima verossimilhança, o que, por ser uma otimização não-linear, exige o uso de métodos numéricos.

Grande parte dos modelos não-Gaussianos existentes na literatura podem ser especificados na forma dos modelos GAS a partir da escolha de parâmetros variantes no tempo e parametrização adequada como: *Generalized Autoregressive Conditional Heterocedasticity Model* (GARCH), *Multiplicative Error Model* (MEM), *Dynamic Exponential Family Models*, entre outros, como exemplificado em Creal *et al.* (2008).

Vale ressaltar que sempre que se fizer referência à função de verossimilhança nessa dissertação, o interesse será na função de log-verossimilhança, uma vez a maximização de ambas as funções é equivalente e a utilização da segunda é mais adequada para cálculos computacionais devido a maior facilidade no cômputo de derivadas da soma do que as derivadas do produto.

A seguir, é apresentada a definição da estrutura do modelo GAS em sua forma geral, com evolução do tipo SARIMA e também com evolução do tipo componentes não-observáveis. Posteriormente, será apresentada a formulação GAS para as distribuições utilizadas nessa dissertação e suas variações.

5.1.1

Especificação do Modelo GAS com estrutura de evolução SARIMA

Seja y_t a variável dependente de interesse no estudo. Define-se f_t como um vetor de parâmetros variantes no tempo, x_t vetor de variáveis exógenas, no tempo t e θ o vetor de parâmetros fixos.

Além disso, pode-se definir os seguintes conjuntos:

- $Y_1^t = \{y_1, \dots, y_t\}$;
- $F_1^t = \{f_1, \dots, f_t\}$;
- $X_1^t = \{x_1, \dots, x_t\}$;

Toda a informação disponível em t pode ser representada por:

$$F_t = \{Y_1^{t-1}, F_1^{t-1}, X_1^t\}, \quad t = 1, \dots, n.$$

Supondo uma densidade preditiva geral para y_t dada por

$$p(y_t|f_t, F_t; \theta) \quad (5-1)$$

para todo $t = 1, \dots, n$.

A dinâmica dos parâmetros variantes no tempo é introduzida através do seguinte mecanismo de atualização:

$$f_t = \omega + \sum_{i=1}^p A_i s_{t-i} + \sum_{j=1}^q B_j f_{t-j} \quad (5-2)$$

onde ω é um vetor de constantes, A_i e B_j são matrizes com dimensões apropriadas para $i = 1, \dots, p$ e $j = 1, \dots, q$, s_t é uma função escalar, que pondera o *score* da log-verossimilhança que por sua vez depende das observações passadas, dos parâmetros variantes no tempo f_t e do vetor de parâmetros fixos θ .

Assim, o mecanismo de atualização dos parâmetros que variam no tempo é dado por uma equação com forma similar a estrutura autoregressiva-médias móveis dos modelos SARIMA de Box e Jenkins.

A função que pondera o vetor *score*, s_t , é dada por:

$$s_t = S_t \cdot \nabla_t \text{ e } \nabla_t = \frac{\partial \ln p(y_t|f_t, F_t; \theta)}{\partial f_t} \quad (5-3)$$

onde S_t é a matriz de ponderação do vetor *score*, e ∇_t o vetor *score*.

A matriz S_t pode ser definida de várias maneiras, ou seja, existem muitas formas de ponderação do vetor *score*. No entanto, segundo Creal *et al.* (2013), é natural que a forma da ponderação dependa da variância do vetor *score*. Assim, uma das escolhas mais realizadas é a matriz de ponderação S_t como sendo a pseudo-inversa da matriz de informação de Fisher $I_{t|t-1}^{-1}$. Vale ressaltar que caso não exista uma forma fechada para esta matriz, é possível utilizar aproximações numéricas. Esta e outras formas de ponderação do vetor *score* como a raiz quadrada da pseudo-inversa da matriz de informação de Fisher e a matriz identidade, são apresentadas abaixo:

$$S_t = I_{t|t-1}^{-1}, \quad S_t = I_{t|t-1}^{-1/2}, \quad S_t = I \quad (5-4)$$

onde

$$\begin{aligned} I_{t|t-1} &= E_{t-1}[\nabla_t \nabla_t'] \\ &= -E_{t-1} \left[\frac{\partial^2 \ln p(y_t|f_t, F_t; \theta)}{\partial f_t^2} \right] \end{aligned} \quad (5-5)$$

5.1.2

Especificação do Modelo GAS com estrutura de evolução CNO

Segundo Creal *et al.* (2008), modelos de componentes não-observáveis pertencem à classe de modelos guiados por parâmetros com uma única fonte de erro. Com isso, é proposto um outro mecanismo de evolução para os modelos GAS, onde este é baseado na estrutura dos modelos de componentes não observáveis. Nesta nova formulação, um sinal univariado ψ_t pode ser extraído da série temporal $\{y_1, y_2, \dots, y_n\}$ e, assim, o vetor $f_t = (f_{1,t}, f_{2,t}, \dots, f_{r,t})'$ é decomposto em r fatores, onde estes representam cada uma das componente não observáveis da série temporal como por exemplo: tendência, sazonalidade e ciclo. Um exemplo é apresentado a seguir.

Seja uma série temporal $\{y_1, y_2, \dots, y_n\}$, a qual pode ser decomposta em componentes de tendência com *drift*($f_{1,t}$) e ciclo ($f_{2,t}$) e, assim, $r = 2$. O fator de tendência é representado por um passeio aleatório, enquanto que o fator de ciclo por um processo autoregressivo de segunda ordem. Assim, tem-se que:

$$\begin{aligned}\psi_t &= f_{1,t-1} + f_{2,t-1} \\ f_{1,t} &= \omega + a_1 s_t + f_{1,t-1} \\ f_{2,t} &= a_2 s_t + \phi_1 f_{2,t-1} + \phi_2 f_{2,t-2}\end{aligned}$$

A junção dos termos anteriores fornece a forma matricial de f_t dada por:

$$f_{t+1} = \begin{pmatrix} w \\ 0 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} s_t + \begin{pmatrix} 1 & 0 \\ 0 & \phi_1 \end{pmatrix} f_t + \begin{pmatrix} 0 & 0 \\ 0 & \phi_2 \end{pmatrix} f_{t-1} \quad (5-6)$$

Portanto, através dessa decomposição obtêm-se o modelo GAS (1,2) representado pela equação de observações $y_t = \psi + \epsilon_t = f_{1,t} + f_{2,t} + \epsilon_t$ com equação de evolução dada por f_t (Eq. (5-6)), onde $p(y_t|\psi_t; \theta)$ é uma distribuição $N(f_{1,t-1} + f_{2,t-1}, \sigma^2)$. Os parâmetros ω , a_1 , a_2 , ϕ_1 e ϕ_2 pertencem ao vetor de parâmetros fixos do modelo e são estimados por máxima verossimilhança, assim como no modelo GAS com evolução SARIMA.

Note que a estrutura de atualização f_t é a mesma utilizada na evolução por tipo SARIMA, porém nessa abordagem a dinâmica da série temporal é dividida em componentes.

Logo, a especificação de um modelo GAS baseia-se principalmente numa escolha adequada dos parâmetros variantes no tempo e da matriz de ponderação do vetor score, S_t , responsável pelo mecanismo de atualização destes parâmetros.

5.2

Estimação e Inferência

A estimação dos parâmetros dos modelos GAS, assim como a dos demais modelos guiados por observações, é realizada a partir da função de log-verossimilhança para uma dada realização de uma série temporal y_t para $t = 1, \dots, n$. Entretanto, uma característica particular do GAS é a facilidade da avaliação desta função.

Assim, tem-se que os estimadores de máxima verossimilhança para cada realização de y_t são obtidos a partir de:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmax}} \sum_{t=1}^n l_t$$

onde $l_t = \ln p(y_t | f_t, F_t; \theta)$ é a função de log-verossimilhança instantânea para uma dada realização de y_t , sendo esta obtida a partir da função de densidade de y_t e Θ é o espaço paramétrico. A avaliação desta função é dada através do mecanismo de atualização f_t , o qual define o modelo GAS avaliado em valor particular θ .

Nessa dissertação, a maximização da função de log-verossimilhança foi obtida por métodos numéricos, tais como: BFGS e Nelder-Mead. Além disso, para o primeiro deles são utilizadas derivadas numéricas, uma vez que as derivadas da função de log-verossimilhança, as quais são obtidas de forma simultânea com os parâmetros variantes em t , podem ser complicadas na prática.

Creal *et al.* (2008) apresenta o cálculo dos erros padrão dos estimadores dos parâmetros fixos do modelo a partir da inversa da matriz de informação de Fisher avaliada no ponto ótimo.

Assim, tem-se que, sob condições de regularidade, o estimador de máxima verossimilhança é consistente e satisfaz a seguinte relação:

$$H^{\frac{1}{2}} \sqrt{n} (\hat{\theta} - \theta) \rightarrow N(0, I) \quad (5-7)$$

onde $H = -E[\frac{\partial^2 l_t}{\partial \theta \partial \theta'}] |_{\theta=\hat{\theta}}$.

A partir do cálculo das estimativas dos parâmetros estáticos do modelo, é possível avaliar sua significância estatística, utilizando (5-7).

5.3 Parametrização

Uma das grandes vantagens em se utilizar o modelo GAS é que o mesmo não se limita a uma parametrização específica de uma função de densidade, isto é, a especificação do modelo GAS se adapta a diferentes parametrizações de distribuições preditivas. Outro ganho deste processo é que o mesmo permite que um modelo restrito possa ser especificado por um modelo irrestrito através de uma parametrização adequada. No entanto, para cada parametrização escolhida, novas especificações dos modelos serão obtidas e isto pode tornar o cálculo da verossimilhança mais complexo (Creal *et al.*, 2013).

É comum a escolha de parametrizações do tipo $f_t = \ln(\alpha_t)$, na qual α_t é um parâmetro de uma determinada distribuição de densidade, pois em muitas situações o parâmetro α_t pode possuir somente valores positivos.

Para cada distribuição de uma variável de interesse y_t (Eq. (5-1)), pode-se obter uma nova parametrização a fim de se facilitar os cálculos necessários para aplicação do modelo GAS.

Segundo Creal *et al.* (2013), a partir da parametrização escolhida para $p(y_t|f_t, F_t; \theta)$ pode-se obter o vetor de score e a matriz de ponderação referentes a esta através da derivada da parametrização em relação ao parâmetro de interesse. Assim, dada uma função derivável, invertível e cuja inversa seja derivável do parâmetro variante no tempo $\tilde{f}_t = h(f_t)$ e considerando a matriz Jacobiana desta função em relação ao parâmetro de interesse, isto é, $\dot{h}_t = \frac{\partial h(f_t)}{\partial f_t}$, pode-se obter o vetor score e sua matriz de ponderação. Portanto, o vetor score pode ser obtido por

$$\tilde{\nabla}_t = \frac{\partial \ln p(y_t|f_t, F_t; \theta)}{\partial \tilde{f}_t} = (\dot{h}_t)^{-1} \nabla_t$$

e a matriz de informação de Fisher que é utilizada no cálculo da matriz de ponderação do score é dada por

$$\begin{aligned} \tilde{I}_{t|t-1} &= E_{t-1}[\tilde{\nabla}_t \tilde{\nabla}_t'] \\ &= E_{t-1}[(\dot{h}_t')^{-1} \nabla_t \nabla_t' (\dot{h}_t^{-1})'] \\ &= (\dot{h}_t')^{-1} E_{t-1}[\nabla_t \nabla_t'] (\dot{h}_t^{-1})' \\ &= (\dot{h}_t')^{-1} I_{t|t-1} (\dot{h}_t^{-1})' \end{aligned} \quad (5-8)$$

Logo, o score ponderado é dado por

$$s_t = \tilde{I}_{t|t-1} \tilde{\nabla}_t$$

Foi apresentado até aqui todo o arcabouço da formulação geral dos modelos GAS para quaisquer densidades preditivas. Assim, nas seções a seguir serão apresentados as formulações dos modelos GAS que consideram as distribuições Weibull, Gama e Beta como $p(y_t|f_t, F_t; \theta)$.

5.4

Modelo Univariado WEIBULL

Para a modelagem do vento, primeiramente foi proposto o modelo GAS para a distribuição Weibull, uma vez que diversos estudos como o de Brown *et al.* (1984) apresentam a Weibull como sendo a distribuição que melhor representa a distribuição da velocidade dos ventos.

Partindo da parametrização da distribuição Weibull proposta em Koopman *et al.* (2012) com parâmetro de forma k_1 e parâmetro de escala β , sugeriu-se uma nova parametrização onde a média condicional representasse o próprio parâmetro variante no tempo. Essa abordagem foi baseada na parametrização proposta por Matos (2013) para a distribuição Gama.

Assim, utilizando a distribuição de Weibull para a modelagem condicional de y_t obtém-se que

$$y_t|k_1, \kappa', \beta \sim Weibull(k_1, \beta\kappa')$$

na qual $\kappa' = \frac{1}{\Gamma(1+\frac{1}{k_1})}$ e assim a função de densidade Weibull é dada por:

$$p(y_t|k_1, \kappa', \beta) = \frac{k_1}{\beta\kappa'} \left(\frac{y_t}{\beta\kappa'}\right)^{k_1-1} \exp - \left(\frac{y_t}{\beta\kappa'}\right)^{k_1}, \quad y_t > 0, \quad \beta, k_1, \kappa' > 0$$

Mostra-se que, a média e a variância são dados por:

$$E[y_t|k_1, \kappa', \beta] = \beta, \quad V[y_t|k_1, \kappa', \beta] = (\kappa')^2 \beta^2 \left[\Gamma\left(1 + \frac{2}{k_1}\right) - (\kappa')^{-2} \right];$$

nessa segunda, $\Gamma(\cdot)$ é a função Gama.

A função de log-verossimilhança é dada por:

$$\begin{aligned} l_t &= \sum_{t=12}^n \ln p(y_t|k_1, \kappa', \beta) \\ &= \sum_{t=12}^n \left[\ln k_1 - \ln \kappa' + (k_1 - 1) (\ln y_t - \ln \kappa') - \left(\frac{y_t}{\beta\kappa'}\right)^{k_1} \right], \quad \theta \in \Theta. \end{aligned}$$

5.4.1

Modelo Weibull - SARIMA com o parâmetro β variante no tempo

Escolhendo neste caso o parâmetro de escala, o qual é a média condicional de y_t tem-se que uma parametrização natural seria $\tilde{f}_t = \ln \beta_t$, preferível à $f_t = \beta_t$ como abordado na Seção 5.3. Assim tem-se que a derivada de $\ln p(y_t|f_t, F_t; \theta)$, que é a contribuição de cada valor de y_t para o vetor *score*, em relação ao parâmetro variante no tempo β_t é dada por:

$$\frac{\partial \ln p(y_t|f_t, F_t; \theta)}{\partial \beta_t} = \frac{k_1}{\beta_t} \left[\left(\frac{y_t}{\beta_t \kappa'} \right)^{k_1} - 1 \right]$$

A partir desta calcula-se o vetor *score* para a parametrização escolhida onde $h(f_t) = \ln \beta_t \Rightarrow \dot{h}_t = \frac{1}{\beta_t}$ e este é dado por:

$$\tilde{\nabla}_t = \frac{\partial \ln p(y_t|f_t, F_t; \theta)}{\partial \tilde{f}_t} = (\dot{h}_t)^{-1} \nabla_t = k_1 \left[\left(\frac{y_t}{\beta_t \kappa'} \right)^{k_1} - 1 \right]$$

Para o cálculo da matriz de informação de Fisher, necessita-se das derivadas parciais de segunda ordem. Como o parâmetro variante no tempo é β_t , isto é, $f_t = \beta_t$, o interesse está no elemento da matriz $I_{t|t-1}$ referente a este parâmetro, sendo assim tem-se que:

$$I_{t|t-1} = -E \left[\frac{\partial^2 \ln p(y_t|f_t, F_t; \theta)}{\partial f_t^2} \right] = \frac{k_1^2}{\beta_t^2}$$

e, assim em relação à \dot{h}_t obtida de $\tilde{f}_t = \ln \beta_t$, pela Eq. (5-8), obtém-se que:

$$\tilde{I}_{t-1}(\beta) = k_1^2$$

Por fim, calculam-se os *scores* ponderados:

$$s_t = \left(\frac{y_t}{\beta_t \kappa'} \right)^{k_1} - 1, \quad s_t = \frac{1}{k_1} \left[\left(\frac{y_t}{\beta_t \kappa'} \right)^{k_1} - 1 \right], \quad s_t = k_1 \left[\left(\frac{y_t}{\beta_t \kappa'} \right)^{k_1} - 1 \right]$$

com as quais forma usadas as seguintes ponderações: raiz quadrada da pseudo-inversa da matriz $I_{t|t-1}$, pseudo-inversa da matriz $I_{t|t-1}$ e matriz identidade I, respectivamente.

Os modelos Weibull-SARIMA propostos com o parâmetro β_t variante no

tempo possuem a seguinte equação de evolução, onde $\beta_t = \exp \tilde{f}_t$:

$$\begin{aligned} \tilde{f}_{t+1} = \omega &+ a_1 s_t + a_2 s_{t-1} + a_3 s_{t-2} + a_4 s_{t-10} + a_5 s_{t-11} \\ &+ b_1 \tilde{f}_t + b_2 \tilde{f}_{t-1} + b_3 \tilde{f}_{t-2} + b_4 \tilde{f}_{t-10} + b_5 \tilde{f}_{t-11} \end{aligned} \quad (5-9)$$

Neste, buscou-se capturar as dinâmicas de curto prazo pelos *lags* 1,2 e 3, bem como dinâmicas de longo prazo, caracterizando a sazonalidade, através dos *lags* 11 e 12.

Devido aos resultados não satisfatórios obtidos através do ajuste do Modelo Weibull-SARIMA, não será apresentada a metodologia referente à estrutura de evolução CNO.

5.5 Modelo Univariado GAMA

Com o intuito de propor uma nova distribuição para as séries de vento, utilizou-se a distribuição Gama com parâmetros de forma α e escala λ por esta possuir suporte $[0, \infty)$, assim como a distribuição Weibull, o qual é adequado ao intervalo de variação das séries em estudo. Utilizando as duas configurações de evolução dos modelos GAS, SARIMA e CNO, a parametrização escolhida para a Gama foi a proposta em Matos (2013), onde a média condicional de y_t é o próprio parâmetro de escala, λ . Assim, optou-se por utilizar a seguinte distribuição condicional para y_t :

$$y_t | \alpha, \lambda \sim Gama \left(\alpha, \frac{\lambda}{\alpha} \right),$$

ou seja,

$$p(y_t | \alpha, \lambda) = \frac{1}{\Gamma(\alpha)(\alpha^{-1}\lambda)^\alpha} y_t^{\alpha-1} \exp \left(-\frac{\alpha}{\lambda} y_t \right), \quad y_t > 0, \quad \alpha, \lambda > 0 \quad (5-10)$$

cuja esperança e variância são:

$$E[y_t | \alpha, \lambda] = \lambda, \quad V[y_t | \alpha, \lambda] = \frac{\lambda^2}{\alpha}$$

e a função de log-verossimilhança é dada por:

$$\begin{aligned} l_t &= \sum_{t=12}^n \ln p(y_t | \alpha, \lambda) \\ &= \sum_{t=12}^n \left[\ln \Gamma(\alpha) + \alpha \ln(\alpha) - \alpha \ln(\lambda) + (\alpha - 1) \ln y_t - \frac{\alpha}{\lambda} y_t \right], \theta \in \Theta. \end{aligned}$$

5.5.1

Modelo Gama - SARIMA com o parâmetro λ variante no tempo

Primeiramente, optou-se pela utilização de um modelo GAS com evolução do tipo SARIMA com apenas um dos parâmetros da Gama variante no tempo. Escolhendo neste caso o parâmetro de escala, o qual é a média condicional de y_t tem-se que uma parametrização natural seria $\tilde{f}_t = \ln \lambda_t$, preferível à $f_t = \lambda_t$, análoga à escolhida para a distribuição de Weibull.

Assim, o interesse está em calcular o vetor *score* e a matriz de informação de Fisher, que será utilizada no cálculo final do *score* ponderado, foco principal da modelagem GAS. Para esta abordagem necessita-se apenas da derivada em relação ao parâmetro variante λ_t .

Sendo assim, a derivada de $\ln p(y_t|f_t, F_t; \theta)$, que é a contribuição de cada valor de y_t para o vetor *score*, em relação à λ_t é dada por:

$$\frac{\partial \ln p(y_t|f_t, F_t; \theta)}{\partial \lambda_t} = \frac{\alpha}{\lambda_t} \left(\frac{y_t}{\lambda_t} - 1 \right)$$

De modo análogo, pode-se obter a derivada em relação ao parâmetro α , a qual é apresentada no Apêndice A.

A partir desta calcula-se o vetor *score* para a parametrização escolhida onde $h(f_t) = \ln \lambda_t \Rightarrow \dot{h}_t = \frac{1}{\lambda_t}$ e este é dado por:

$$\tilde{\nabla}_t = \frac{\partial \ln p(y_t|f_t, F_t; \theta)}{\partial \tilde{f}_t} = (\dot{h}_t)^{-1} \nabla_t = \alpha \left(\frac{y_t}{\lambda_t} - 1 \right)$$

Como o parâmetro variante no tempo é λ_t , isto é, $f_t = \lambda_t$ o interesse está no elemento da matriz $I_{t|t-1}$ referente a este parâmetro, sendo assim, tem-se que:

$$I_{t|t-1} = -E \left[\frac{\partial^2 \ln p(y_t|f_t, F_t; \theta)}{\partial f_t^2} \right] = -E \left[\frac{\alpha}{\lambda_t^2} - \frac{\alpha y_t}{\lambda_t^3} \right] = \frac{\alpha}{\lambda_t^2}$$

e, assim em relação à \dot{h}_t obtida de $\tilde{f}_t = \ln \lambda_t$, mostra-se que:

$$\tilde{I}_{t-1}(\lambda) = \alpha$$

A partir das fórmulas descritas acima, pode-se obter o *score* ponderado para algumas escolhas de ponderação como:

$$s_t = \sqrt{\alpha} \left(\frac{y_t}{\lambda_t} - 1 \right), \quad s_t = \left(\frac{y_t}{\lambda_t} - 1 \right), \quad s_t = \alpha \left(\frac{y_t}{\lambda_t} - 1 \right)$$

as quais são obtidas a partir das seguintes ponderações: raiz quadrada da pseudo-inversa da matriz $I_{t|t-1}$, pseudo-inversa da matriz $I_{t|t-1}$ e matriz

identidade I , respectivamente.

Os modelos Gama-SARIMA propostos com o parâmetro λ_t variante no tempo possuem a seguinte equação de evolução para \tilde{f}_t , onde $\lambda_t = \exp \tilde{f}_t$:

$$\begin{aligned} \tilde{f}_{t+1} = \omega &+ a_1 s_t + a_2 s_{t-1} + a_3 s_{t-2} + a_4 s_{t-10} + a_5 s_{t-11} \\ &+ b_1 \tilde{f}_t + b_2 \tilde{f}_{t-1} + b_3 \tilde{f}_{t-2} + b_4 \tilde{f}_{t-10} + b_5 \tilde{f}_{t-11} \end{aligned} \quad (5-11)$$

Neste buscou-se capturar as dinâmicas de curto prazo pelos *lags* 1,2 e 3, bem como dinâmicas de longo prazo, caracterizando a sazonalidade, através dos *lags* 11 e 12.

A inclusão de variáveis exógenas no modelo foi realizada diretamente no cálculo da média condicional λ_t na forma aditiva, como proposto por Creal *et al.* (2013), apresentado a seguir:

$$\tilde{f}_t = h(\lambda_t) = - \left[\sum_{k=1}^r \phi_k g(x_{t-k+1}) \right] + \ln \lambda_t$$

ou de forma equivalente

$$\lambda_t = \exp \left[\tilde{f}_t + \sum_{k=1}^r \phi_k g(x_{t-k+1}) \right]$$

na qual \tilde{f}_t é o parâmetro variante no tempo e x_t é o vetor de variáveis exógenas a serem incorporadas no modelo.

5.5.2 Modelos Gama - CNO

Os modelos GAS com evolução por componentes não-observáveis propostos nessa dissertação buscam capturar a dinâmica das séries de vento através de uma componente de sazonalidade (Eq. (5-12)), característica que é presente e se destaca nas séries vento, bem como uma componente autoregressiva de ordem 1 (Eq. (5-13)), como é apresentado a seguir.

$$\tilde{f}_{1,t+1} = - \sum_{k=1}^{11} \tilde{f}_{1,t-k} + a_1 s_t, \quad (5-12)$$

$$\tilde{f}_{2,t+1} = \theta \tilde{f}_{2,t} + a_2 s_t, \quad (5-13)$$

nas quais θ , a_1 e a_2 parâmetros estáticos relacionados aos vetores \tilde{f}_t e s_t ,

estimados por máxima verossimilhança. E assim, a função de atualização do modelo GAS é dada por:

$$\tilde{f}_{t+1} = \omega + \tilde{f}_{1,t+1} + \tilde{f}_{2,t+1} \quad (5-14)$$

$$\tilde{f}_{t+1} = \begin{pmatrix} \tilde{f}_{1,t+1} \\ \tilde{f}_{2,t+1} \end{pmatrix} = \begin{pmatrix} w \\ 0 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} s_t + \begin{pmatrix} -\sum_{k=1}^{11} \tilde{f}_{1,t-k} & 0 \\ 0 & \theta \tilde{f}_{2,t} \end{pmatrix} \quad (5-15)$$

5.6 Modelo Univariado BETA

A distribuição Beta é comumente utilizada na versão em que seu suporte é o intervalo $[0,1]$. Esta possui dois parâmetros de forma, β e α . Matos (2013) propõe a utilização da distribuição Beta com suporte $[0,k]$ para modelagem do fator de capacidade de uma usina eólica no Nordeste. Leite *et al.* (2007) propõe a utilização da distribuição Beta para o ajuste da série de velocidade do vento para uma região no estado do Paraná. No mesmo é possível observar que a distribuição Beta apresenta bons ajustes às séries de velocidade do vento. Tal análise é feita a partir do teste de Kolmogorov-Sminorv, o qual forneceu resultados satisfatórios. Vale ressaltar que, assim como nessa dissertação, os dados modelados eram médias diárias e não horárias, como em muitos trabalhos que utilizam a distribuição de Weibull. Este fato motivou o utilização da mesma no processo de modelagem dos ventos na região de estudo.

Assim, seja a distribuição condicional para y_t dada por:

$$\frac{y_t}{k} | \beta, \alpha \sim Beta(\beta, \alpha),$$

ou seja,

$$p(y_t | \beta, \alpha) = \frac{1}{k^{\beta+\alpha-1}} \frac{\Gamma(\beta + \alpha)}{\Gamma(\beta)\Gamma(\alpha)} y_t^{\beta-1} (k - y_t)^{\alpha-1}, \quad y_t \in [0, k], \quad (5-16)$$

$$\beta, \alpha > 0$$

cuja esperança e variância são dados respectivamente por:

$$E[y_t | \beta, \alpha] = k \frac{\beta}{\beta + \alpha}, \quad V[\beta, \alpha] = k^2 \frac{\beta\alpha}{(\beta + \alpha)^2(\beta + \alpha + 1)}$$

e a função de log-verossimilhança é dada por:

$$\begin{aligned} l_t &= \sum_{t=12}^n \ln p(y_t | \beta, \alpha) \\ &= \sum_{t=12}^n -(\beta + \alpha - 1) \ln k + \ln \left(\frac{\Gamma(\beta + \alpha)}{\Gamma(\beta)\Gamma(\alpha)} \right) + (\beta - 1) \ln y_t \\ &\quad + (\alpha - 1) \ln(k - y_t), \quad \theta \in \Theta. \end{aligned}$$

5.6.1

Modelo Beta - SARIMA com o parâmetro β variante no tempo

Escolhendo o parâmetro de forma β como variante no tempo, à semelhança do parâmetro λ na distribuição Gama, tem-se que uma parametrização adequada seria $\tilde{f}_t = \ln \beta_t$.

Sendo assim, a derivada de $\ln p(y_t | f_t, F_t; \theta)$, que é a contribuição de cada valor de y_t para o vetor *score*, em relação à β_t é dada por:

$$\frac{\partial \ln p(y_t | f_t, F_t; \theta)}{\partial \beta_t} = \ln y_t - \ln k - \psi_1(\beta_t) + \psi_1(\beta_t + \alpha)$$

A partir desta calcula-se o vetor *score* para a parametrização escolhida onde $h(f_t) = \ln \beta_t \Rightarrow \dot{h}_t = \frac{1}{\beta_t}$ e este é dado por:

$$\tilde{\nabla}_t = \frac{\partial \ln p(y_t | f_t, F_t; \theta)}{\partial \tilde{f}_t} = (\dot{h}_t)^{-1} \nabla_t = \beta_t \{ \ln y_t - [\ln k + \psi_1(\beta_t) - \psi_1(\beta_t + \alpha)] \}$$

Como o parâmetro variante no tempo é β_t , isto é, $f_t = \beta_t$, o interesse é o elemento da matriz referente a este parâmetro, sendo assim tem-se que:

$$I_{t|t-1} = -E \left[\frac{\partial^2 \ln p(y_t | f_t, F_t; \theta)}{\partial f_t^2} \right] = \beta^2 \{ \psi_2(\beta_t) - \psi_2(\beta_t + \alpha) \}$$

e, assim em relação à \dot{h}_t obtida $\tilde{f} = \ln \beta_t$, mostra-se que:

$$\tilde{I}_{t-1}(\beta) = \psi_2(\beta_t) - \psi_2(\beta_t + \alpha)$$

Assim, pode-se obter o *score* ponderado para algumas escolhas de ponderação, como:

$$s_t = \frac{\ln y_t - [\ln k + \psi_1(\beta_t) - \psi_1(\beta_t + \alpha)]}{\beta_t^{1-2d} [\psi_2(\beta_t) - \psi_2(\beta_t + \alpha)]^{1-d}}, \quad d = 1/2, 0, 1. \quad (5-17)$$

onde estas são obtidas a partir das seguintes ponderações: raiz quadrada da pseudo-inversa da matriz $I_{t|t-1}$, pseudo-inversa da matriz $I_{t|t-1}$ e matriz identidade I , respectivamente.

A inclusão de variáveis exógenas no modelo se deu da mesma forma, descrita anteriormente para a distribuição Gama. Assim, tem-se que:

$$\tilde{f}_t = h(\beta_t) = - \left[\sum_{k=1}^r \phi_k g(x_{t-k+1}) \right] + \ln \beta_t$$

ou de forma equivalente

$$\beta_t = \exp \left[\tilde{f}_t + \sum_{k=1}^r \phi_k g(x_{t-k+1}) \right]$$

onde \tilde{f}_t é o parâmetro variante no tempo e x_t é o vetor de variáveis exógenas a serem incorporadas no modelo.

A configuração do modelo Beta-CNO é a mesma utilizada no modelo Gama-CNO, descrita na subseção 5.5.2 nas equações (5-11) e (5-12), onde o mesmo possui as componentes autoregressiva de ordem 1 e de sazonalidade.

5.7

Critérios de seleção e medidas de aderência do modelo

Para a escolha do melhor modelo são utilizados os critérios de seleção de modelos *Akaike Information Criterion* (AIC) e *Bayesian Information Criterion* (BIC) definidos por:

$$AIC = -2 \ln(L) + 2\nu,$$

$$BIC = -2 \ln(L) + \nu \ln(n),$$

nas quais L é função de verossimilhança do modelo avaliada no máximo, ν é o número de parâmetros estimados no modelo e n o tamanho da série.

Ao se utilizar os critérios citados anteriormente para a comparação entre modelos, o melhor modelo é aquele que possui os menores valores de AIC ou BIC.

Após a especificação dos modelos a serem testados, é necessário realizar uma análise do poder preditivo dos mesmos. Para isto, são utilizados tanto recursos gráficos, quanto medidas de aderência, a fim de se analisar se o modelo proposto conseguiu captar a dinâmica da série, isto é, se o mesmo possuiu um bom ajuste à série temporal em estudo. As medidas utilizadas neste trabalho

foram: RMSE (*Root Mean Square*), MAE (*Mean Absolute Error*), MASE (*Mean Absolute Scaled Error*), sMAPE (*Symmetric Mean Absolute Percent Error*) e *Pseudo R*², que são dadas pelas fórmulas abaixo.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_{t|t-1})^2}$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_{t|t-1}|$$

$$MASE = \left(\frac{n-1}{n}\right) \frac{\sum_{t=1}^n |y_t - \hat{y}_{t|t-1}|}{\sum_{t=2}^n |y_t - y_{t-1}|}$$

$$sMAPE = \left(\frac{1}{n}\right) \frac{\sum_{t=1}^n |y_t - \hat{y}_{t|t-1}|}{\sum_{t=2}^n |y_t + \hat{y}_{t|t-1}|}$$

$$PseudoR^2 = [corr(y_t, \hat{y}_{t|t-1})]^2$$

5.8

Teste de Diebold-Mariano para comparação entre previsões

Um dos procedimentos mais importantes antes da tomada de decisão em relação à previsão de uma série temporal está na comparação entre previsões geradas por diferentes modelos. Assim, em Diebold & Mariano (1995) é proposto um teste (Diebold-Mariano) para a comparação entre previsões. Este é realizado através dos resíduos obtidos em cada um dos modelos e busca verificar se ambos os modelos possuem o mesmo grau de aderência ou, ainda se um dos modelos fornece previsões mais, ou menos precisas, do que o outro. Sendo assim, teste pode ser realizado em sua versão bilateral ou unilateral.

Considere as previsões $\{\hat{y}_{1,t+h|t}\}_{t=1}^T$ e $\{\hat{y}_{2,t+h|t}\}_{t=1}^T$ obtidas de diferentes modelos e, $\{e_{1,t+h|t}\}_{t=1}^T$ e $\{e_{2,t+h|t}\}_{t=1}^T$ os resíduos ou erros de previsão destes, respectivamente.

As hipóteses nula e alternativa do teste bilateral são definidas como:

$$H_0 : E[g(\{e_{1,t+h|t}\}_{t=1}^T)] = E[g(\{e_{2,t+h|t}\}_{t=1}^T)], \forall t$$

$$H_1 : E[g(\{e_{1,t+h|t}\}_{t=1}^T)] \neq E[g(\{e_{2,t+h|t}\}_{t=1}^T)]$$

nas quais $g(\{e_{i,t+h|t}\}_{t=1}^T)$ para $i = 1, 2$ é definida como uma função perda, a qual deve ser escolhida pelo usuário. Nessa dissertação foi considerada a função $g(\{e_{i,t+h|t}\}_{t=1}^T) = e_{i,t+h|t}^2$.

Assim o teste Dielbold-Mariano é baseado na perda diferencial d_t que é dada por:

$$d_t = g(e_{1,t+h|t}) - g(e_{2,t+h|t})$$

A estatística de teste é

$$S = \frac{\bar{d}}{(\widehat{LVR}_{\bar{d}}/T)^{\frac{1}{2}}} \sim N(0, 1)$$

na qual

$$\bar{d} = \frac{1}{T} \sum_{t=1}^T d_t \quad \text{e} \quad LVR = \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j, \quad \gamma_j = cov(d_t, d_{t-j})$$

sendo LVR uma estimativa assintoticamente consistente da variância de $\sqrt{T}\bar{d}$.

Logo, a hipótese nula será rejeitada ao nível de 5% de significância se o p-valor do teste for menor que 5%.

Vale ressaltar que devido ao foco dessa dissertação ser o ajuste de uma série temporal e não a previsão desta, o horizonte de previsão do teste será 1, ou seja, $h = 1$, que caracteriza a previsão *in-sample*.

5.9

Resíduos do modelo

A análise dos resíduos é uma das partes mais importantes do diagnóstico de um modelo. Na modelagem de séries lineares e Gaussianas, os resíduos utilizados para tal análise são os resíduos de Pearson. Para os modelos GAS, que modelam séries não lineares e não-Gaussianas, não existe uma definição consensual para os resíduos a serem avaliados. Recentemente foi proposta na literatura (Kalliovirta, 2009) a utilização dos resíduos quantílicos a fim de superar a não adequação dos resíduos convencionais na ausência de normalidade e linearidade das séries.

O resíduo quantílico teórico é dado por:

$$R_{t,\theta} = \Phi^{-1}(F(y_t|f_t, F_t; \theta)) \quad (5-18)$$

onde $\Phi^{-1}(\cdot)$ é a inversa da função de distribuição normal e $F(y_t|f_t, F_t; \theta)$ é a função de distribuição condicional do processo.

Sendo assim, nessa dissertação foram utilizados os resíduos quantílicos (Eq. (5-18)) para a investigação da correta especificação dos modelos. A análise destes resíduos dá-se através de testes estatísticos que buscam averiguar a presença de autocorrelação serial, heterocedasticidade condicional e não-normalidade dos mesmos. Frequentemente são utilizados na literatura dos modelos para séries gaussianas, testes como: Ljung-Box para a autocorrelação serial e heterocedasticidade condicional; e Jarque Bera para a normalidade. No entanto, para a análise de resíduos quantílicos é necessária a utilização de testes adequados à natureza destes. Todavia, devido a dificuldade na implementação de testes para os resíduos quantílicos, fato este abordado por Matos (2013), foram utilizados os testes convencionais citados anteriormente nessa seção.

5.10

Mensuração do efeito das variáveis exógenas nos modelos GAS

A partir do ajuste dos modelos descritos anteriormente, é possível estimar o efeito exercido pelas variáveis exógenas nas séries em estudo. O cálculo deste efeito para os modelos Gama e Beta é realizado através da variação percentual na média condicional de y_{t-1} para y_t . Este procedimento para ambas as distribuições é apresentado a seguir.

Seja o modelo univariado Gama com a inclusão da variável exógena x_t , cuja média condicional é igual a λ_t :

$$\begin{aligned} E[y_t | f_t, F_t; \theta] &= \lambda_t = e^{(\tilde{f}_t + \phi x_t)}, \\ E[y_{t-1} | f_{t-1}, F_{t-1}; \theta] &= \lambda_{t-1} = e^{(\tilde{f}_{t-1} + \phi x_{t-1})} \end{aligned}$$

Desta forma, a variação percentual na média condicional de y_t será dada por:

$$\begin{aligned} \Delta\% &= \left(\frac{E[y_t | f_t, F_t; \theta]}{E[y_{t-1} | f_{t-1}, F_{t-1}; \theta]} - 1 \right) .100 \\ &= \left(\frac{\lambda_t}{\lambda_{t-1}} - 1 \right) .100 \\ &= \left(e^{(\tilde{f}_t - \tilde{f}_{t-1}) + \phi(x_t - x_{t-1})} - 1 \right) .100 \\ &= \left(e^{(\Delta\tilde{f}_t) + \phi(\Delta x_t)} - 1 \right) .100 \end{aligned}$$

e desprezando o efeito das demais variáveis, isto é, fazendo $\Delta\tilde{f}_t = 0$, o efeito marginal da variável x_t será dado por

$$\Delta\% = (e^{\phi(\Delta x_t)} - 1) .100, \quad (5-19)$$

na qual ϕ é um parâmetro conhecido, estimado por máxima verossimilhança.

Para os modelos Beta, o cálculo do efeito da variável exógena x_t fornece uma forma menos elegante que a dos modelos Gama, uma vez que na parametrização do mesmo, a média condicional de y_t não é função apenas de um dos parâmetros da distribuição.

Seja o modelo univariado Beta com a inclusão da variável exógena x_t , cuja média condicional é igual a $k \frac{\beta_t}{\beta_t + \alpha}$ tem-se que:

$$\begin{aligned} E[y_t | f_t, F_t; \theta] &= k \frac{\beta_t}{\beta_t + \alpha} = k \frac{e^{(\phi x_t + \tilde{f}_t)}}{e^{(\phi x_t + \tilde{f}_t)} + \alpha}, \\ E[y_{t-1} | f_{t-1}, F_{t-1}; \theta] &= k \frac{\beta_{t-1}}{\beta_{t-1} + \alpha} = k \frac{e^{(\phi x_{t-1} + \tilde{f}_{t-1})}}{e^{(\phi x_{t-1} + \tilde{f}_{t-1})} + \alpha} \end{aligned}$$

Sendo assim, a variação percentual na média condicional de y_t é dada por:

$$\begin{aligned} \% \Delta &= \left[\frac{E[y_t | f_t, F_t; \theta]}{E[y_{t-1} | f_{t-1}, F_{t-1}; \theta]} - 1 \right] .100 \\ &= \left[\frac{1 + \alpha e^{-\phi x_{t-1} - \tilde{f}_{t-1}}}{1 + \alpha e^{-\phi x_t - \tilde{f}_t}} - 1 \right] .100 \end{aligned} \quad (5-20)$$

Usando que $\Delta x_t = x_t - x_{t-1}$ então

$$\begin{aligned} x_t &= \Delta x_t + x_{t-1} \\ \tilde{f}_t &= \Delta \tilde{f}_t + \tilde{f}_{t-1} \end{aligned}$$

e substituindo na Eq. (5-20) tem-se que

$$\Delta \% = \left[\frac{1 + \alpha e^{-\phi x_{t-1} - \tilde{f}_{t-1}}}{1 + \alpha e^{-\phi[\Delta x_t + x_{t-1}] - [\Delta \tilde{f}_t + \tilde{f}_{t-1}]}} - 1 \right] .100$$

Como o interesse está no efeito de x_t , bem como no modelo Gama, $\Delta \tilde{f}_t$ é igual a zero e, assim a variação percentual será dada por:

$$\Delta \% = \left[\frac{1 + \alpha e^{-\phi x_{t-1} - \tilde{f}_{t-1}}}{1 + \alpha e^{-\phi[\Delta x_{t-1} + x_{t-1}] - \tilde{f}_{t-1}}} - 1 \right] .100 \quad (5-21)$$

onde, x_{t-1} é a média dos valores da variável exógena num determinado mês j e \tilde{f}_{t-1} é a média dos \tilde{f}_t 's no mês j , $j = 1, \dots, 12$.

Em ambos os modelos é possível encontrar o efeito marginal da variável regressora x_t através de variações da mesma, isto é, fazendo $\Delta x_t = C$, onde C são valores arbitrários para a variação da variável regressora.

Em suma, a metodologia apresentada nesse capítulo é uma alternativa à modelagem de séries temporais de forma mais aprimorada, tornando flexível sua utilização em variáveis que possuem, ou não, características gaussianas. Sendo assim, essa dissertação propõe a utilização desse arcabouço em séries meteorológicas, as quais possuem como maior característica, a sazonalidade. Contudo, trabalhos que utilizam os modelos GAS para a modelagem de séries que possuem comportamento sazonal ainda são escassos na literatura. Um estudo que faz uso de tal abordagem é Matos (2013), que foi descrito brevemente no capítulo anterior. Neste trabalho, devido à presença de estrutura sazonal nas séries, foram definidos modelos GAS de ordens p e q superiores, por exemplo, GAS (1, 11), o que torna a estimação dos parâmetros mais complexa, uma vez que o número destes se torna mais alto, conforme as ordens de p e q aumentam.

6 Resultados

As séries de vento utilizadas nessa dissertação são oriundas de uma fonte de reanálise ERA-*Interim* do *European Center for Medium Range Weather Forecast* (ECMRWF), na qual estas são geradas através de um método científico capaz de gerar séries em uma escala global a partir da assimilação de dados, incorporando dados observados a modelos numéricos capazes de simular diversos aspectos da Terra. (Martins *et al.* (2008); *Reanalysis* (2013)). Estas séries foram disponibilizadas pela empresa 3TIER.

A decisão pela utilização dessas séries foi em razão da necessidade de uma série longa de dados para o estudo da relação do vento com o El Niño, fenômeno este que começou a ser monitorado a partir de 1950. Os empreendimentos eólicos se consolidaram após esse período, o que justifica a ausência de séries de vento longas para a região. Outra importante razão que tornou inviável a utilização de dados observados é que estes apresentam períodos sequentes com muitas observações faltantes e não se tinha certeza acerca de sua confiabilidade.

Os dados de vento se apresentavam em frequência horária para doze usinas eólicas localizadas na Zona Litorânea Norte-Nordeste definida em Amarante (2001a), a qual compreende a região entre a faixa costeira que possui cerca de 100 km de largura, que se inicia no extremo norte da costa do Amapá e se estende ao Cabo de São Roque com inclusão da cidade de Natal, no estado do Rio Grande do Norte. As séries de velocidade do vento (m/s) modeladas a 100 metros de altura compreendem o período de 01/01/1982 a 31/12/2012.

Para a investigação da relação da velocidade do vento com o ENOS, foram utilizadas séries mensais de anomalias e valores brutos de TSM da região Niño 3.4, retirados do site do NOAA [<http://www.cpc.ncep.noaa.gov>]. A escolha desta série, para representar o fenômeno El Niño, se deve em razão da relação direta desta com a definição do fenômeno, fato já abordado no Capítulo 3. Além disso, a região Niño 3.4 é considerada uma região que fornece boas medidas de mudanças importantes na TSM (NOAA, 2014).

É possível observar na Figura 6.1, de forma resumida, todos os procedimentos realizados no presente trabalho que serão apresentados no decorrer

desse capítulo.

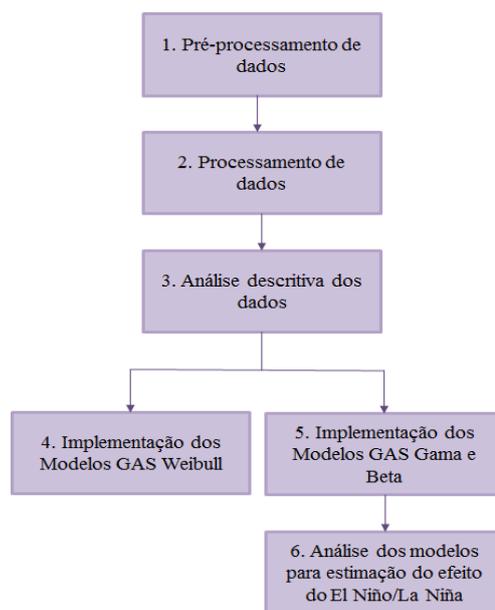


Figura 6.1: Fluxograma das etapas realizadas na modelagem.

O pré-processamento dos dados engloba as etapas de coleta dos dados de vento e TSM e validação dos dados de vento. Esta validação foi realizada a fim de se explorar a representatividade dos dados de reanálise, comparando-os com dados de bóias que estão localizadas próximas às usinas eólicas em estudo. Assim, foi feita uma extração de várias amostras de ambas as séries em concordância com o período disponível. Esta análise forneceu coeficientes de correlação entre 0,80 e 0,90. Com isso, os dados de reanálise foram considerados representativos para o estudo. Além da validação realizou-se uma análise de correlação entre as usinas a fim de verificar a possível semelhanças do comportamento dos ventos nestas.

A fase de processamento dos dados consistiu da conversão dos dados horários de velocidade do vento para frequência mensal. Desta forma, foram obtidas médias horárias para cada mês ao longo dos 31 anos de dados para que se pudesse relacioná-las com o ENOS, um fenômeno de baixa frequência. Como, neste caso, a amostra é grande suficientemente e os desvios padrão das variáveis possuem valores baixos, a média é um estimador coerente e de fácil interpretação para o estudo. Assim, foram utilizadas 372 observações considerando o período de análise.

Após o processamento dos dados, foi realizada uma análise descritiva das séries de vento e TSM através da observação das séries ao longo do tempo, função de autocorrelação, histograma e Boxplot, além de testes de estacionariedade e estimação do Periodograma das mesmas.

Para a implementação dos modelos GAS, foram testadas distribuições univariadas do tipo Weibul, Gama e Beta, uma vez que a velocidade do vento possui apenas valores reais positivos sendo, portanto, coerente com o suporte destas distribuições. O ajuste da distribuição de Weibull inicialmente se deve a grande utilização da mesma como melhor ajuste para a distribuição da velocidade do vento, como abordado em Justus *et al.* (1976).

Em seguida, os modelos GAS para as distribuições Gama e Beta com estruturas de evolução SARIMA e CNO foram analisados e suas previsões *in-sample* comparadas através do teste Diebold-Mariano, apresentado do Capítulo 5.

Por fim, foi calculado o efeito do El Niño/ La Niña através da variável exógena TSM da região Niño 3.4 para os modelos GAS com distribuição Gama e Beta.

Toda implementação desse trabalho foi realizada no *software* estatístico R versão 3.0.2.

6.1

Análise descritiva

As doze usinas de energia eólica escolhidas para esse estudo apresentam as melhores características para a geração eólica devido à localização no litoral dos estados do Ceará e Rio Grande do Norte, a qual sofre pouca influência do atrito e rugosidade do relevo (Amarante *et al.*, 2001a). Além disso, muitas destas estão localizadas próximas entre si e isto se dá em alguns casos, em função de algumas usinas pertencerem ao mesmo parque eólico. Esta característica motivou a realização de uma análise da correlação entre as usinas a fim de se reduzir o número destas a ser considerado no trabalho.

A análise da correlação entre as variáveis que caracterizam as usinas mostrou que o comportamento do vento na localização destas é muito semelhante, uma vez que todas as correlações foram superiores a 0,81. A correlação foi estimada considerando as variáveis: velocidade e direção do vento a 100 metros de altitude, além da temperatura e pressão do ar a 2 metros de altitude. Assim, selecionou-se uma das usinas, a qual se acredita ser suficiente para a representação do comportamento do vento na região de interesse.

6.1.1

Análise descritiva dos dados de vento

Analisando a série de velocidade do vento anual da usina eólica em estudo (Figura 6.2) localizada no estado do Ceará na latitude: -2.908 e longitude:

-41.039, nota-se que em grande parte dos anos a intensidade dos ventos permaneceu acima de 7 m/s. Observa-se ainda na Figura 6.2 que nos anos de 1985 e 1989, anos de ocorrência do fenômeno La Niña, a velocidade do vento possuiu os menores valores quando comparado aos demais anos da série.

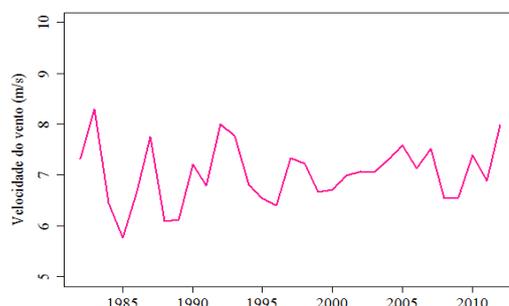


Figura 6.2: Médias anuais de velocidade do vento para a usina do Nordeste para o período de 1982 a 2012.

Na Figura 6.3 pode-se observar as médias mensais para a usina de estudo. Nota-se que o período com velocidades médias mensais superiores na região coincide com o período de seca da mesma, como abordado no Capítulo 2. Assim, no segundo semestre de cada ano, que corresponde às estações de inverno e primavera (períodos de seca), os ventos apresentam alta intensidade.

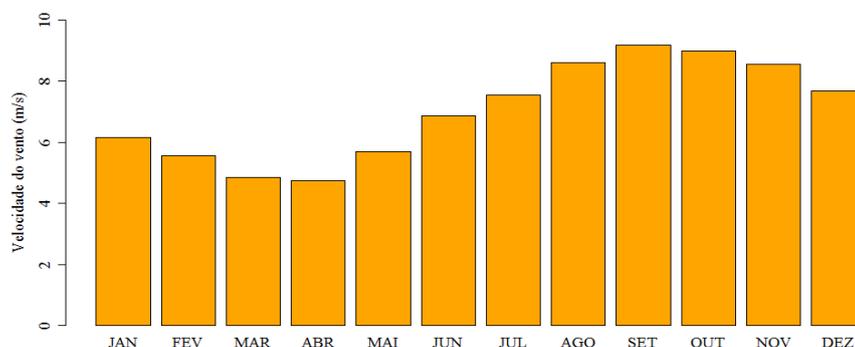


Figura 6.3: Valores mensais médios para a velocidade do vento no período de 1982 a 2012.

Para a análise descritiva da direção do vento, foi realizada uma representação da mesma pela rosa dos ventos (Figura 6.4) para o período de dados analisados.

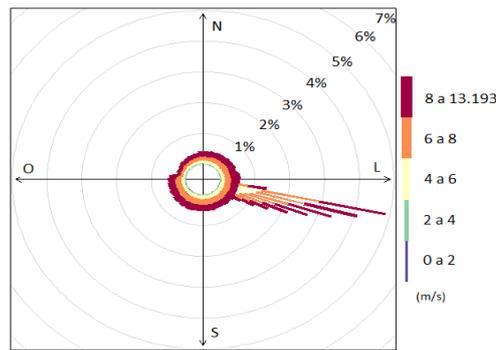


Figura 6.4: Rosa dos ventos da usina eólica.

Através da Figura 6.4, pode-se observar que ventos de leste são predominantes na região. Isto se deve ao fato da usina eólica em questão estar localizada no litoral do Ceará, onde os ventos são caracterizados pelos ventos alísios e brisas constantes. Nota-se ainda que os ventos apresentam maior frequência de intensidade num intervalo entre 6 m/s e 13 m/s, confirmando assim o ótimo potencial eólico da região.

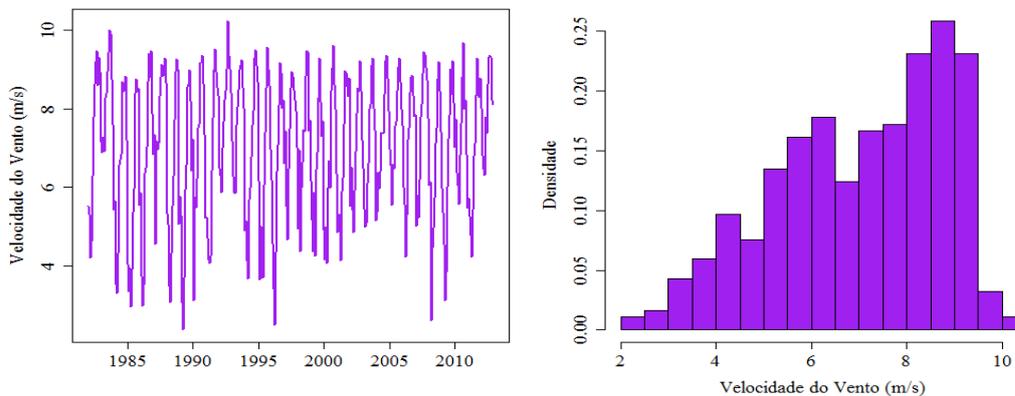


Figura 6.5: Série mensal e histograma da velocidade do vento na usina eólica no período de 01/1982 a 12/2012.

Na Figura 6.5, é apresentada a evolução da série de velocidade do vento da usina eólica, onde é possível notar um comportamento sazonal da mesma. Observa-se que o nível da série não apresentam grande variação ao longo do tempo, sugerindo uma estacionariedade da série, fato este também abordado por Amaral(2011).

A Figura 6.6 apresenta os boxplots da série mensal da velocidade do vento ao longo dos 31 anos, com os quais é possível observar a distribuição da velocidade em cada um dos meses.

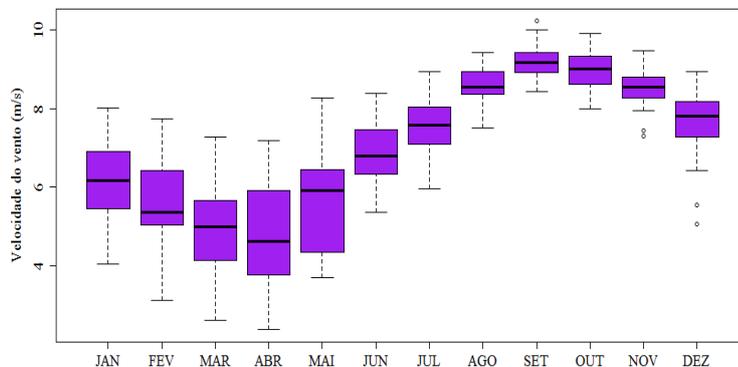


Figura 6.6: Box Plot das médias horárias por mês da velocidade do vento para a usina eólica entre 01/1982 e 12/2012.

Através da Figura 6.6, observa-se a presença de alguns *outliers*¹ nos meses de setembro, novembro e dezembro. Pode-se observar a maior dispersão dos dados de velocidade do vento nos primeiros meses do ano, de janeiro a maio, enquanto o segundo semestre apresenta dados menos dispersos.

Foi realizado o teste de Dickey Fuller nas séries de velocidade do vento a fim de se testar a estacionariedade destas e o mesmo forneceu um p-valor menor que 0,01. Assim, ao nível de 5% significância há evidências de que a série é estacionária, confirmando a observação feita anteriormente.

A fim de investigar a periodicidade desta foi estimado o periodograma, procedimento muito utilizado em séries meteorológicas para a determinação do período sazonal de uma série. No periodograma a série é decomposta em várias frequências senoidais, possibilitando a interpretação de uma série temporal no domínio da frequência, segundo Morettin & Toloí (2006). Para o cálculo do periodograma, utilizou-se as 372 observações disponíveis, isto é, período de 01/1982 a 12/2012. Pela Figura 6.7 pode-se notar que o periodograma apresenta um pico na frequência 31, caracterizando 31 ciclos em 372 observações, indicando assim a periodicidade de 12 meses da série da velocidade do vento.

¹Os valores atípicos ou *outliers* de um conjunto de dados no gráfico Boxplot são classificados através dos seguintes limites: $x < Q_1 - 1,5 IQR$ ou $x > Q_3 + 1,5 IQR$, onde Q_1 é o 1º quartil, Q_3 o 3º quartil e IQR é o intervalo interquartil dado por $Q_3 - Q_1$.

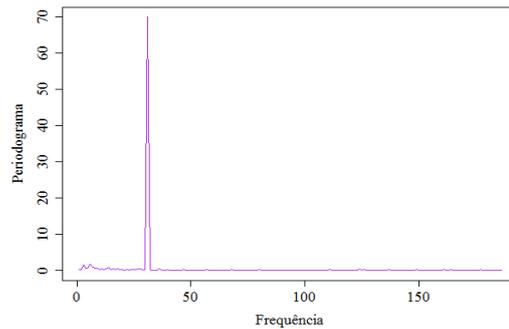


Figura 6.7: Periodograma da série mensal da velocidade do vento entre 01/1982 e 12/2012.

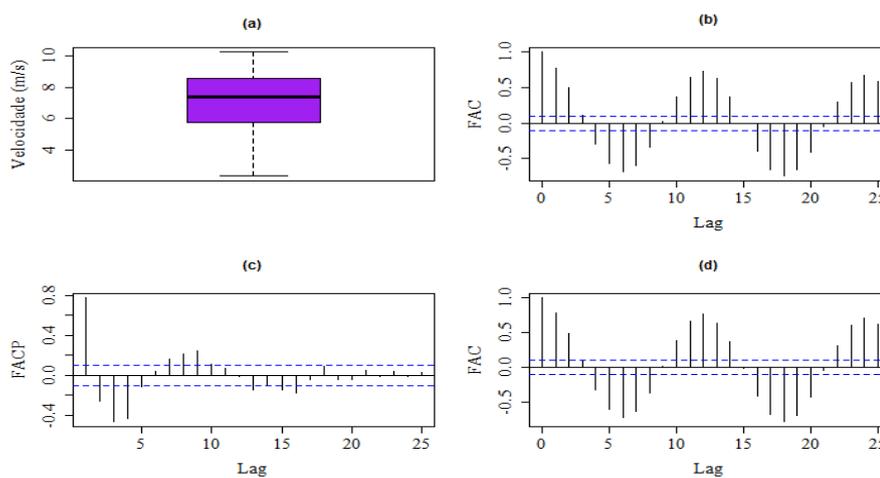


Figura 6.8: (a) Box Plot das médias mensais de velocidade do vento; (b) Função de Autocorrelação; (c) Função de Autocorrelação Parcial; (d) Função de autocorrelação da série ao quadrado.

Analisando a Figura 6.8, é possível observar que a série da velocidade do vento mensal não possui *outliers*. É possível notar que a velocidade do vento possui comportamento típico de séries temporais com componente sazonal devido ao comportamento senoidal da FAC como abordado por Morettin (2006) e comentado anteriormente na interpretação da Figura 6.7. Nota-se ainda o comportamento da FAC da velocidade ao quadrado que evidencia uma dependência temporal bem pronunciada.

6.1.2

Análise descritiva dos dados de El Niño

Para mensurar o efeito do fenômeno ENOS na velocidade do vento, foi utilizada como variável exógena dos modelos propostos nessa dissertação a série de TSM da região Niño 3.4, apresentada no Capítulo 3. Esta região fornece boas medidas de mudanças importantes na TSM, sendo comumente utilizada para a classificação dos fenômenos El Niño e La Niña em índices elaborados pelo NOAA (NOAA, 2014).

A análise descritiva da série de TSM sugere que a mesma (Figura 6.9) é aparentemente estacionária. Para investigar tal hipótese foi realizado o teste da raiz unitária de Dickey-Fuller e o mesmo forneceu um p-valor de 0,01. Sendo assim, há evidências de que a série de TSM é estacionária.

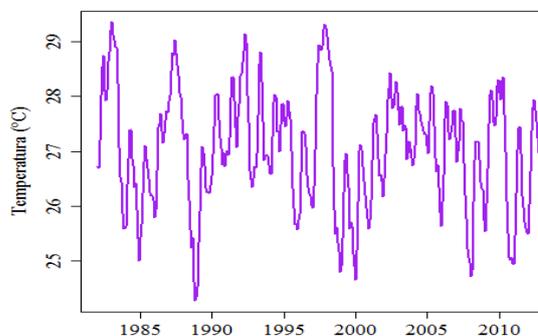


Figura 6.9: Série mensal de Temperatura da Superfície do Mar na região Niño 3.4.

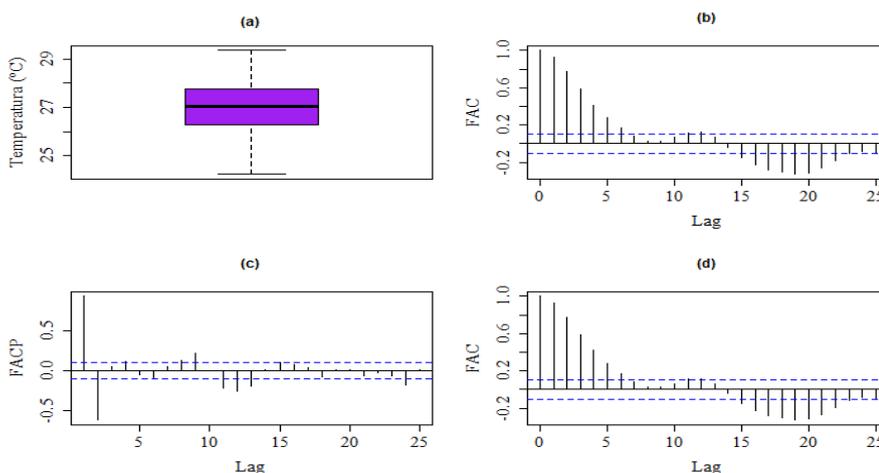


Figura 6.10: (a) Box Plot das médias mensais de TSM da região Niño 3.4; (b) Função de Autocorrelação; (c) Função de Autocorrelação Parcial; (d) Função de autocorrelação da série ao quadrado.

Pela Figura 6.10 nota-se que a série de TSM da região Niño 3.4 não possui *outliers*, sua distribuição é aproximadamente simétrica e seu correlograma apresenta um decaimento exponencial nas primeiras defasagens.

6.1.3 El Niño x Vento

A fim de investigar a relação entre a série de vento e a série de TSM da região Niño 3.4, foi realizada a correlação cruzada, a qual mensura a dependência temporal entre séries temporais.

Assim, tem-se que a correlação cruzada entre x_{t+k} (TSM) e y_t (vento) é o valor do *lag* k no gráfico desta (Figura 6.11). Pode-se observar no gráfico que novamente o comportamento senoidal de séries meteorológicas está presente. Além disso, nota-se uma dependência temporal significativa entre estas séries, com maiores valores nas defasagens de ordem 4 e 5.

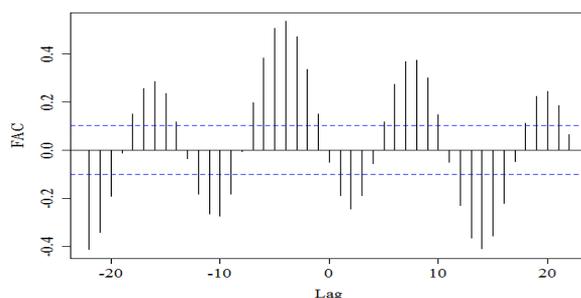


Figura 6.11: Correlação cruzada entre velocidade do vento e TSM na região Niño 3.4.

A partir desta inspeção optou-se por utilizar nos primeiros modelos apenas as defasagens 4 e 5 da variável exógena (TSM da região Niño 3.4). Contudo, os testes mostraram que a maior dependência é de ordem 4 e, portanto, foram considerados para a análise final apenas os modelos com a inclusão da variável exógena TSM região Niño 3.4 com defasagem 4.

Posteriormente, buscou-se analisar o perfil da velocidade ao longo dos anos através de médias horárias da velocidade do vento condicionada a três fases: El Niño, La Niña e neutro. Para essa classificação considerou-se como ocorrência de El Niño valores de anomalia de TSM da região Niño 3.4 maiores ou iguais a 0,5; ocorrência de La Niña menores ou iguais a -0,5 e a fase neutra como valores contidos no intervalo de -0,5 e 0,5. Essa classificação é similar à utilizada na classificação do fenômeno ENOS através do ONI, descrito no Capítulo 3. Isto foi realizado com o intuito de se observar possíveis diferenças mensais, o que caracterizaria o efeito do fenômeno ENOS no vento.

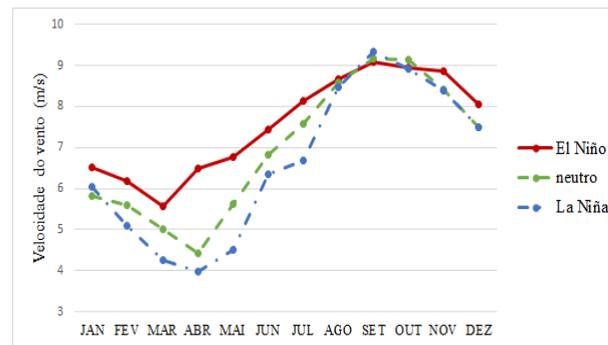


Figura 6.12: Médias horárias da velocidade do vento por mês condicionada a fase de ENOS para o período de 01/1982 a 12/2012.

Analisando a Figura 6.12 é possível notar diferenças no padrão de ocorrências do El Niño e La Niña entre fevereiro e julho, sendo a maior diferença observada no mês de abril, tanto para a fase fria do fenômeno (La Niña), quanto para a fase quente (El Niño), característica também encontrada por Harper *et al.* (2007).

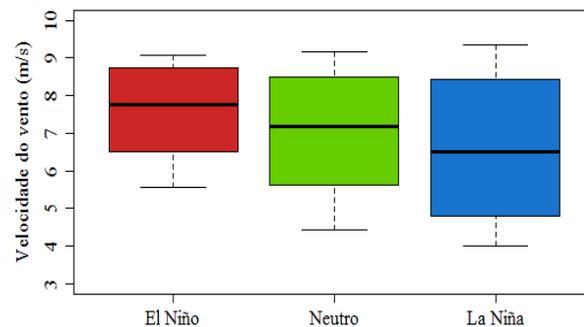


Figura 6.13: Boxplot das médias horárias da velocidade do vento por mês condicionada a fase de ENOS para o período de 01/1982 a 12/2012.

Pela Figura 6.13 observa-se que tanto as fases de El Niño e de La Niña quanto, a fase neutra, não possuem valores extremos e são aproximadamente simétricas. Além disso, nota-se que a fase quente do fenômeno ENOS apresenta valores mais altos em grande parte dos meses, diferentemente da fase fria.

Pelas análises anteriores pode-se notar que o comportamento da velocidade no período de presença do fenômeno ENOS são aparentemente diferentes dos anos neutros em alguns meses. No entanto, como já observado, não espera-se que essas diferenças sejam de grande magnitude, não só pelo resultados já encontrados na literatura (Harper *et al.*, 2007) mas também por causa das características da região de estudo, tal como ventos constantes durante todo o ano. Assim, observa-se que o efeito do ENOS no vento é de difícil mensuração e como forma de explorar mais esse efeito, foi proposta a construção de uma

tabela de contingência, na qual buscou-se associar a ocorrência ou não de anomalias extremas de TSM entre diferentes intensidades de vento.

Primeiramente, foi calculado tercis da velocidade do vento, onde estes foram classificados como sendo: abaixo, quase-normal e acima da média, durante todo o período em estudo. Posteriormente utilizou-se as anomalias de TSM na região Niño 3.4 para a classificação dos eventos de ENOS. Para eventos fortes de El Niño e La Niña definiu-se anomalias maiores ou iguais que 1,2 e menores ou iguais a -1,2, respectivamente. Essa definição foi realizada de forma análoga à classificação do SOI, também utilizado pelo NOAA para monitoramento do ENOS. No entanto, em termos do SOI, valores positivos indicam El Niño, caso contrário La Niña. Além disso, nota-se que nos intervalos definidos anteriormente encontra-se 20 % das ocorrências de El Niño/La Niña, como também foi apresentado por Sansigolo *et al.* (2006).

Assim, os períodos de ocorrência ou não de valores extremos de anomalia de TSM foram associados a estas classificações do vento, como pode ser observado na Tabela 6.1.

Tabela 6.1: Tabela de Contingência ENOS x Vento.

			Anomalia Extrema		Total
			Sim	Não	
Vento	<i>Abaixo</i>	Observado	18	105	123
		Esperado	25,8	97,2	123,0
		Resíduo	-1,5	0,8	
	<i>Quase-Normal</i>	Observado	28	94	122
		Esperado	25,6	96,4	122
		Resíduo	0,5	-0,2	
	<i>Acima</i>	Observado	32	95	127
		Esperado	26,6	100,4	127,0
		Resíduo	1,0	-0,5	
Total		Observado	78	294	372
		Esperado	78	294	372

Analisando a tabela acima, nota-se que em 372 meses dos 31 anos de dados foram observados 78 meses em que ocorreram anomalias extremas de TSM, onde destes 18 meses o vento se apresentou de forma abaixo da média; em 28 meses acima da média, e assim sucessivamente.

Por fim, foi realizado um teste Qui-quadrado, no qual buscou-se testar a existência de dependência entre a ocorrência ou não de valores extremos de anomalia de TSM em diferentes níveis de vento. O teste forneceu um p-valor de 0.098 e foi fixado um nível de significância de 10 % pois diante das

evidências anteriores trata-se de um evento extremo, o que torna razoável o valor fixado. Sendo assim, através desse teste vê-se que há evidências de que a TSM e velocidade do vento possuem associação, isto é, o comportamento do vento é influenciado pela TSM.

Além disso, é possível verificar se esta associação está na direção esperada, isto é, se os valores observados são maiores ou menores que os valores esperados sob a hipótese da ausência de associação. Esta é analisada através dos resíduos padronizados, os quais também são apresentados na Tabela 6.1. Nota-se que em 18 dos 78 meses de ocorrência de anomalias extremas, os ventos estavam abaixo da média. Logo, tem-se que o número de casos observados é inferior ao esperado, o que pode ser analisado pelo resíduo negativo de -1,5. O contrário pode ser observado para ventos acima da média nos períodos de ocorrência de anomalias extremas onde tem-se um resíduo de 1,0, que leva a uma conclusão de que o número de casos observados superaram o número de casos esperados sob hipótese de não associação entre as variáveis.

Vale ressaltar que essa análise mensal não está diretamente relacionada à ocorrência do ENOS em si, uma vez que este é caracterizado somente no comportamento anormal da TSM em pelo menos 5 meses consecutivos (NOAA, 2014). No entanto, o interesse está na observação do comportamento do vento na presença de um extremo de anomalia TSM.

6.2

Modelos Univariados GAS

Para os modelos apresentados no decorrer deste capítulo, tanto no que diz respeito à densidade preditiva (Weibull, Gama e Beta), quanto aos seus respectivos mecanismos de evolução (SARIMA e CNO), serão referenciados como **Modelo 1**, os modelos sem variáveis exógenas e **Modelo 2** como modelos com a variável exógena TSM da região 3.4.

Em relação aos valores iniciais dos parâmetros fixos do modelo (A_i , B_j e α), estes foram escolhidos por tentativa e erro, através da comparação do comportamento dos resíduos quantílicos para diversas combinações de valores dos parâmetros. As estimativas dos parâmetros de todos os modelos e seus respectivos erros-padrão são apresentados no Apêndice A, disponível no final deste documento.

A seguir, serão apresentados os resultados obtidos para os modelos GAS para as distribuições Weibull, Gama e Beta.

6.2.1 Modelos GAS para a distribuição Weibull

Como antecipado no Capítulo 5, os modelos realizados com a distribuição de Weibull não foram satisfatórios. Os principais resultados são apresentados nessa seção.

Na Figura 6.14, é apresentado o ajuste do modelo Weibull-SARIMA (linha vermelha pontilhada), à série de velocidade do vento (linha roxa sólida) para os 31 anos de observações e para os últimos 10 anos.

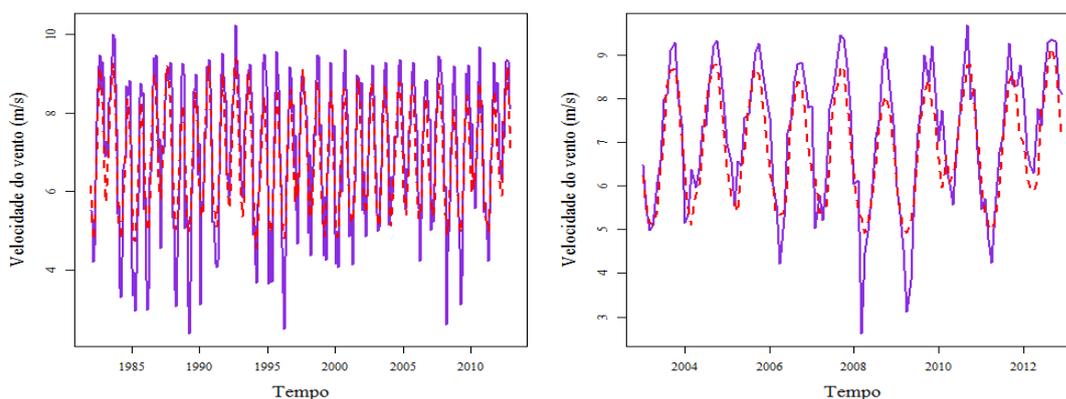


Figura 6.14: Ajuste da série de vento para período de estimação do Modelo 1 Weibull-SARIMA.

Pela Figura 6.14, pode-se observar que o Modelo Weibull-SARIMA aparentemente capturou de uma forma geral a dinâmica da série. No entanto, nota-se a presença bem pronunciada de deslocamentos do ajuste em relação à série, principalmente para valores baixos desta, os quais ocorrem, em geral, no início de cada ano.

Na Tabela 6.2 são apresentados as medidas de ajuste do modelo Weibull-SARIMA sem variáveis exógenas, onde se pode ter conhecimento da magnitude destas medidas.

Tabela 6.2: Medidas de ajuste do modelo Weibull-SARIMA.

	RMSE	MAE	MASE	SMAPE	Pseudo - R²
Modelo 1	0.86	0.67	0.71	5.51	0.80

Para esse modelo, utilizou-se o *score* ponderado pela matriz pseudo-inversa da matriz de informação de Fisher, que em geral, na literatura dos modelos GAS fornece melhores resultados.

Os testes de Ljung-Box e Jarque Bera, os quais foram utilizados para testar a presença de autocorrelação serial, heterocedasticidade condicional e normalidade nos resíduos quantílicos do modelo Weibull-SARIMA, são apresentados na Tabela 6.3.

Tabela 6.3: P-valor dos testes para os resíduos quantílicos do modelo Weibull-SARIMA.

	Autocorrelação	Heterocedasticidade	Normalidade
Modelo 1	0.00	0.00	0.00

Nota-se pela Tabela 6.3 que para todos os testes realizados, as hipóteses nulas dos mesmos devem ser rejeitadas, evidenciando assim que os resíduos quantílicos apresentam autocorrelação serial, heterocedasticidade condicional e não-normalidade. Estas características também podem ser observadas nas Figuras 6.15 e 6.16, nas quais pode-se concluir que a dependência linear e não linear, não foram capturadas satisfatoriamente, uma vez que nos correlogramas o comportamento senoidal desses resíduos é bem evidente.

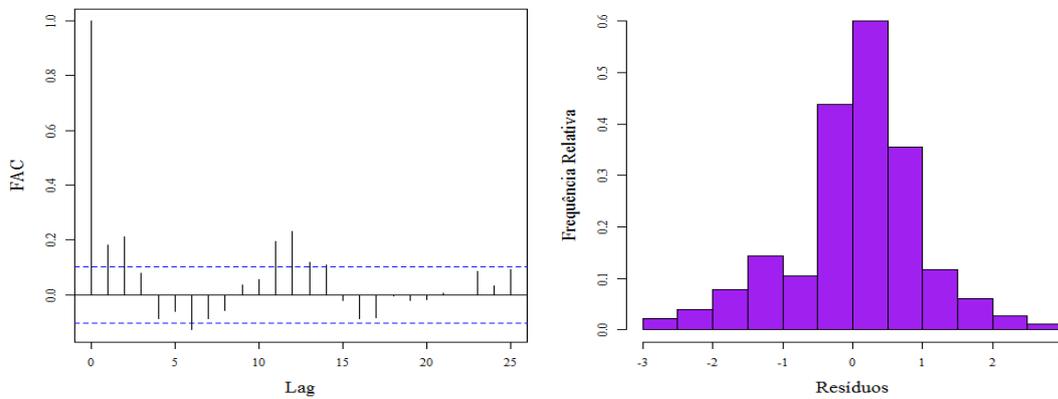


Figura 6.15: FAC e histograma dos resíduos quantílicos do Modelo 1 - Weibull-SARIMA.

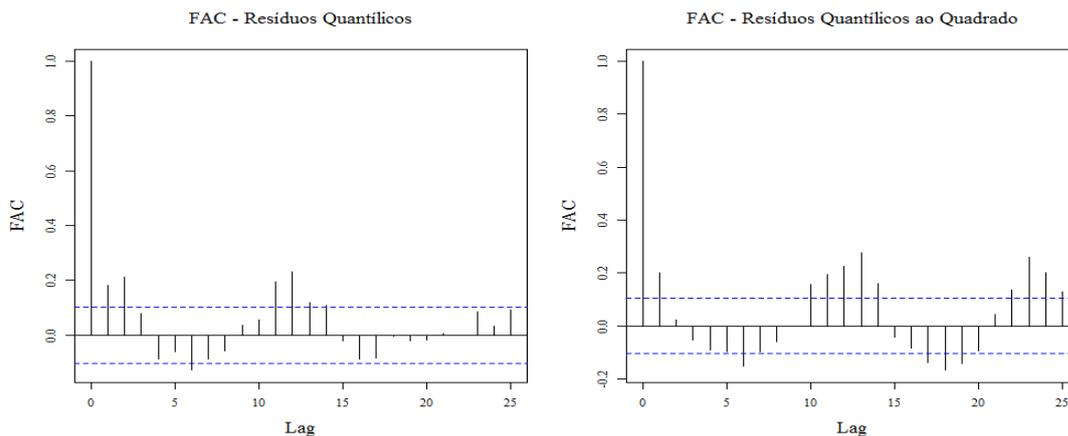


Figura 6.16: FAC dos resíduos quantílicos do Modelo 1 - Weibull-SARIMA.

Logo, ainda que a distribuição de Weibull seja a mais utilizada na literatura para descrever o comportamento da velocidade do vento, esta não se ajustou de forma satisfatória às séries de vento de frequências mensais.

6.2.2

Modelos GAS para a distribuição Gama

Nessa seção, são apresentados os resultados para a modelagem das séries de vento para uma usina eólica de interesse para os modelos univariados Gama, apresentados no Capítulo 5.

Modelos Gama com estrutura de evolução SARIMA

A configuração dos modelos Gama trata apenas o parâmetro de escala λ como variante no tempo, sendo este definido pela equação de evolução 5-11, na página 40. Em primeiro momento, o modelo foi aplicado somente à série de vento, sem a inclusão de variáveis exógenas. Posteriormente, a variável exógena TSM da região Niño 3.4 com defasagem 4 foi incluída no modelo com o objetivo de quantificar o efeito do El Niño na série de velocidade do vento. Vale ressaltar que a inclusão desta variável foi significativa no modelo, corroborando para a hipótese de existência do efeito da mesma na velocidade dos ventos. Os resultados para esses modelos são apresentados a seguir.

Na Figura 6.17 é apresentado o ajuste do modelo (linha vermelha pontilhada) à série de velocidade do vento (linha roxa sólida), onde é possível notar que o modelo consegue capturar grande parte da dinâmica da série apenas por informações passadas da mesma, isto é, apresenta uma boa capacidade preditiva. Entretanto, nota-se alguns descolamentos do ajuste à série de velocidade do vento, já observados anteriormente no modelo Weibull.

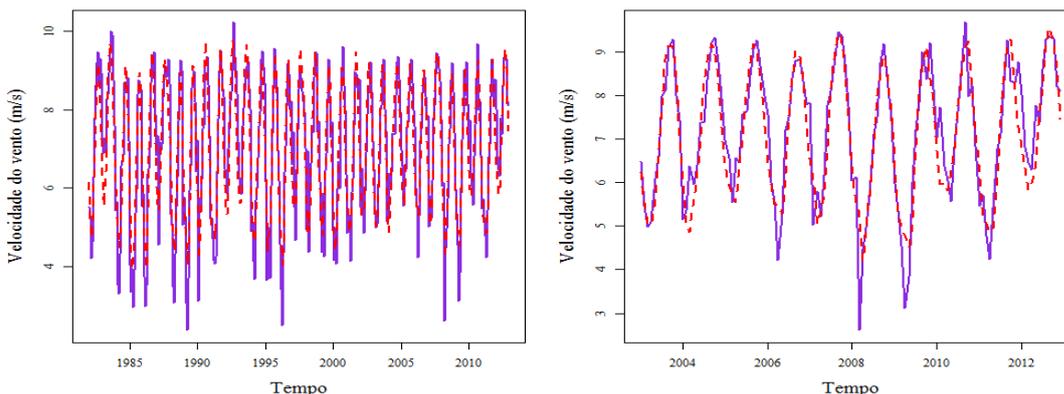


Figura 6.17: Ajuste da série de vento para período de estimação do Modelo 1 - Gama-SARIMA.

Na Tabela 6.4 são apresentados os critérios de adequação dos modelos Gama-SARIMA.

Tabela 6.4: Critérios de adequação dos modelos Gama-SARIMA.

	AIC	BIC	Log-likelihood
Modelo 1	1012.402	1059.035	-494.2009
Modelo 2	1009.60	1060.12	-491.80

Através da Tabela 6.4 nota-se que o modelo 2 possui características um pouco superiores aos do modelo 1, como menor AIC e maior log-verossimilhança. Isto indica que, de uma forma geral, a inclusão da variável exógena no modelo Gama forneceu melhorias no desempenho do mesmo. No entanto, esta melhoria é aparente, uma vez que esta não se traduz em um aumento do poder preditivo do modelo, o que pode ser observado na Tabela 6.5, na qual são apresentadas as medidas de ajuste dos modelos Gama-SARIMA.

Tabela 6.5: Medidas de ajuste dos modelos Gama-SARIMA.

	RMSE	MAE	MASE	SMAPE	Pseudo - R²
Modelo 1	0.78	0.58	0.61	4.87	0.81
Modelo 2	0.80	0.63	0.66	5.06	0.80

Na Tabela 6.5 pode-se observar que o modelo 2 apresentou medidas de ajuste inferiores às do modelo 1. Entretanto, este fato pode ser justificado pela nova combinação de valores gerados a cada estimação dos parâmetros dos modelos, isto é, para cada especificação do modelo os parâmetros são reestimados, o que pode levar a uma variação dos valores dos mesmos.

Como o objetivo central dessa dissertação é a estimação do efeito da variável exógena TSM na velocidade dos ventos, a significância desta variável no modelo já será suficiente para esta análise, ainda que isso não se traduza em um ganho de poder preditivo ao modelo.

Os gráficos de ajuste do modelo 2 às séries de velocidade foram omitidos, uma vez que este forneceu resultados semelhantes ao modelo 1.

A análise dos resíduos quantílicos mostrou que a dependência linear da série foi capturada, o que pode ser concluído através do p-valor fornecido pelo teste de Ljung-Box (Tabela 6.6) e observado na Figura 6.18, na qual todos os lags da FAC não são significativos, possuindo um comportamento desejável, sendo similar a um processo ruído branco. No entanto, estes resíduos não apresentam distribuição normal, segundo o teste Jarque-Bera (Tabela 6.6), o que também pode ser observado no histograma dos resíduos, onde os mesmos apresentam aparentemente uma assimetria negativa.

Sendo assim, nota-se através dos testes apresentados na Tabela 6.6, que os resíduos não apresentam autocorrelação serial e normalidade, entretanto, possuem heterocedasticidade condicional (Figura 6.19).

Tabela 6.6: P-valor dos testes para os resíduos quantílicos dos modelos Gama-SARIMA.

	Autocorrelação	Heterocedasticidade	Normalidade
Modelo 1	0.76	0.00	0.00
Modelo 2	0.40	0.00	0.00

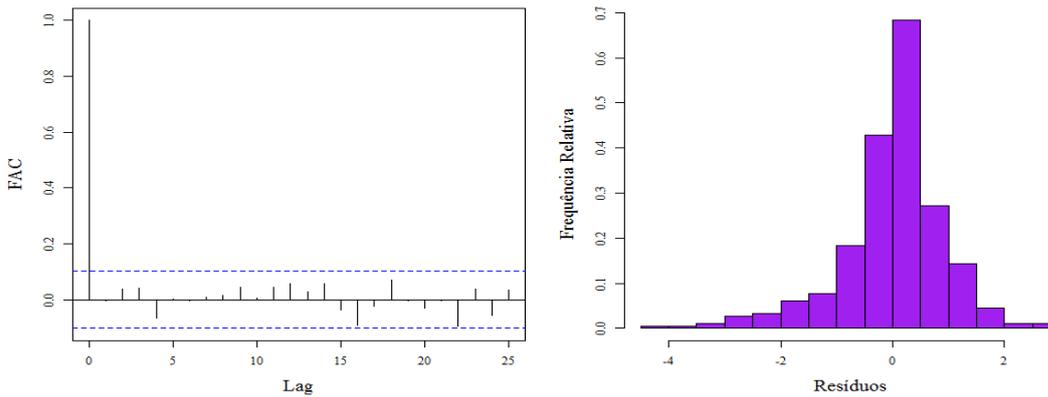


Figura 6.18: FAC e Histograma dos resíduos quantílicos do Modelo 1 - Gama-SARIMA.

Na Figura 6.19, é apresentada a FAC dos resíduos ao quadrado, onde é possível notar uma forte estrutura senoidal, comportamento este esperado pela natureza dessa série, sugerindo que a variância condicional não é explicada satisfatoriamente pelo modelo adotado.

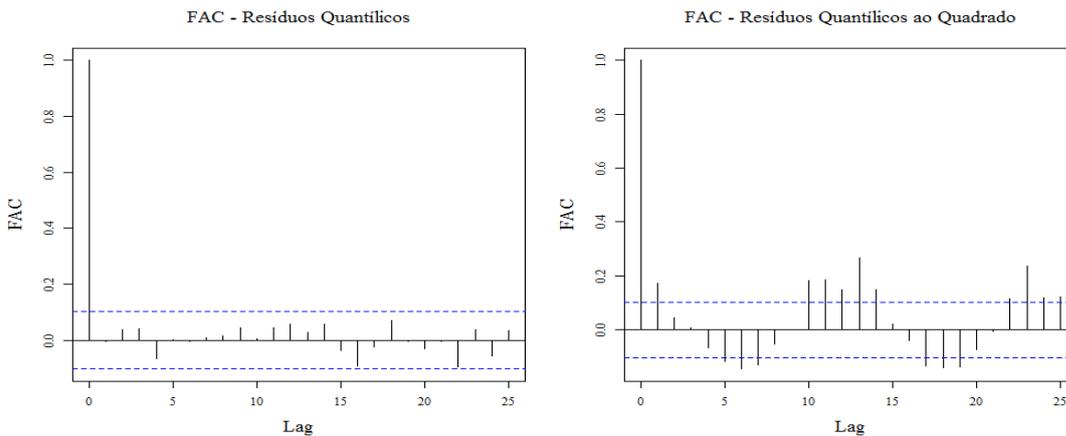


Figura 6.19: FAC dos resíduos do Modelo 1 - Gama-SARIMA.

Os resíduos quantílicos do modelo 2 apresentaram as mesmas características dos resíduos do modelo 1, como observado na Tabela 6.6 e por esse motivo os gráficos da análise dos mesmos foram omitidos.

Para a variável exógena TSM, o modelo forneceu uma estimativa para o parâmetro ϕ igual a 0.019, com um erro padrão de 0.005, o que resultou em uma estatística de teste igual a 3.81, sendo esta variável significativa no modelo ao nível de 5%. Assim, o efeito da TSM da região Niño 3.4 na velocidade do vento, obtido a partir Eq. (5-19) considerando-se x_{t-4} ao invés de x_t e, supondo $\Delta f_t = 0$ e $\Delta x_{t-4} = 1^\circ\text{C}$ é de 1.92%. Já para um decréscimo de 1°C na TSM entre os meses, isto é, $\Delta x_{t-4} = -1^\circ\text{C}$ o que seria caracterizado pelo La Niña o efeito deste na velocidade do vento seria de -1.88% , sendo este bem próximo da magnitude do efeito anterior, o qual caracteriza o El Niño.

Devido ao comportamento senoidal observado nos resíduos ao quadrado dos modelos Gama, foi proposta uma nova configuração do mesmo onde os dois parâmetros da distribuição, α e β , variam no tempo a fim de procurar reduzir a forte dependência não-linear presente nos resíduos. Entretanto, pelos testes realizados com esse modelo notou-se que este não forneceu resultados superiores ao modelo Gama-SARIMA com o parâmetro de escala variante no tempo, no que diz respeito a melhorias do ajuste do modelo. Caso o leitor tenha interesse sobre tal abordagem, as principais manipulações dessa configuração do modelo Gama, bem como sua ilustração são apresentadas no Apêndice A.

Modelos Gama com estrutura de evolução CNO

A especificação do modelo GAS com evolução por componentes não-observáveis foi definida nas equações na Subseção 5.5.2 do capítulo 5. A seguir são observados os principais resultados do modelo Gama-CNO.

A principal observação realizada nos modelos Gama-CNO é que a estimação do modelo Gama por componentes que buscam capturar a sazonalidade e a parte autoregressiva da série fornece uma melhoria para o ajuste da série de velocidade do vento, o que pode ser observado nas tabelas 6.7 e 6.8 e na Figura 6.20.

Tabela 6.7: Critérios de adequação dos modelos Gama-CNO.

	AIC	BIC	Log-likelihood
Modelo 1	965.47	984.90	-477.73
Modelo 2	957.47	980.79	-472.74

Tabela 6.8: Medidas de ajuste dos modelos Gama-CNO.

	RMSE	MAE	MASE	SMAPE	Pseudo - R ²
Modelo 1	0.74	0.56	0.59	4.64	0.82
Modelo 2	0.74	0.57	0.59	4.65	0.83

Ao se fazer uma comparação entre a Tabela 6.5, provenientes dos resultados dos modelos Gama-SARIMA, e a Tabela 6.8, pode-se notar que os modelos Gama por componentes não-observáveis apresentaram menores valores para a log-verossimilhança, bem como os critérios AIC e BIC do que os modelos com evolução SARIMA. Além disso, a redução das métricas RMSE, MAE, MASE, SMAPE, de ajuste do modelo, mostram que essa configuração do modelo Gama apresentou melhorias no que diz respeito a captura da dinâmica da série. Além disso, observa-se que o modelo 2 com a variável exógena TSM da região de Niño 3.4 apresentou melhores critérios de seleção dos modelos e um pequeno aumento do R² de 0.82 para 0.83, sendo este maior que o encontrado nos modelos Gama-SARIMA.

Para a comparação da precisão dos modelos Gama em relação às previsões *in-sample*, foi realizado o teste Diebold-Mariano unilateral com H_1 : “O Modelo Gama-SARIMA possui menor grau de aderência que o Modelo Gama-CNO”, apresentado em sua forma bilateral na Seção 5.8, onde este forneceu um p-valor de 0.016 e estatística de teste de -2,14. Assim, através do teste, ao nível de significância de 5%, a hipótese nula deve ser rejeitada e, portanto, o modelo Gama-CNO apresenta maior precisão do que o modelo Gama-SARIMA. Tal resultado concorda com as observações feitas anteriormente para o modelo Gama-CNO.

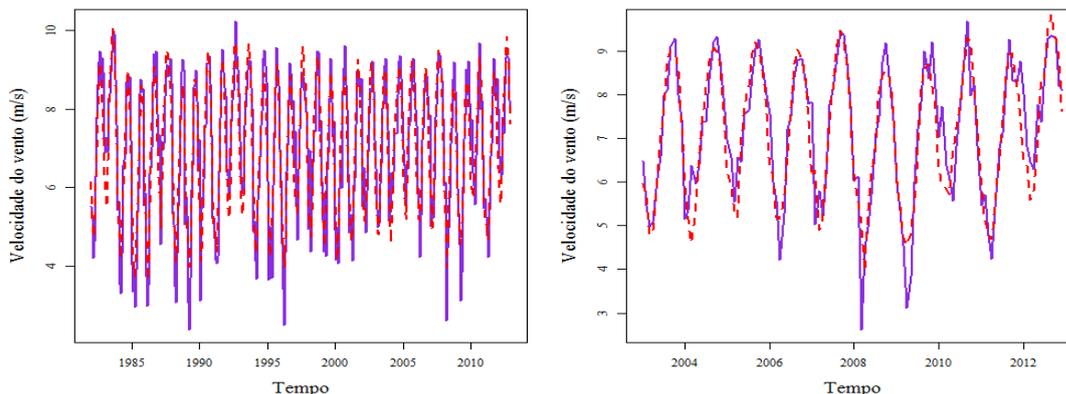


Figura 6.20: Ajuste da série de vento para período de estimação do Modelo 1 - Gama-CNO com o parâmetro de escala variante em t .

O modelo Gama-CNO possui maior interpretabilidade que o modelo Gama-SARIMA, uma vez que suas componentes podem ser interpretadas separadamente. Na Figura 6.21, apresenta-se a componente autoregressiva para todo o período de estudo e a sazonalidade para os últimos dez anos da série de velocidade do vento.

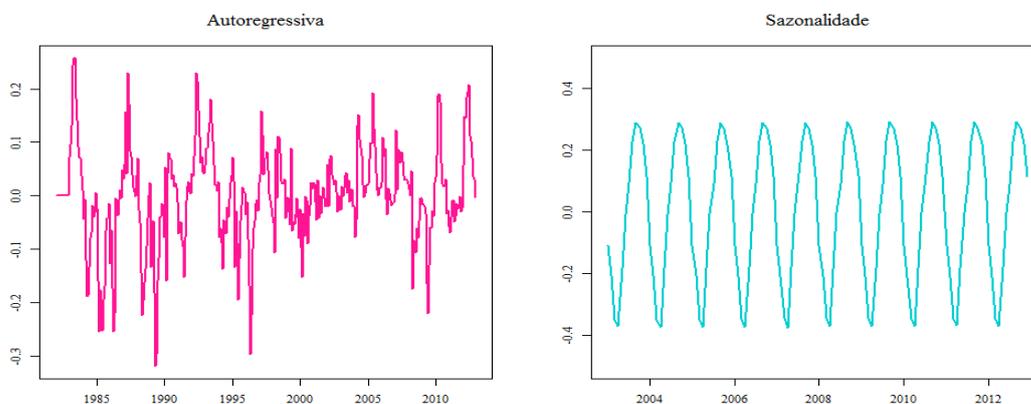


Figura 6.21: Componente sazonal e autoregressiva do ajuste da série de vento para período de estimação do Modelo 1 - Gama-CNO.

Os resíduos quantílicos dos modelos Gama-CNO se comportam de forma semelhante aos dos modelos Gama-SARIMA, apresentando heterocedasticidade condicional e não-normalidade, como pode ser visto na Tabela 6.9 e Figura 6.22.

Tabela 6.9: P-valor dos testes para os resíduos dos modelos Gama-CNO.

	Autocorrelação	Heterocedasticidade	Normalidade
Modelo 1	0.18	0.00	0.00
Modelo 2	0.26	0.00	0.00

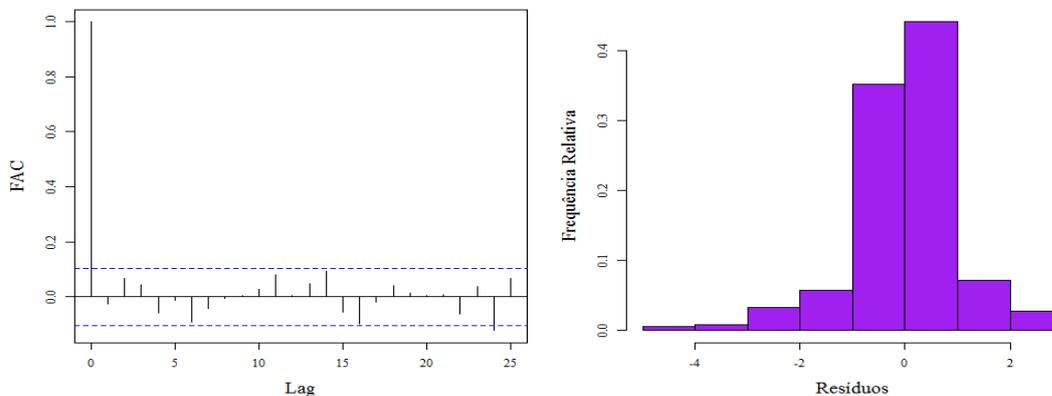


Figura 6.22: FAC e Histograma dos resíduos quantílicos do Modelo 1 - Gama-CNO.

Para o modelos Gama-CNO, a estimativa do parâmetro ϕ relacionado à variável exógena TSM da região Niño 3.4 foi igual a 0.039 e seu erro padrão 0.011, o que resultou em uma estatística t igual a 3.26, sendo esta variável significativa no modelo ao nível de 5%. Assim, o efeito da TSM da região Niño 3.4 na velocidade do vento, obtido a partir da Eq. (5-19), com as mesmas suposições para o modelo Gama-SARIMA, é de 3.98 %, ou seja, para um aumento de 1°C na TSM entre meses consecutivos, tem-se que o efeito desta na velocidade do vento será de 3.98 %. Já para um decréscimo de 1°C na TSM entre os meses, isto é, $\Delta x_{t-4} = -1^\circ\text{C}$ o que seria caracterizado pelo La Niña o efeito deste na velocidade do vento seria de -3.82% , sendo este bem próximo da magnitude do efeito anterior, o qual caracteriza o El Niño.

Por fim, vale ressaltar que foram realizadas intervenções do tipo *dummy* em observações mal-ajustadas pelos modelos Gama. No entanto, devido a evidente estrutura de assimetria dos resíduos destes modelos, estas não foram suficientes para corrigir a heterocedasticidade condicional e a não-normalidade presente nos mesmos.

6.2.3

Modelos GAS para a distribuição Beta

Yilmaz & Çelik (2008) e Azrulhisham *et al.* (2012) utilizam a distribuição Beta para a modelagem de séries de velocidade do vento. A motivação para a utilização dessa distribuição é devido ao comportamento assimétrico dos dados, fato este observado na Figura 6.5 no início desse capítulo.

Os modelos Beta foram descritos na seção 5.6. Nestes é necessário fixar um valor para k , relacionado à nova parametrização da distribuição Beta. Para isso foram realizados vários testes em busca do melhor k a ser utilizado. Vale ressaltar que os k 's testados foram fixados, considerando que as médias diárias de vento na série em estudo não ultrapassassem os valores dos mesmos. Os resultados mais satisfatórios foram obtidos utilizando $k = 11$, ou seja, velocidade máxima 11 m/s.

Os resultados para os modelos Beta são apresentados no decorrer dessa seção.

Modelos Beta com estrutura de evolução SARIMA

Na Figura 6.23, é apresentado o ajuste do modelo 1 (linha vermelha pontilhada) à série de velocidade do vento (linha roxa sólida). Observa-se que o mesmo consegue capturar de forma geral a dinâmica da série, bem como os modelos Weibull e Gama.

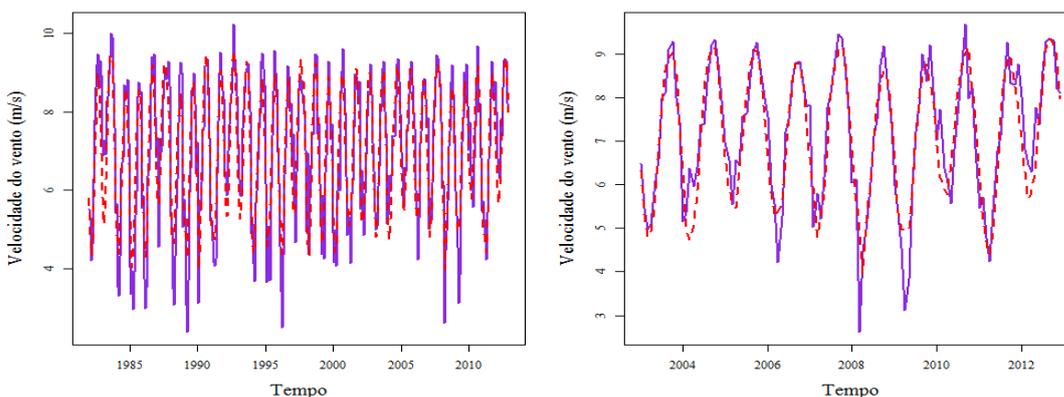


Figura 6.23: Ajuste da série de vento para período de estimação do Modelo 1 - Beta-SARIMA.

Nas Tabelas 6.10 e 6.11 são apresentados os critérios de adequação do modelo e medidas de ajuste dos mesmos para os modelos 1 e 2. Nestas é possível observar que os modelos Beta possuem melhores propriedades que os modelos Gama, uma vez que apresentam bem menores de AIC, BIC, maior verossimilhança e melhores medidas de ajuste.

Tabela 6.10: Critérios de adequação dos modelos Beta-SARIMA.

	AIC	BIC	Log-likelihood
Modelo 1	769.20	815.83	-372.60
Modelo 2	743.09	793.61	-358.54

Tabela 6.11: Medidas de ajuste dos modelos Beta-SARIMA.

	RMSE	MAE	MASE	SMAPE	Pseudo - R²
Modelo 1	0.76	0.57	0.60	4.83	0.82
Modelo 2	0.74	0.56	0.58	4.65	0.83

Em relação ao modelo 2, o qual inclui a variável exógena TSM, esta forneceu melhorias significativas ao modelo. Isto é concluído pela menor verossimilhança que o mesmo apresenta, bem como menores critérios AIC e BIC e melhores medidas de ajuste. Na Figura 6.24, é apresentado o ajuste do modelo 2, o qual considera a inclusão da variável TSM Niño 3.4, à série de velocidade do vento.

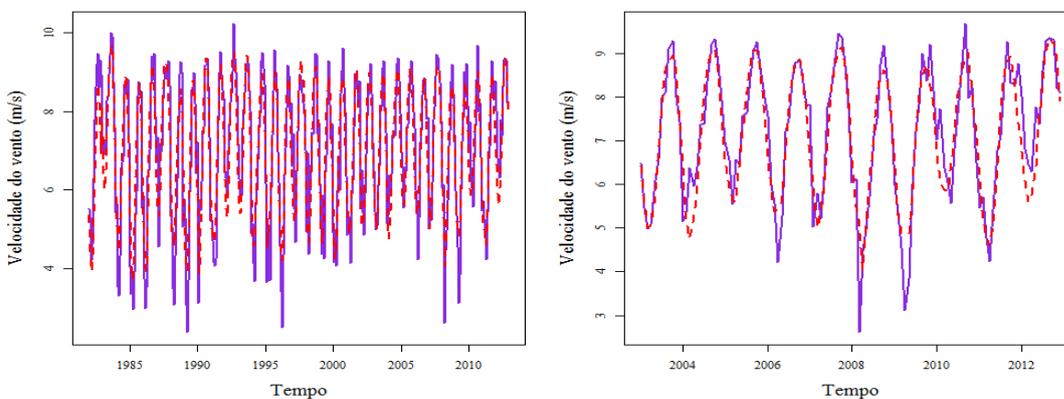


Figura 6.24: Ajuste da série de vento para período de estimação do Modelo 2 - Beta-SARIMA.

Como decorrência das observações feitas anteriormente, nota-se o bom desempenho dos modelos Beta através dos resíduos quantílicos dos mesmos. Estes apresentam um comportamento desejável, uma vez que não apresentam autocorrelação serial, são homocedásticos e normais, segundo os testes realizados e apresentados na Tabela 6.12.

Tabela 6.12: P-valor dos testes para os resíduos dos modelos Beta-SARIMA.

	Autocorrelação	Heterocedasticidade	Normalidade
Modelo 1	0.13	0.90	0.33
Modelo 2	0.31	0.35	0.25

Portanto, nota-se que o p-valor dos testes nos resíduos de ambos os modelos é maior que 5%, sugerindo que as hipóteses nulas não são rejeitadas e, assim os pressupostos de independência, homocedasticidade e normalidade dos resíduos são mantidos. Tais características podem ser observadas na Figura 6.25.

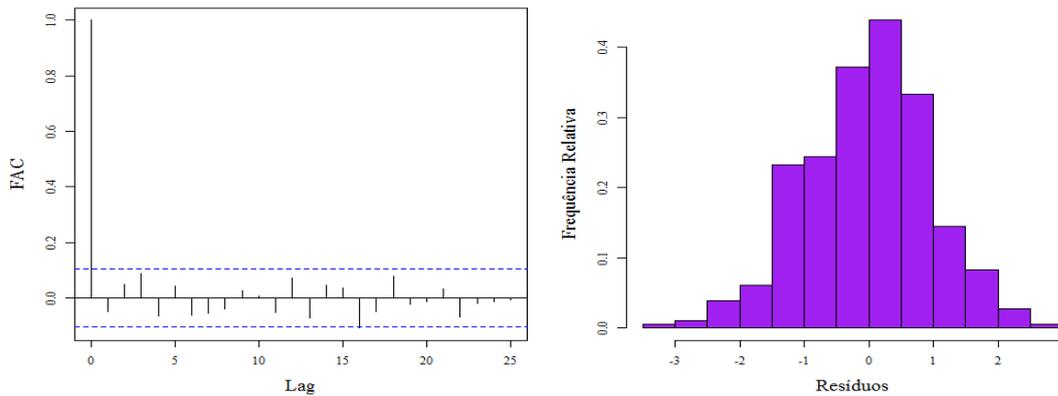


Figura 6.25: FAC e Histograma dos resíduos quantílicos do Modelo 2 - Beta-SARIMA.

Na Figura 6.26 são apresentados ambos os resíduos onde se pode notar que a forte dependência não-linear nos modelos Gama não está presente nos modelos Beta-SARIMA, corroborando para a correta especificação deste modelo.

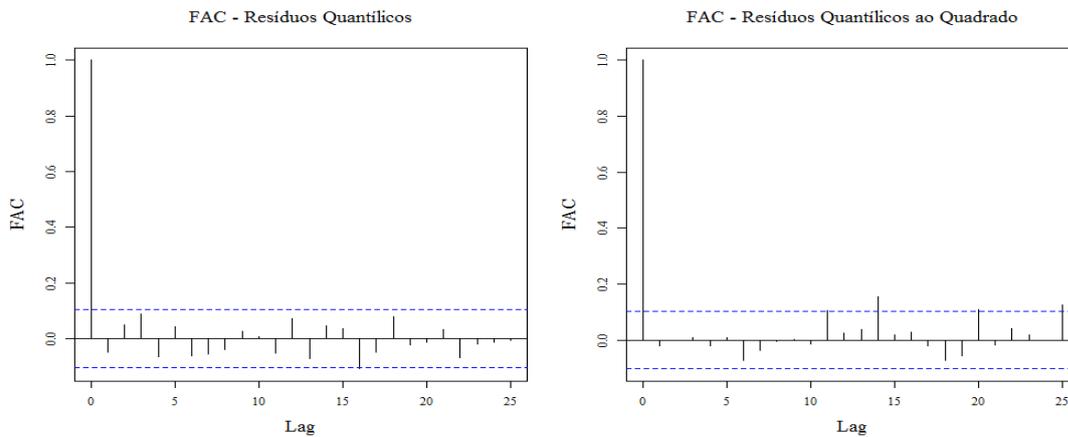


Figura 6.26: FAC dos resíduos quantílicos do Modelo 2 - Beta-SARIMA.

Para o modelo Beta-SARIMA, a estimativa do parâmetro ϕ relacionado à variável exógena TSM da região Niño 3.4 foi igual a 0.047 e seu erro padrão 0.012, o que resulta em uma estatística t igual a 3.92, sendo esta variável significativa no modelo ao nível de 5%. Assim, o efeito da TSM da região Niño 3.4 na velocidade do vento, obtido a partir da Eq. (5-21) para o modelo Beta-SARIMA é apresentado na Figura 6.27 para diferentes valores de variação mensal de TSM.

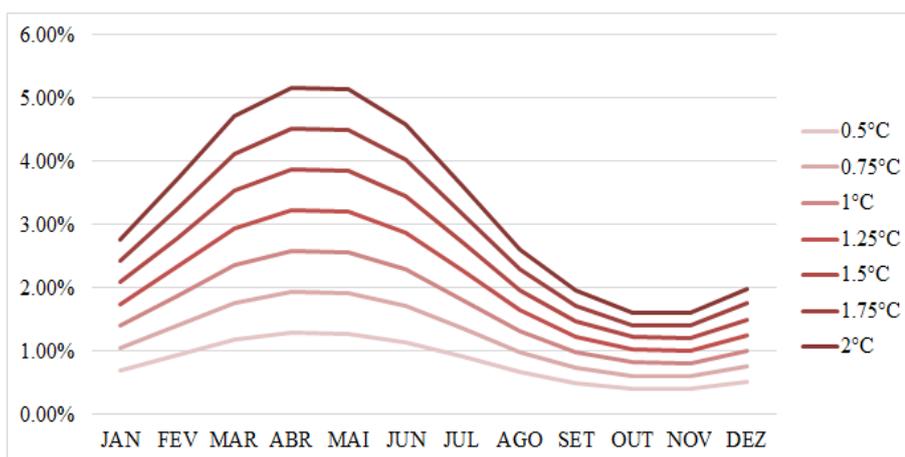


Figura 6.27: Efeito da variável exógena TSM na velocidade do vento para diferentes variações mensais de temperatura - Modelo Beta-SARIMA.

Pela Figura 6.27 é possível observar que à medida que o valor de variação de temperatura cresce, o efeito da variável exógena TSM na intensidade dos ventos também aumenta. Além disso, os maiores efeitos são observados nos meses de março a maio, enquanto os menores, de setembro a novembro, sendo o maior efeito observado no mês de abril (2.58%), fato também notado anteriormente na análise descritiva da relação entre o El Niño e a velocidade do vento (Figura 6.12).

Em relação a decréscimos significativos da TSM entre os meses, o que caracteriza o La Niña, a curva de efeitos é similar à do El Niño em termos de magnitude. No entanto, esta apresenta efeitos negativos como no caso da Gama, apresentado anteriormente. Assim, para um decréscimo de 1°C da TSM entre os meses, isto é, $\Delta x_{t-4} = -1^\circ\text{C}$ o que seria caracterizado pelo La Niña o efeito deste na velocidade do vento seria de -2.57% , sendo este bem próximo da magnitude do efeito anterior, o qual caracteriza o El Niño.

Vale ressaltar que a estimação do efeito da variável de temperatura da superfície do mar no Oceano Pacífico Equatorial foi baseada na estimação do coeficiente relacionada a mesma na defasagem 4, o que justifica o Δx_{t-4} .

Modelos Beta com estrutura de evolução CNO

Assim como nos modelos Gama, foi ajustado à série de velocidade do vento o modelo Beta por componentes não-observáveis. Nas Tabelas 6.13 e 6.14 são apresentados os critérios de adequação e medidas de ajuste dos modelos Beta-CNO, respectivamente.

Tabela 6.13: Critérios de adequação dos modelos Beta-CNO.

	AIC	BIC	Log-likelihood
Modelo 1	724.05	743.48	-357.03
Modelo 2	719.22	742.54	-353.61

Tabela 6.14: Medidas de ajuste dos modelos Beta-CNO.

	RMSE	MAE	MASE	SMAPE	Pseudo - R^2
Modelo 1	0.73	0.54	0.57	4.57	0.83
Modelo 2	0.72	0.54	0.57	4.52	0.84

Nota-se primeiramente que os modelos Beta-CNO, assim como os modelos Gama-CNO, apresentam melhores critérios de adequação e medidas de ajuste quando comparado ao modelo Beta-SARIMA, ainda que estas medidas não sejam tão superiores aos modelos com estrutura de evolução SARIMA.

Na Figura 6.28 é apresentado o ajuste (linha vermelha pontilhada) do modelo Beta-CNO à série de velocidade do vento (linha roxa sólida).

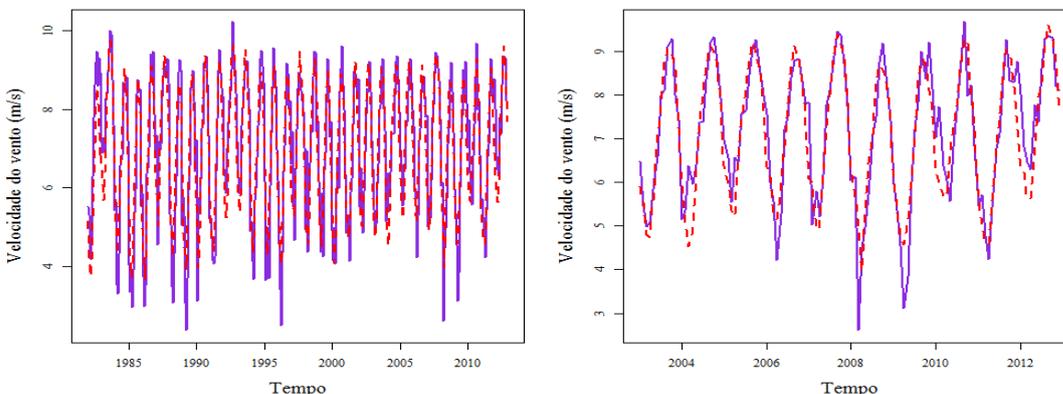


Figura 6.28: Ajuste da série de vento para período de estimação do Modelo 1 - Beta-CNO.

Através da análise dos resíduos nota-se que estes possuem propriedades inferiores aos dos modelos da seção anterior, uma vez que a hipótese de normalidade é rejeitada no modelo 1 e os p-valores do teste Ljung Box para a heterocedasticidade condicional são inferiores aos dos modelos Beta-SARIMA (Tabela 6.15).

Tabela 6.15: P-valor dos testes para os resíduos dos modelos Beta-CNO.

	Autocorrelação	Heterocedasticidade	Normalidade
Modelo 1	0.29	0.06	0.00
Modelo 2	0.32	0.09	0.17

Na Figura 6.29 são apresentados a FAC e o histograma dos resíduos onde pode-se observar que a dependência linear da série foi capturada.

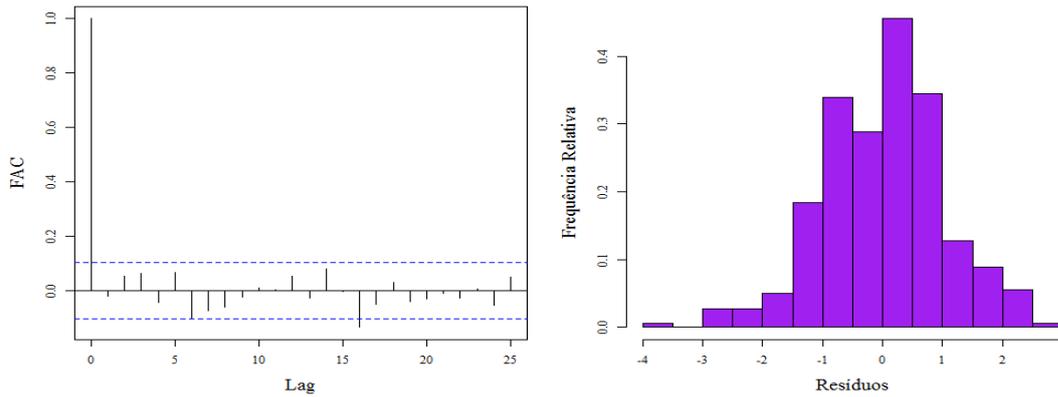


Figura 6.29: FAC e Histograma dos resíduos quantílicos do Modelo 2 - Beta-CNO.

Além disso, através dos resultados anteriores nota-se que a inclusão da variável exógena TSM melhorou o ajuste do modelo, o que pode ser observado nas tabelas apresentadas anteriormente, onde o modelo 2 apresenta melhores critérios e medidas de ajuste quando comparado ao modelo 1.

Na Figura 6.30 são apresentadas as componentes autoregressiva para os 31 anos de dados e de sazonalidade para os últimos 10 anos.

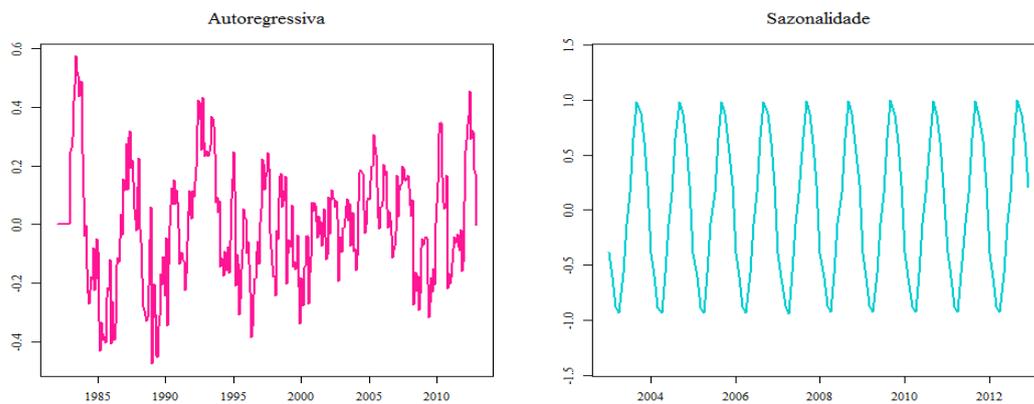


Figura 6.30: Componente sazonal e autoregressiva do ajuste da série de vento para período de estimação do Modelo 1 - Beta-CNO.

Para o modelo Beta-CNO, a estimativa do parâmetro ϕ relacionado à variável exógena TSM da região Niño 3.4 foi igual a 0.062 e seu erro padrão 0.024, o que resulta em uma estatística t igual a 2.58, sendo esta variável significativa ao nível de 5%. Assim, o efeito da TSM da região Niño 3.4 na velocidade do vento, obtido a partir Eq. (5-21) para o modelo Beta-CNO é apresentado na Figura 6.31 para diferentes variações mensais de temperatura.

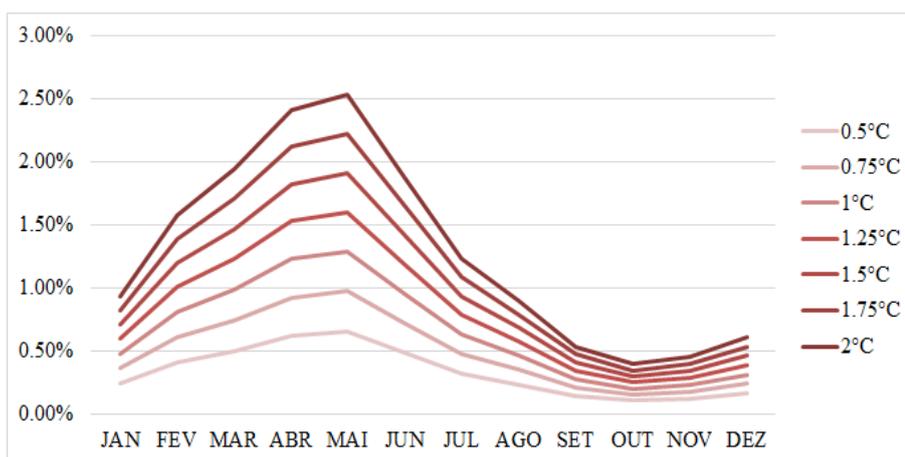


Figura 6.31: Efeito estimado da variável exógena TSM na velocidade do vento para diferentes variações mensais de temperatura - Modelo Beta-CNO.

Através da Figura 6.31 pode-se observar que o efeito da variável exógena TSM no modelo Beta por componentes não-observáveis possui dinâmica similar ao efeito estimado para o modelo Beta com estrutura de evolução SARIMA, uma vez que os maiores efeitos são também observados entre os meses de março e maio, bem como os menores valores entre setembro e novembro. Observa-se também que à medida que a variação mensal da temperatura cresce, o efeito percentual também aumenta. Entretanto, nota-se que o efeito no modelo Beta-CNO é menos suavizado, apresentando menor magnitude que os encontrados no modelo Beta-SARIMA.

O maior efeito da TSM na velocidade do vento foi notado nos meses de abril e maio, onde os mesmos apresentaram os valores de 1.23% e 1.29%, respectivamente. Em relação a decréscimos significativos da TSM entre os meses, o que caracteriza o La Niña, a curva de efeitos é novamente similar à do El Niño em termos de magnitude. No entanto, esta apresenta efeitos negativos. Assim, para uma diminuição de 1°C da TSM entre os meses, o efeito deste na velocidade do vento seria de -1.27% (abril) e -1.33% (maio), sendo estes bem próximos da magnitude dos efeitos anteriores, relacionado ao El Niño.

Vale ressaltar, ainda, que para uma variação da TSM menor que 0.5°C em módulo, assim como nos índices descritos no Capítulo 3, é considerado um estado neutro, isto é, sem a ocorrência dos fenômenos El Niño e La Niña.

Por fim, ao se realizar o teste Diebold-Mariano unilateral com H_1 : “O Modelo Beta-SARIMA possui menor grau de aderência que o Modelo Beta-CNO”, descrito na Seção 5.8, para a comparação da precisão dos modelos Gama em relação às previsões *in-sample* fornecidas pelos mesmos, o teste forneceu um p-valor de 0.026 e uma estatística de teste de -1,95. Assim, considerando um nível de significância de 5%, a hipótese nula deve ser rejeitada por se tratar de um teste unilateral, indicando assim que o modelo Beta-CNO possui maior precisão em relação ao modelo Beta-SARIMA.

7

Conclusões e Extensões

Nessa dissertação utilizou-se uma classe de modelos não-Gaussianos para séries temporais com o objetivo de estimar o impacto do El Niño/La Niña na velocidade dos ventos no Nordeste brasileiro.

Os resultados obtidos para o modelo Weibull com evolução SARIMA mostraram que o mesmo não se ajustou bem às séries mensais de velocidade do vento, o que pôde ser notado pelo forte estrutura senoinal nos resíduos do mesmo. Assim, notou-se que a distribuição de Weibull não foi capaz de capturar satisfatoriamente o comportamento das séries mensais de velocidade.

Em relação aos modelos Gama, estes capturaram boa parte da dinâmica da série de velocidade. A configuração deste modelo por componentes não-observáveis forneceu melhores propriedades como: maior R^2 ; menores RMSE, MAE, MAPE e critérios de informação; maior verossimilhança, quando comparado aos modelos Gama com estrutura de evolução SARIMA. Corroborando com tais observações, o teste Diebold-Mariano unilateral, realizado para a comparação de ambos os modelos, apontou que o modelo Gama-CNO é mais preciso que o modelo Gama-SARIMA ao nível de significância de 5%.

O impacto do El Niño/La Niña na velocidade do vento, calculado através da Temperatura da Superfície do Mar, foi de 1.92% para a variação de 1°C na TSM no modelo Gama com evolução SARIMA e 3.98% para o modelo Gama com evolução CNO, para a mesma variação da TSM. Em relação ao La Niña o efeito foi de -1.88% em relação ao primeiro modelo e -3.82% no segundo.

Os modelos Beta, assim como em outros estudos citados nessa dissertação, apresentaram melhores ajustes às séries de velocidade mensais em relação às demais distribuições aqui estudadas. Assim, a distribuição Beta mostrou-se mais adequada para a modelagem das séries em estudo, uma vez que essa foi capaz de capturar de forma eficaz a dependência linear e não-linear da velocidade do vento, propriedades superiores aos modelos Weibull e Gama.

Além disso, notou-se que a inclusão da variável exógena no modelo Beta melhorou o ajuste do mesmo, concluindo-se que esta possui uma relação significativa com a intensidade dos ventos.

Assim como nos modelos Gama, foi realizado o teste de Diebold-Mariano a fim de comparar a precisão dos modelos Beta-SARIMA e Beta-CNO. O teste forneceu um p-valor de 0.026, evidenciando que, ao nível de 5% de significância, o modelo Beta-CNO possui maior grau de precisão que o modelo Beta-SARIMA.

Devido à maior precisão do modelos Beta-CNO, o efeito do fenômeno El Niño/La Niña pelos modelos Beta foi extraído do mesmo. Com isso, a partir desse modelo o valor estimado para o parâmetro relacionado ao El Niño/La Niña através da TSM foi igual a 0,062. Assim, notou-se que os maiores efeitos da variável exógena TSM na intensidade dos ventos ocorrem nos meses de março, abril e maio, com destaque para o mês de maio. Enquanto os menores foram observados nos meses de outubro e novembro. Além disso, tem-se que, à medida que variasse as diferenças médias entre a TSM de meses consecutivos, esse efeito aumenta.

Logo, pôde-se inferir que a TSM da região Niño 3.4 apresentou um efeito positivo de aproximadamente 1.3%, considerando uma variação de 1°C na TSM entre os meses. Tal efeito corrobora com a hipótese da influência do El Niño e da La Niña na velocidade do vento, uma vez que estes fenômenos se originam da diferença entre temperaturas da superfície do mar do Oceano Pacífico Equatorial. Já em relação a diminuições significativas na TSM, isto é, para uma variação de -1°C, o efeito seria de aproximadamente -1.3%, caracterizando o La Niña.

Destaca-se ainda, que o efeito do fenômeno El Niño/La Niña na velocidade do vento, encontrado através do modelo Beta-SARIMA, foi de aproximadamente o dobro do fornecido pelo modelo Beta-CNO. Isto se deve pela configuração distinta destes dois modelos, e com isso, a escolha do segundo para a extração do efeito final, foi em virtude do modelo Beta-CNO ter apresentado melhores propriedades que o Beta-SARIMA, como abordado anteriormente.

Este efeito poderia acarretar em influência de ordem maior na potência gerada pela usina, uma vez que a potência é função da velocidade do vento ao cubo.

A partir do efeito do El Niño/La Niña na velocidade do vento mensurado neste trabalho, é possível realizar uma análise do seu impacto na receita futura de um parque eólico, considerando os anos de ocorrência de tais fenômenos. Isto poderia ser realizado a partir da previsão, por exemplo, da energia gerada para uma determinada usina eólica, convertido em US\$/KWh, considerando a atuação ou não do El Niño/La Niña num determinado período futuro. As projeções futuras da velocidade dos ventos realizadas pelos modelo GAS

são obtidas através de simulações de Monte Carlo, como em Matos (2013). Essas projeções seriam então obtidas condicionadas a ocorrência ou não de El Niño/La Niña. Portanto, as receitas estimadas nesse período considerariam dois cenários distintos: com e sem a presença dessas anomalias climáticas. Um maior refinamento desse resultado poderia ser obtido considerando a incerteza estatística na mensuração desse efeito na receita futura do parque eólico. Isto pode ser realizado através da construção de intervalos de confiança para a receita futura. Assim, ao invés de simular a velocidade do vento considerando apenas um valor dos parâmetros associados ao El Niño/La Niña, as simulações seriam realizadas a partir desse conjunto de parâmetros, incorporando na previsão da velocidade do vento, a incerteza das estimativas dos parâmetros. A partir dessas novas simulações, seriam construídos intervalos de confiança para a receita futura do empreendimento.

Cabe ressaltar que o foco dessa dissertação não é a previsão da velocidade do vento e sim a busca pelo melhor ajuste desta série para inferir sobre o possível efeito do fenômeno El Niño e La Niña na velocidade. Para fins de previsão, pode-se explorar outras variáveis exógenas tais como temperatura e pressão do ar, que são fortemente correlacionadas com a velocidade.

A influência dos fenômenos El Niño e La Niña através da temperatura da superfície do mar do Oceano Atlântico no regime de ventos na região nordeste do Brasil não foi abordado nesse trabalho e esta é uma relação interessante que pode ser explorada. Além disso, a influência desses fenômenos em usinas de energia eólica que estão situadas fora da zona litorânea possivelmente é diferenciada.

Outro interesse está em desenvolver melhorias no que diz respeito à inicialização dos modelos GAS através de novas heurísticas, uma vez que nessa dissertação os valores iniciais para os parâmetros estáticos e variantes do modelo foi obtida através de tentativa e erro.

Referências Bibliográficas

- [1] Advancing Reanalyses. **Reanalyses**. Disponível em: <<http://reanalyses.org/>>. Acesso em: Fevereiro de 2013.
- [2] AMARAL, B. M. **Modelos VARX para geração de cenários de vento e vazão aplicados à comercialização de energia**. 2011. 155f. Dissertação (Mestrado em Engenharia Elétrica) - PUC-Rio, Rio de Janeiro, 2011.
- [3] AMARANTE, O. A. C.; BROWER, M.; ZACK, J.; SÁ, A. L. 2001, **Atlas do Potencial Eólico Brasileiro**. Camargo Schubert Engenharia Eólica, TrueWind Solutions, Centro de Referência para Energia Solar e Eólica Sérgio de Salvo Brito/CRESESB, Centro de Pesquisas de Energia Elétrica/CEPEL, Eletrobrás, Ministério de Minas e Energia, Governo Federal, Brasília, Brasil.
- [4] ---- ; SCHULTZ, D.J.; BITTENCOURT, R.M.; ROCHA, N.A. Wind/hydro complementary seasonal regimes in Brazil. **DEWI Magazine**, n. 19, p. 79-86, 2001.
- [5] AHRENS, C. D. **Essentials of Meteorology: an invitation to the atmosphere**. 6.ed. New York, 2008. 506p.
- [6] ARAGÃO, J.; Roucou, P.; Harzallah, A.; Fontaine, B.; Janicot, S. A Influência dos Oceanos Atlântico e Pacífico sobre a Circulação atmosférica e a chuva na Região semi-árida do Nordeste do Brasil: Simulação e Observação. In: Congresso Brasileiro de Meteorologia, 9, 1996, Campos do Jordão. **Anais ...** 1996, p.830-833.
- [7] Associação Brasileira de Energia Eólica/ABEEólica. Números do setor. Disponível em: <<http://www.portalabeeolica.org.br/>>. Acesso em: Abril de 2014.
- [8] AZRULHISHAM, E. A.; ZAKARIA, K. P.; SAMIZEE, A.; JUHARI, M. B. M. Pearson System Distribution Approximation in Wind Energy Potential Analysis. In: International Conference on Future Environment and

- Energy, 2. 2012, Singapore. **International Proceedings of Chemical, Biological and Environmental Engineering (IPCBEE)**. Singapore: IACSIT Press, 2012. v.28, P.61-65.
- [9] Balanço Energético Nacional 2013 – Ano base 2012: Relatório Síntese. Rio de Janeiro: EPE, 2013. Anual. 55 p. 2013.
- [10] BLISS, E.W.; WALKER, G.T. World Weather V. **Memoirs of the Royal Meteorological Society**, v.4, n.36, p. 53-84, 1932.
- [11] Boletim Mensal de Dados do Setor Eólico - Público. São Paulo: ABEEólica, 2013. Mensal. n.7, 10 p. 2013.
- [12] Boletim Mensal de Dados do Setor Eólico - Público. São Paulo: ABEEólica, 2014. Mensal. n.1, 13 p. 2014.
- [13] BROWN, B. G.; KATZ, R. W.; MURPHY, A. H., Time Series Models to Simulate and Forecast Wind Speed and Wind Power. **Journal of Climate and Applied Meteorology**, v. 23, p. 1184-1195, 1984.
- [14] CAVALCANTI, I.F. de A.; FERREIRA, N.J.; SILVA, M.G.A.J. da; DIAS, M.A.F. da S. (Org) **Tempo e clima no Brasil**. São Paulo: Oficina de Textos, 2009. 463p.
- [15] Climate Prediction Center/CPC-NOAA. El Niño - Southern Oscillation (ENSO). Disponível em: <<http://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/enso.shtml>>. Acesso em: Março de 2013.
- [16] _____. Monitoring & Data. Disponível em:<<http://www.cpc.ncep.noaa.gov/data/>>. Acesso em: Julho de 2013.
- [17] COVEY, D. L.; HASTENRATH,S. The Pacific El Niño phenomenon and the Atlantic circulation. **Monthly Weather Review**, v. 106, p. 1280–1287, 1978.
- [18] COX, D. R. Statistical analysis of time series: some recent developments. **Scandinavian Journal of Statistics**, v. 8, p. 93-115, 1981.
- [19] CREAL, D.; KOOPMAN, S.J.; LUCAS, A. A General Framework for Observation Driven Time-Varying Parameter Models. **Tinbergen Institute Discussion**, n. 108, 2008. Não paginado.

- [20] ____ . Generalized Autoregressive Score Models with Applications. **Journal of Applied Econometrics**, v. 28, 2013. 777-795.
- [21] DAHLMAN, L. 2009. **Climate Variability: Oceanic Niño Index**. Disponível em: <<http://www.climate.gov/news-features/understanding-climate/climate-variability-oceanic-ni%C3%B1o-index>>. Acesso em: Agosto de 2013.
- [22] DIEBOLD, F.X.; MARIANO, R. Comparing Predictive Accuracy. **Journal of Business and Economic Statistics**, n. 13, p.253-265, 1995.
- [23] Earth System Research Laboratory/ESRL. Multivariate ENSO Index (MEI). Disponível em: <<http://www.esrl.noaa.gov/psd/enso/mei/>>. Acesso em: Abril de 2013.
- [24] GRUBB, M. J.; MEYER, N. I. Wind energy: resources, systems and regional strategies. In: JOHANSSON, T. B. *et al.* **Renewable energy: sources for fuels and electricity**. Washington, D.C.: Island Press, 1993. cap. 3, p. 157-212.
- [25] Global Wind Report: Annual Market Update 2012. GLOBAL WIND ENERGY COUNCIL/GWEC. Anual Disponível em: <<http://www.gwec.net>>. Acesso em: Abril de 2013.
- [26] HARPER, B.R.; KATZ, R.W.; HARRIS, R.C., Statistical Methods for Quantifying the Effect of the. El Niño–Southern Oscillation on Wind Power in the Northern Great Plains of the United States. **WIND ENGINEERING**. v. 31, n. 3, p. 123–137, 2007.
- [27] HASTENRATH, S.; HELLER, L., Dynamics of climatic hazards in northeast Brazil. **Quarterly Journal of the Royal Meteorological Society**, n. 103, p. 77-92, 1977.
- [28] Instituto Brasileiro de Geografia e Estatística/IBGE, 2013. Disponível em: <<http://www.censo2010.ibge.gov.br/sinopse>>. Acesso em: Agosto de 2013.
- [29] JUSTUS, C. G.; HARGRAVES, W. R.; YALCIN, A. Nationwide Assessment of Potential Output from Wind-Powered Generators. **Journal of Applied Meteorology**, n. 15, p.673–678, 1976.
- [30] KALLIOVIRTA, L. **Diagnostic Tests Based on Quantile Residuals for Nonlinear Time Series Models**. 2009. 165 f. Thesis (Doctorate in Economics) - University of Helsinki, Helsinki, 2009.

- [31] KOOPMAN, S. J.; LUCAS, A.; SCHARTH, M. Predicting Time-Varying Parameters with Parameter-Driven and Observation-Driven Models. **Tinbergen Institute Discussion Paper**, n.20, 2012. Não paginado.
- [32] LEITE, M. L.; VIRGENS FILHO, J. S. Avaliação da distribuição beta como modelo probabilístico para análise de dados de velocidade do vento para Ponta Grossa. – PR. **Publ.UEPG Ci. Exatas Terra, Ci. Agr. Eng.**, v.13, n.1, p.51-57, 2007.
- [33] LOPES, F. Z.; DINIZ, G. B.; MARQUES, J. R. Relação entre o Multivariate Enso Index (MEI) e a TSM das Regiões dos Niños com a Precipitação em Regiões Homogêneas do Estado do Rio Grande do Sul. **Anuário do Instituto de Geociências - UFRJ**, v.30, n.2, p. 11-22, 2007.
- [34] MARINHO, Manoel Henrique da Nóbrega; AQUINO, Ronaldo Ribeiro Barbosa de. Oferta de energia através da complementariedade zazonal Hidro-Eólica. **Revista PCH Notícias & SHP NEWS**, Ano 11, n.40, p. 36-42, jan./fev./mar. 2009
- [35] MARKHAM, C. G.; MCLAIN, D. R. Sea surface temperature related to rain in Ceará, north-eastern Brazil. **Nature**. v. 265, p. 320-323, 1977.
- [36] MARTINS, F.R.; GUARNIERI, R.A.; PEREIRA, E.B. O aproveitamento da energia eólica. **Revista Brasileira de Ensino de Física**, São Paulo, v.30, n.1, 1304, 2008.
- [37] MATOS, G. G. **Modelos GAS Aplicados a Séries Temporais de Vazão e Vento**, Rio de Janeiro, 2013. 103p. Dissertação (Mestrado em Engenharia Elétrica) - Pontifícia Universidade Católica do Rio de Janeiro/PUC-Rio
- [38] Ministério de Minas e Energia/MME. O PROINFA. Disponível em:<<http://www.mme.gov.br/programas/proinfa/>>. Acesso em: Março de 2013.
- [39] MORETTIN, P. A.; Toloi, C.M.C. **Análise de Séries Temporais**. 1.ed. São Paulo: Projeto Fisher, 2006, 564p.
- [40] National Oceanic and Atmospheric Administration/NOAA. Equatorial Pacific Sea Surface Temperatures. Disponível em:<www.ncdc.noaa.gov/teleconnections/enso/indicators/sst.php>. Acesso em: Abril de 2014.

- [41] NETO, O. L. M.; OLIVEIRA, J. L.; COSTA, A. A.; SOMBRA, S. S. Impactos da Circulação geral em casos de El Niño e La Niña no Potencial Eólico do Nordeste Brasileiro. In: Congresso Brasileiro de Meteorologia, 14. 2006, Florianópolis. **Anais eletrônicos...** Florianópolis: SBMET, 2006. Disponível em: <<http://www.cbmet.com/>>. Acesso em: Fevereiro de 2013.
- [42] OLIVEIRA, G.S. **El Niño e você: O fenômeno Climático**. 1.ed. São José dos Campos, 1999. 115p.
- [43] OLIVEIRA, J. L.; COSTA, A. A. Estudo de variabilidade do vento em escala sazonal sobre o Nordeste brasileiro utilizando o RAMS: os casos de 1973-1974 e 1982-1983. **Revista Brasileira de Meteorologia**, v.26, n.1, p. 53-66, 2011.
- [44] PINTO, L.; SICA, D.; SZCZUPAK, J.; SILVA, D.; AMBRIZZI, T. Projeção de Cenários Climatológicos Hidro-Eólicos: Desafios, Perigos, Riscos e Mitigações. In: SNPTEE SEMINÁRIO NACIONAL DE PRODUÇÃO E TRANSMISSÃO DE ENERGIA ELÉTRICA, 21. 2011, Florianópolis. **Anais eletrônicos...** Florianópolis: CIGRÉ, 2011. Disponível em: <<http://www.cigre.org.br/>>. Acesso em: Fevereiro de 2013.
- [45] RAO, V. B.; LIMA, M. C.; FRANCHITO, S.H., 1993. Seasonal and Interannual Variations of Rainfall over Eastern Northeast Brazil. **Journal of Climate**, v.6, p. 1754-1763, 1993.
- [46] SANSIGOLO, C.A.; DINIZ, G.B.; MACHADO J.P. Impacto das fases extremas do ENOS nas temperaturas médias sazonais do Rio Grande do Sul. In: Congresso Brasileiro de Meteorologia, 14. 2006, Florianópolis. **Anais eletrônicos...** Florianópolis: SBMET, 2006. Disponível em: <<http://mtc-m15.sid.inpe.br/>>. Acesso em: Maio de 2013.
- [47] SCHAEFFER, R.; SZKLO, A. S.; LUCENA, A. F. P.; SOUZA, R. R.; BORBA, B. S. M. C.; COSTA, I. V. L.; JÚNIOR, A. O. P.; CUNHA, S. H. F. Mudanças Climáticas e Segurança Energética no Brasil. Rio de Janeiro: PPE - UFRJ, 2008, 67p. Disponível em: <<http://mudancasclimaticas.cptec.inpe.br>>. Acesso em: Março de 2013. 67p.
- [48] SOUSA, E. P. P. **Relações entre as anomalias de TSM do Atlântico e Pacífico e as precipitações na Amazônia oriental**. São José dos

Campos, 2003. 78p. Dissertação (Mestrado em Meteorologia) - Instituto Nacional de Pesquisas Espaciais/INPE.

- [49] YILMAZ, V.; ÇELİK, H. E. A statistical approach to estimate the wind speed distribution: the case of Gelibolu region. **Doğuş Üniversitesi Dergisi**, v.9, n.1, p. 122-132, 2008.

A Apêndice A

A.1

Modelo Gama - SARIMA com os parâmetros λ e α variantes no tempo

Com o intuito de aprimorar a captura da dinâmica das séries de vento, propõe-se um modelo com os parâmetros de forma e escala variantes no tempo.

Nesta abordagem, utilizou-se a mesma parametrização para a distribuição Gama apresentada na subseção anterior. No entanto, o interesse presente é considerar os parâmetros α e λ variantes no tempo. Assim, tem-se que:

$$f_{1,t} = \alpha_t \quad f_{2,t} = \lambda_t$$

E uma escolha natural e preferível à f_t seria:

$$\tilde{f}_t = \begin{bmatrix} \tilde{f}_{1,t} \\ \tilde{f}_{2,t} \end{bmatrix} = \begin{bmatrix} \ln \alpha_t \\ \ln \lambda_t \end{bmatrix}$$

Nota-se novamente a utilização da parametrização pelo logaritmo neperiano, para garantir a não-negatividade dos valores dos parâmetros originais de densidade Gama.

Assim, obtém-se o vetor *score* através da derivada de $\ln p(y_t|f_t, F_t; \theta)$, que é a contribuição de cada valor de y_t , em relação a ambos os parâmetros variantes no tempo, como apresentado abaixo.

$$\nabla_t = \begin{bmatrix} \frac{\partial \ln p(y_t|f_t, F_t; \theta)}{\partial f_{1,t}} \\ \frac{\partial \ln p(y_t|f_t, F_t; \theta)}{\partial f_{2,t}} \end{bmatrix}$$

e através das relações apresentadas na seção 5.3 tem-se que:

$$\tilde{f}_t = h(f_t) = \begin{bmatrix} \ln \alpha_t \\ \ln \lambda_t \end{bmatrix}$$

e, portanto, para a parametrização escolhida

$$\dot{h}(f_t) = \begin{bmatrix} \frac{1}{\alpha_t} \\ \frac{1}{\lambda_t} \end{bmatrix}$$

E a partir deste, calcula-se o vetor *score* referente à mesma como é apresentado a seguir:

$$\tilde{\nabla}_t = \frac{\partial \ln p(y_t | f_t, F_t; \theta)}{\partial \tilde{f}_t} = (\dot{h}_t)^{-1} \nabla_t = \begin{bmatrix} \alpha_t \left[\ln y_t - \frac{y_t}{\lambda_t} + \ln \frac{\alpha_t}{\lambda_t} + 1 - \psi_1(\alpha_t) \right] \\ \frac{\alpha_t (y_t - \lambda_t)}{\lambda_t} \end{bmatrix}$$

Portanto, a matriz de Informação de Fisher é representada pela seguinte matriz diagonal:

$$I_{t|t-1}(\alpha, \lambda) = \begin{bmatrix} \psi_2(\alpha) - \frac{1}{\alpha} & 0 \\ 0 & \frac{\alpha}{\lambda^2} \end{bmatrix}$$

esta em relação à \dot{h}_t (Equação 5-8) será dada por:

$$\tilde{I}_{t|t-1}(\alpha_t, \lambda_t) = \begin{bmatrix} \alpha_t^2 \psi_2(\alpha_t) - \alpha_t & 0 \\ 0 & \alpha_t \end{bmatrix}$$

onde $\psi_k(x) = \frac{\partial^k \ln \Gamma(x)}{\partial x^k}$.

Logo, o *score* ponderado pode ser obtido a partir da equação $s_t = \tilde{I}_{t|t-1}^{-1} \tilde{\nabla}_t$ para diferentes ponderações como: raiz quadrada da pseudo-inversa da matriz $I_{t|t-1}$, pseudo-inversa da matriz $I_{t|t-1}$ e matriz identidade I , já apresentadas anteriormente.

Os modelos GAS com evolução SARIMA propostos com os parâmetros λ e α_t variantes no tempo possuem a seguinte forma:

$$\begin{aligned} \tilde{f}_{t+1} = & W + A_1 s_t + A_2 s_{t-1} + A_3 s_{t-2} + A_4 s_{t-10} + A_5 s_{t-11} \\ & + B_1 \tilde{f}_t + B_2 \tilde{f}_{t-1} + B_3 \tilde{f}_{t-2} + B_4 \tilde{f}_{t-10} + B_5 \tilde{f}_{t-11} \end{aligned}$$

onde A_i e B_j são matrizes representadas por:

$$A_1 = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{21} \end{bmatrix}, \quad A_2 = \begin{bmatrix} a_{12} & 0 \\ 0 & a_{22} \end{bmatrix}, \quad A_3 = \begin{bmatrix} a_{13} & 0 \\ 0 & a_{23} \end{bmatrix},$$

$$A_4 = \begin{bmatrix} a_{111} & 0 \\ 0 & a_{211} \end{bmatrix}, \quad A_5 = \begin{bmatrix} a_{121} & 0 \\ 0 & a_{222} \end{bmatrix}$$

$$B_1 = \begin{bmatrix} b_{11} & 0 \\ 0 & b_{21} \end{bmatrix}, \quad B_2 = \begin{bmatrix} b_{12} & 0 \\ 0 & b_{22} \end{bmatrix}, \quad B_3 = \begin{bmatrix} b_{13} & 0 \\ 0 & b_{23} \end{bmatrix},$$

$$B_4 = \begin{bmatrix} b_{111} & 0 \\ 0 & b_{211} \end{bmatrix}, \quad B_5 = \begin{bmatrix} b_{121} & 0 \\ 0 & b_{222} \end{bmatrix}$$

Nesta configuração do modelo GAS, buscou-se capturar as dinâmicas de curto prazo pelos *lags* 1,2 e 3, bem como dinâmicas de longo prazo, caracterizando a sazonalidade, através dos *lags* 11 e 12, para ambos os parâmetros.

A inclusão de variáveis exógenas nesse modelo dá-se de forma análoga ao anterior à apresentada no Capítulo 5, sendo estas incorporadas de forma aditiva na média condicional λ_t , ou seja,

$$\tilde{f}_t = \begin{bmatrix} \tilde{f}_{1,t} \\ \tilde{f}_{2,t} \end{bmatrix} = \begin{bmatrix} \ln \alpha_t \\ \ln \lambda_t - \sum_{k=1}^r \phi_k g(x_{t-k+1}) \end{bmatrix}$$

ou de forma equivalente,

$$\begin{bmatrix} \alpha_t \\ \lambda_t \end{bmatrix} = \begin{bmatrix} \exp \tilde{f}_{1,t} \\ \exp\{\tilde{f}_{2,t} + \sum_{k=1}^r \phi_k g(x_{t-k+1})\} \end{bmatrix}$$

onde \tilde{f}_t é o vetor de parâmetros variantes no tempo e x_t é o vetor de variáveis exógenas a serem consideradas no modelo.

Os testes realizados com essa configuração do modelo Gama mostraram que o mesmo foi eficaz na correção da forte dependência nos resíduos ao quadrado (Figura A.1). No entanto, o ajuste desse modelo gerou outros problemas como a presença de algumas defasagens significativas no Correlograma dos resíduos e medidas de ajuste inferiores, fatos estes não observados anteriormente nos modelos com apenas o parâmetro de escala variante no tempo. Além disso, tal modelo possui o dobro de parâmetros das configurações anteriores, o que acarretou num grande aumento do tempo computacional da estimação desse modelo. Sendo assim, optou-se por não utilizar esse modelo

nas análises finais dessa dissertação, uma vez que este não apresentou resultados satisfatórios, dificultando assim sua viabilidade.

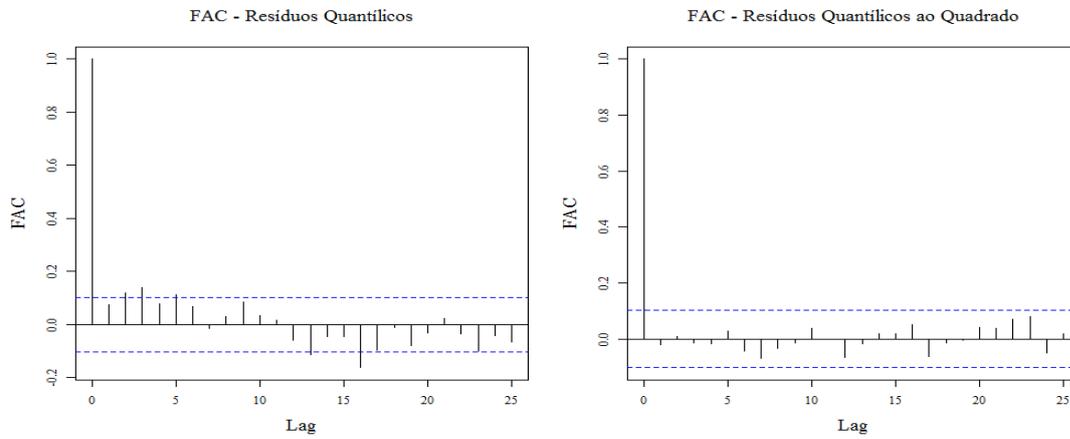


Figura A.1: FAC dos resíduos quantílicos Gama-SARIMA com dois parâmetros variantes no tempo.

Tabela A.1: Estimativas dos parâmetros dos modelos Gama-SARIMA e erros padrão.

Parâmetro	Modelo 1		Modelo 2	
	Est	E.P.	Est.	E.P.
ω	0.130	0.046	0.081	0.020
a_1	0.420	0.051	0.387	0.055
a_2	-0.591	0.086	-0.466	0.120
a_3	0.320	0.052	0.160	0.104
a_{11}	0.050	0.029	0.022	0.067
a_{12}	-0.106	0.028	-0.075	0.073
b_1	2.204	0.027	1.867	0.190
b_2	-1.854	0.024	-1.182	0.342
b_3	0.551	0.033	0.243	0.163
b_{11}	0.095	0.027	0.299	0.077
b_{12}	-0.064	0.028	-0.285	0.076
α	45.154	3.351	47.008	3.491
$\phi_4 TSM$	-	-	0.019	0.005

Tabela A.2: Estimativas dos parâmetros dos modelos Gama-CNO e erros padrão.

Parâmetro	Modelo 1		Modelo 2	
	Est.	E.P.	Est.	E.P.
ω	1.922	0.022	0.873	0.321
a_1	-0.009	0.013	-0.007	0.013
a_2	0.434	0.047	0.394	0.048
ϕ_1	0.796	0.048	0.754	0.059
α	54.442	4.045	56.594	4.206
$\phi_4 TSM$	-	-	0.038	0.011

Tabela A.3: Estimativas dos parâmetros dos modelos Beta-SARIMA e erros padrão.

Parâmetro	Modelo 1		Modelo 2	
	Est	E.P.	Est.	E.P.
ω	0.362	0.077	0.118	0.054
a_1	0.294	0.051	0.289	0.053
a_2	-0.358	0.085	-0.365	0.089
a_3	0.208	0.052	0.208	0.064
a_{11}	0.108	0.031	0.086	0.036
a_{12}	-0.170	0.034	-0.142	0.036
b_1	1.757	0.050	1.931	0.071
b_2	-0.980	0.090	-1.369	0.083
b_3	0.107	0.060	0.340	0.050
b_{11}	0.273	0.012	0.238	0.020
b_{12}	-0.268	0.020	-0.193	0.036
α	14.881	1.023	17.486	1.289
$\phi_4 TSM$	-	-	0.047	0.012

Tabela A.4: Estimativas dos parâmetros dos modelos Beta-CNO e erros padrão.

Parâmetro	Modelo 1		Modelo 2	
	Est.	E.P.	Est.	E.P.
ω	3.569	0.088	1.892	0.657
a_1	-0.003	0.003	-0.003	0.003
a_2	0.099	0.012	0.094	0.013
ϕ_1	0.857	0.038	0.830	0.047
α	18.833	1.390	19.234	1.420
$\phi_4 TSM$	-	-	0.062	0.024