Henrique Helfer Hoeltgebaum

**Previsão da Densidade Conjunta de Fator de
Capacidade Eólico via Modelo GAS
Multivariado**

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

Rio de Janeiro
Março de 2015

Henrique Helfer Hoeltgebaum

**Previsão da Densidade Conjunta de Fator de
Capacidade Eólico via Modelo GAS
Multivariado**

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. Alexandre Street de Aguiar
Departamento de Engenharia Elétrica – PUC-Rio

Prof. Eduardo Fraga Lima de Melo
SUSEP

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico

Rio de Janeiro, 25 de Março de 2015

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

Henrique Helfer Hoeltgebaum

Graduou-se em Estatística pela Universidade Federal do Rio Grande do Sul, em 2012. Desde Fevereiro de 2014, é analista de operações estruturadas na Wilson Sons.

Ficha Catalográfica

Hoeltgebaum, Henrique Helfer

Previsão da Densidade Conjunta de Fator de Capacidade Eólico via Modelo GAS Multivariado / Henrique Helfer Hoeltgebaum; orientador: Cristiano Augusto Coelho Fernandes. — Rio de Janeiro : PUC–Rio, Departamento de Engenharia Elétrica, 2015.

v., 97 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. Modelos GAS. 3. Energia Eólica. 4. Parâmetros variantes no tempo. 5. Cópulas. I. Fernandes, Cristiano Augusto Coelho. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD:621.3

Aos meus pais, Elizabel e Emílio.

Agradecimentos

Chegando ao final dessa jornada, não posso deixar de agradecer a algumas pessoas especiais que estiveram, de alguma maneira, presentes durante a realização desse trabalho. Saibam que nada disso teria sido possível sem vocês.

Início os meus agradecimentos fazendo menção ao grande amor da minha vida, Adriana. Obrigado por ter aceitado embarcar comigo nessa jornada, relevando sempre com bom humor os meus momentos ausentes causados pela dissertação e principalmente em aceitar o desafio de vir morar no Rio de Janeiro.

Agradeço imensamente a minha família. Iniciando pela minha mãe, Elizabel, sem a qual nada disso teria sido possível. Obrigado por me dar todo o apoio necessário quando eu mais precisei para me mudar em função do mestrado. Ao meu pai Emílio, pelos sábios conselhos dados durante a minha jornada acadêmica, sem os quais não teria tomado as decisões que me trouxeram até aqui hoje. E finalmente minha irmã Laura, que sempre esteve ao meu lado nos momentos difíceis acreditando no meu potencial quando mais precisei. Amo vocês!

Dedico um agradecimento especial ao meu orientador Cristiano Fernandes, inicialmente pela grande influência na minha formação acadêmica e por toda a atenção dada durante a realização deste trabalho, desde atendimento aos sábados, aos horários mais complicados durante a semana. Muito obrigado!

Agradeço aos meus amigos e também companheiros do LAMPS pelas valiosas discussões teóricas e pela amizade. Sem vocês muito desse trabalho não teria sido feito. Por ordem alfabética, Alexandre Moreira da Silva, Bruno Fânzeres, César Neves, Gustavo Amaral, Joaquim Lacombe, Lucas Freire, Mario Souto e Tiago Barata.

Ao pessoal da minha equipe da Wilson Sons, André Ferreira, Isabela Motta e Pedro Caldas, por me proporcionarem um ambiente de trabalho favorável à realização desse trabalho.

A todo o corpo docente do Departamento de Engenharia Elétrica da PUC-Rio que tiveram participação na minha formação acadêmica. E também à PUC-Rio e CAPES pelos auxílios concedidos.

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

Resumo

Hoeltgebaum, Henrique Helfer; Fernandes, Cristiano Augusto Coelho. **Previsão da Densidade Conjunta de Fator de Capacidade Eólico via Modelo GAS Multivariado**. Rio de Janeiro, 2015. 97p. Dissertação de Mestrado — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Neste trabalho usamos o arcabouço dos modelos GAS para gerar previsões conjuntas de fator de capacidade eólico, pertencentes a diferentes usinas localizadas em áreas geográficas distintas. Esses cenários são insumos para gerar uma distribuição de fluxo de caixa associada a um portfólio de contratos atrelados aos parques eólicos em questão. Inicialmente modelamos as densidades marginais via um modelo GAS, supondo densidade Beta. De maneira a capturar a estrutura de dependência entre esses fatores de capacidade, usamos uma cópula t-Student com a matriz de correlação também sendo atualizada via mecanismo GAS. Uma das contribuições importantes desse trabalho para o setor elétrico está na geração de cenários conjuntos apenas em um passo, evitando a necessidade de modelar variáveis transformadas e posteriormente transforma-las para retornar às suas respectivas escalas originais. Assim como é feito no caso supondo normalidade para as marginais. Como é sabido, exponenciar valores simulados a partir de uma densidade normal pode gerar resultados equivocados para fatores de capacidade eólico, e por propagação, isso pode afetar severamente as medidas de risco que são obtidas a partir da distribuição simulada de fluxo de caixa associada com o portfólio das usinas eólicas. Nossos resultados mostram que quando a dependência é levada em consideração, os fluxos de caixa tendem a ser maiores do que quando ignora-se a dependência.

Palavras-chave

Modelos GAS ; Energia Eólica ; Parâmetros variantes no tempo ; Cópu-
las.

Abstract

Hoeltgebaum, Henrique Helfer; Fernandes, Cristiano Augusto Coelho (advisor). **Forecast of the joint density of Wind Capacity Factor through the use of a multivariate GAS model.** Rio de Janeiro, 2015. 97p. MSc. Dissertation — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

In this work we use the framework of GAS models to generate joint forecasts for capacity factors of several wind plants belonging to different geographical areas. Such scenarios are then used as input to raise the distribution of cash flows associated with a portfolio of contracts attached to these wind plants. We first model the marginal density of each capacity factor using a GAS model with Beta density. In order to capture the observed dependence among these capacity factors, we use a copula t-Student with correlation matrix evolving through a GAS mechanism. One of the important contributions of our framework is that generation of scenarios is accomplished in just one step, avoiding the need of transforming back variables to its original scale, as it is the case under a Gaussian assumption for the marginals. As it is known, exponentiation of simulated Gaussian values can result in unrealistic sampling paths for the wind capacity factor, and by propagation, this can badly affect the risk measures obtained from the simulated distribution of the cash flows associated with a particular portfolio of wind plants. Our results shows that when taking into account dependence the cash flows are higher than when ignoring dependence.

Keywords

GAS Models; Wind Energy; Time Varying Parameters; Copulas.

Sumário

1	Introdução	13
1.1	Motivação e Relevância do Tema	13
1.2	Revisão Bibliográfica	15
1.3	Objetivos e Contribuições	16
1.4	Organização	17
2	Generalized Autoregressive Score Models	18
2.1	Introdução	18
2.2	Especificações do modelo e propriedades	19
2.3	Parametrização	21
2.4	Estimação por Máxima Verossimilhança	22
3	Modelo Beta SARIMA	24
3.1	Especificação	24
3.2	Modelo Beta univariado	25
3.3	Diagnósticos	27
3.4	Previsão fora da amostra	29
4	Copulas e Dependência	30
4.1	Introdução à teoria de cópulas	30
4.2	Cópulas Condicionais	32
5	Modelo <i>t-Student</i> multivariado	37
5.1	Densidade <i>t-Student</i> multivariada	37
5.2	Notações matriciais e definições	38
5.3	Atualização de Σ_t via GAS	39
5.4	Especificação de Cópula variante no tempo	41
5.5	Previsão do modelo GAS com densidade preditiva $t_{(\nu, \Sigma_t)}$	42
5.6	Estimação e propriedades estatísticas	44
5.7	Simulação	50
6	Uso dos modelos GAS para simular e prever históricos de FC de usinas eólicas	53
6.1	Introdução	53
6.2	Análise descritiva	56
6.3	Modelagem do FC das usinas via Beta GAS SARIMA(p,q)	59
6.4	Modelagem da estrutura de dependência via GAS <i>t-Student</i> (p,q)	63
6.5	Previsão Condicional	72
6.6	Análise de risco no Ambiente de Comercialização Livre (ACL)	76
6.7	Análise dos cenários condicionais gerados	79
7	Considerações finais	87

A	Prova da Cópula ser invariante a transformações crescentes	95
B	Casos especiais da densidade <i>t-Student</i> multivariada	96
C	Conditional Sampling Bivariado	97

Lista de figuras

5.1	Ilustração gráfica dos valores condicionais de $y_{t+k}^{*(m)}$.	44
5.2	Diagrama de blocos para a modelagem incluindo a estrutura de dependência especificada pelo modelo $GAS t-Student(p,q)$.	45
5.3	Gráficos dos cinco PGD simulados com $T=300$ e $T=1000$.	52
6.1	Distribuição geográfica das três usinas usinas. <i>Fonte: Google Maps.</i>	54
6.2	Mosaico de imagens de satélite, sobreposto ao modelo de relevo retirado de Amarante <i>et al.</i> (2001).	56
6.3	Gráfico das séries de FC para as usinas de Rio do Fogo, Icaraizinho e Enacel, resultantes do modelo de extensão de histórico com seus respectivos Box-Plots mensais entre janeiro de 1981 a dezembro de 2011.	57
6.4	Gráfico de dispersão entre as três usinas.	58
6.5	Histogramas e QQ-plots para dos resíduos quantílicos produzidos pelos modelos Beta GAS SARIMA(p,q) aplicados nas séries de FC das três usinas eólicas.	61
6.6	FAC e FAC^2 das séries de resíduos quantílicos associados aos FC eólicos das usinas RF, IC e EN.	62
6.7	Diagrama de dispersão das pseudo-observações u_t , produzidas pelos modelos univariados duas a duas. Na primeira linha estão os diagramas de dispersão entre as pseudo-observações de EN e IC na esquerda e o diagrama de dispersão entre as pseudo-observações de EN e RF na direita. Na segunda linha encontra-se o diagrama de dispersão entre as pseudo-observações de IC e RF.	66
6.8	Diagrama de dispersão 3D das pseudo-observações produzidas pelos modelos Univariados.	67
6.9	Comparação entre a densidade de uma normal padrão e uma densidade $t-Student$ com $\nu = 340$.	69
6.10	Dinâmica de w_t ao longo do tempo no modelo de cópula $t-Student$ com $\hat{\nu} = 340$.	70
6.11	Elementos de $\hat{\Sigma}_t$ atualizados pelo mecanismo GAS entre $\tilde{y}_{1t}$ e $\tilde{y}_{2t}$, que representam as usinas de RF e IC respectivamente.	70
6.12	Elementos de $\hat{\Sigma}_t$ atualizados pelo mecanismo GAS entre $\tilde{y}_{1t}$ e $\tilde{y}_{3t}$, que representam as usinas de RF e EN respectivamente.	71
6.13	Elementos de $\hat{\Sigma}_t$ atualizados pelo mecanismo GAS entre $\tilde{y}_{2t}$ e $\tilde{y}_{3t}$, que representam as usinas de IC e EN respectivamente.	71
6.14	Previsões 12 passos à frente dos FC pelos modelos Beta GAS SARIMA(p,q) e VARX para o período de janeiro à dezembro de 2011.	75
6.15	Análise descritiva dos cenários de FC gerados para a usina de RF no ano de 2011 entre os modelos Beta $t-GAS$ e VARX contra a série histórica.	81

- 6.16 Análise descritiva dos cenários de FC gerados para a usina de IC no ano de 2011 entre os modelos Beta t -GAS e VARX contra a série histórica. 82
- 6.17 Análise descritiva dos cenários de FC gerados para a usina de EN no ano de 2011 entre os modelos Beta t -GAS e VARX contra a série histórica. 83

Lista de tabelas

5.1	Medidas de aderência do estudo de simulação.	51
6.1	Matriz de distância entre as usinas eólicas.	53
6.2	Logarítmo da verossimilhança e critérios de informação para os modelos Beta GAS SARIMA(p,q).	59
6.3	Estimativas de máxima verossimilhança e erros padrão das estimativas - modelos Beta univariados para as séries de FC das usinas RF, IC e EN para o período de janeiro de 1981 à dezembro de 2010.	60
6.4	P-valores para os testes de normalidade, autocorrelação e efeitos ARCH dos resíduos quantílicos produzidos pelos modelos univariados para as séries de FC das usinas RF, IC e EN.	61
6.5	Teste da Razão de Verossimilhança (RV).	63
6.6	P-valor dos testes de aderência à distribuição uniforme das variáveis PIT.	63
6.7	Teste BDS na variável x_{1t} referente à usina RF.	64
6.8	Teste BDS na variável x_{2t} referente à usina IC.	64
6.9	Teste BDS na variável x_{3t} referente à usina EN.	64
6.10	Matriz dos coeficientes de correlação de Spearmann.	65
6.11	Matriz dos coeficientes de correlação de Kendall.	65
6.12	Verossimilhança e critérios de Informação para o modelo GAS t -Student(p,q).	68
6.13	Estimativas de máxima verossimilhança e erro padrão das estimativas - GAS t -Student(p,q).	69
6.14	Medidas de aderência das Previsão fora da amostra k -passos frente, levantadas por simulação dos modelos Beta Gas SARIMA(p,q), Beta Gas SARIMA(p,q) com GAS t -Student(p,q) e VARX.	74
6.15	CVaR das distribuições de renda vinculados a um contrato de um ano (2011) no ACL via uso dos modelos GAS. Os valores estão em mil R\$.	78
6.16	Estatísticas descritivas das densidades dos FC avaliadas para o primeiro semestre de 2011 para as usinas RF, IC e EN.	85
6.17	Estatísticas descritivas das densidades dos FC avaliadas para o segundo semestre de 2011 para as usinas RF, IC e EN.	86

1

Introdução

1.1

Motivação e Relevância do Tema

Com condições físicas e geográficas favoráveis no Brasil, um mercado de energias renováveis se estabeleceu no país. Tais fontes alternativas apresentam uma série de vantagens adicionais quando comparadas às demais fontes de produção de energia (térmicas e hidroelétricas), como (i) usinas de menor porte, (ii) facilidade de obtenção de licenças ambientais e ainda (iii) Créditos de Carbono. Como consequência da queda no preço dos equipamentos devido a crise econômica mundial de 2008-2009, a energia eólica despertou em especial a atenção dos investidores que iniciaram o desenvolvimento dessa tecnologia no país em larga escala (Bezerra *et al.*, 2010).

Em função desta crescente e rápida expansão da energia eólica na matriz energética brasileira, é imprescindível a proposição de um modelo estatístico capaz de simular cenários conjuntos que capturem e reproduzam corretamente o padrão sazonal das séries temporais de fator de capacidade eólico (FC) a partir de um número arbitrário de usinas. Estas simulações podem ser utilizadas para estudos em três frentes distintas; (i) operação do sistema sob alta inserção de renováveis; (ii) propósitos regulatórios, para o desenvolvimento de uma nova metodologia de cálculo da garantia física de parques eólicos e (iii) comercialização, na qual as simulações serão utilizadas para confeccionar estratégias de comercialização no Ambiente de Comercialização Livre de energia (ACL). Este ambiente faz parte do mercado de energia elétrica brasileira, em que os investidores adquirem contratos associados à produção das usinas eólicas via um sistema de leilões. Consequentemente, é natural em tal processo decisório quantificar o nível de exposição, a eventuais perdas monetárias, que o investidor fica exposto ao adquirir um portfólio de contratos vinculados a produção de diferentes usinas eólicas. Esta avaliação econômico-financeira é feita com base nas simulações da densidade conjunta de FC.

A literatura vigente em modelagem conjunta de FC no Brasil possui como

referencial o modelo *VARX*, introduzido por Street et al. (2012). Em resumo, é um modelo *Vector Autoregressive* (VAR) com variáveis exógenas que é aplicado para gerar cenários de FC condicionados a séries de afluições (ENAS)¹. O motivo desse condicionamento é o de produzir cenários que reproduzam o perfil de complementariedade entre o regime dos ventos e a vazão dos rios. No entanto, o *VARX* modela o vetor do log dos FC, assumindo normalidade para as variáveis. Por conseguinte, é imprescindível aplicar a função exponencial nos valores simulados, de maneira a respeitar os limites físicos de produção. Caso não seja efetuada essa transformação, é possível gerar cenários de FC negativos. Contudo, ao adotar esta estratégia, os valores dos cenários produzidos podem superar os limites físicos de produção, o que acaba prejudicando a estimação do valor em risco ao qual o investidor está exposto.

A limitação supracitada do modelo *VARX* motivou a proposição da adoção de uma nova classe de modelos de séries temporais para a modelagem da densidade conjunta de FC's, os modelos *Generalized Autoregressive Scores* (GAS). Estes modelos foram inicialmente introduzidos por Creal et al. (2008) e posteriormente detalhados com aplicações em Creal et al. (2013). Essa classe de modelos figura como alternativa para a modelagem de séries temporais não-gaussianas. Logo, serão empregados na modelagem dos modelos univariados de FC. No GAS, modela-se a densidade preditiva como um todo, possibilitando um, ou mais, dos seus parâmetros variarem no tempo. Esses parâmetros são atualizados condicionados à informação contida no score da densidade preditiva. Deste modo, exporta-se toda a estrutura da distribuição preditiva para o mecanismo de atualização do(s) parâmetro(s) e não apenas a média ou momentos de ordem superior para a atualização dos parâmetros.

No entanto como o objetivo final da dissertação é o de simular valores a partir de uma densidade conjunta de FC, de maneira a realizar uma avaliação financeira em um portfólio de contratos de diferentes usinas, é adotada a metodologia de cópulas condicionais variantes no tempo introduzida por Patton (2002). No trabalho de Creal et al. (2011), os autores expandiram os modelos GAS a um arcabouço multivariado, incluindo a possibilidade de se especifica-lo como um modelo de cópulas variante no tempo da classe elíptica. Por conseguinte, esse modelo será empregado para a modelagem da estrutura de dependência estocástica entre as séries de FC. É valoroso destacar que a proposta apresentada nessa dissertação é inédita, e se estende a outros arcabouços, tomando a densidade preditiva dos modelos marginais

¹Nessa dissertação não serão utilizadas variáveis de energia natural afluyente (ENA) como variáveis exógenas. A introdução dessas variáveis exógenas na classe de modelos que será empregada nesse trabalho é apresentada em Matos (2013).

como genérica. Porém, pela particularidade da amplitude de valores das séries de FC, será aplicada a densidade preditiva Beta assim como feito em Matos (2013). Não obstante seria possível utilizar uma outra densidade qualquer, dependendo dos valores assumidos pela série temporal em questão, fato que será abordado posteriormente.

1.2

Revisão Bibliográfica

Os autores em Creal *et al.* (2013) postularam o modelo GAS, classificando-o como um modelo guiado por observações dentro da classe de modelos com parâmetros variantes no tempo de Cox (1981). Em Harvey (2013) o autor propõe um método similar para a modelagem de parâmetros variantes no tempo que o autor denomina de *Dynamic Conditional Score*. Ainda, em Creal *et al.* (2013), os autores demonstram que muitos dos modelos com parâmetros variantes no tempo, hoje consolidados na literatura de séries temporais e econometria, são casos particulares do GAS, incluindo os modelos autoregressivos condicionalmente heterocedásticos (GARCH) de Engle e Bollerslev (1986), os modelos de erro multiplicativo (MEM) de Engle (2002B), além dos modelos com fonte única de erros, de Ord *et al.* (1997).

Ampliando a aplicação dos modelos GAS a um contexto de dados sazonais, no trabalho de Matos (2013), a metodologia supracitada foi estendida impondo uma estrutura similar aos modelos SARIMA de Box e Jenkins na atualização do parâmetro variante no tempo. Ainda em Matos (2013), foram apresentados modelos não-gaussianos e não-lineares com densidade preditiva Beta e Gama no âmbito univariado, como também foi derivada uma densidade Gama-Beta no contexto bivariado. Esse último explorando a complementariedade entre as fontes eólica e hidroelétrica, assunto relevante para o setor elétrico. Além disso, no trabalho é detalhado o uso de variáveis exógenas no ambiente do modelo GAS.

Buscando um maior detalhamento acerca da estrutura de dependência intrínseca as variáveis, Patton (2002) detalha em seu artigo como tornar essa informação mensurável em um contexto de séries temporais usando a teoria de Cópulas proposta por Sklar (1959). Tal teoria foi estendida pelo autor para um contexto de cópulas condicionais, permitindo a inclusão da modelagem da estrutura de dependência, gerando cenários mais coerentes para variáveis envolvidas, conforme comentado em Embrechts *et al.* (2005), Mendes (2004) e Cherubini *et al.* (2004).

Desta maneira, inserindo o trabalho de Patton (2002) dentro do arca-

bouço dos modelos GAS, Creal *et al.* (2011) formularam um modelo multivariado com uma densidade preditiva *t-Student*, partindo das evidências empíricas de que ambas, volatilidade e correlações se alteram ao longo do tempo. Partindo dessa premissa, os autores adotaram o mecanismo GAS para atualizar as variâncias e correlações. Além disso, o modelo contempla distintas especificações para o vetor de parâmetros variantes no tempo, como por exemplo, a atualização do logaritmo da variância, bem como as correlações podem ser temporalmente atualizadas baseadas em coordenadas hiperesféricas. Outra característica interessante do modelo é a flexibilidade de trabalhar com fatores, estrutura muito útil ao aplicar o modelo em um arcabouço onde se faz necessária a redução de dimensão de dados.

1.3

Objetivos e Contribuições

A energia eólica é uma fonte com um perfil intermitente não controlável, ao contrário das fontes hídricas e térmicas convencionais, que podem ser utilizadas pelo operador do sistema elétrico como um recurso controlável para atender a demanda ao longo do dia. Partindo dessa premissa, o objetivo dessa dissertação é o de proporcionar um modelo estatístico que figure como alternativa ao modelo *VARX*, fornecendo previsões e simulações contemplando aspectos adicionais da modelagem como a questão da não normalidade, sazonalidade e modelagem da estrutura de dependência das séries de FC. Tornando possível a elaboração de estratégias tanto no âmbito da operação do sistema elétrico como no da avaliação econômico-financeira de empreendimentos de geração.

De maneira a considerar a modelagem da estrutura de dependência estocástica contida entre as séries temporais envolvidas no problema, o estudo foi dividido em dois estágios. No primeiro é aplicado o modelo GAS univariado, derivado em Matos (2013) com densidade preditiva Beta nas séries dos FC de cada usina eólica. Posteriormente, para o diagnóstico desse modelo não-gaussiano, aplicam-se os resíduos quantílicos descritos em Kalliovirta (2009). Após a transformação dos dados marginais em Uniformes $[0,1]$, modela-se a estrutura de dependência multivariada intrínseca entre as séries de FC via o modelo de cópula GAS *t-Student* postulado em Creal *et al.* (2011).

Nossa contribuição para a literatura de modelos GAS é dupla, a saber: (i) estendemos a aplicação do modelo multivariado GAS *t-Student* a um arcabouço de dados sazonais. (ii) Implementamos a estrutura de cópula variante no tempo para construir uma densidade conjunta a partir dos modelos univariados. Onde tanto o parâmetro de forma das densidades preditivas Beta, quanto

os elementos da matriz de correlação das variáveis formadas a partir da explicação das variáveis marginais (pseudo-observações), são variantes no tempo e atualizadas via mecanismo GAS.

Além das contribuições teóricas dos modelos GAS, é importante destacar a contribuição prática desse trabalho para o setor elétrico. A partir da metodologia proposta, será possível fornecer um modelo alternativo ao *VARX*, de maneira a gerar uma distribuição conjunta de FC que produza cenários que serão introduzidos em estudos de operação, comercialização e regulamentação.

1.4

Organização

O presente capítulo configura o problema a ser resolvido, motiva o leitor mostrando a sua relevância e caracteriza a inovação através de busca na literatura científica por trabalhos correlatos.

O Capítulo 2 apresenta o modelo GAS, suas principais propriedades, previsão, reparametrização e estimação via máxima verossimilhança dos parâmetros estáticos.

O Capítulo 3 refere-se à metodologia proposta por Matos (2013), especificações do modelo, como introduzir uma dinâmica sazonal e a especificação de valores iniciais, necessários para a otimização dos parâmetros.

O Capítulo 4 se propõem a passar para o leitor uma noção básica de cópulas condicionais.

O Capítulo 5 apresenta o modelo de Creal *et al.* (2011) que será aplicado em uma especificação de cópulas variantes no tempo.

O Capítulo 6 apresenta um estudo de caso onde é aplicada a metodologia desenvolvida durante os capítulos anteriores.

O Capítulo 7 apresenta as conclusões obtidas com o uso da metodologia proposta.

2

Generalized Autoregressive Score Models

2.1

Introdução

A família de modelo GAS, é classificada como pertencente aos “modelos guiados por observações”, o qual permite aos parâmetros da distribuição condicional (ou preditiva), tais como média condicional, variância condicional, dentre outros, a possibilidade de variarem no tempo. A atualização desses parâmetros considera informações da densidade preditiva via o uso da função *score*.

O modelo GAS é definido da seguinte forma. Considere um modelo arbitrário $p(y_t; \psi)$, pertinente ao conjunto de observações $\{y_1, y_2, \dots, y_T\}$, onde ψ é o vetor de parâmetros estáticos desse modelo. Em um contexto clássico de séries temporais, a função de verossimilhança é avaliada da seguinte forma,

$$p(y_t; \psi) = p(y_1; \psi) \prod_{t=2}^n p(y_t | y_1, \dots, y_{t-1}; \psi).$$

Já no arcabouço dos modelos GAS, delimita-se um novo subconjunto do vetor ψ variante no tempo, i.e., $\psi_t = (f_t; \theta)$. Onde f_t representa o vetor de parâmetros variantes no tempo e θ um vetor de parâmetros estáticos estimados via máxima verossimilhança que são responsáveis pela atualização de f_t .

Desta forma, a ideia dessa metodologia segundo Koopman *et al.* (2012), é a de modelar o vetor de parâmetros variantes no tempo da densidade preditiva, de uma forma autoregressiva tal como

$$f_{t+1} = \omega + A \cdot \text{alguma inovação} + Bf_t \quad (2-1)$$

no qual a t-ésima contribuição para a verossimilhança seria contabilizada por

$$\ell_t = \ln p(y_t | y_1, \dots, y_{t-1}, f_1, \dots, f_t; \theta) \quad (2-2)$$

onde assumem-se que os valores de $f_1, \dots, f_t$ já foram realizados e por consequên-

cia são conhecidos.

O conceito apresentado pelos autores de Creal *et al.* (2013) é a de utilizar o *score* da verossimilhança com relação a f_t como inovação para a atualização dos parâmetros variantes no tempo. Ao conjecturar f_t dessa maneira, define-se um algoritmo recursivo para a estimação dos parâmetros variantes no tempo.

Ainda em Creal *et al.* (2013), algumas explicações ao uso do *score* como inovação para a atualização autoregressiva de f_t em (2-1) são apresentadas. Uma delas é que o *score* define a direção ascendente mais íngreme para aprimorar o ajuste local do modelo em termos da verossimilhança no tempo t , condicionada à atual posição do parâmetro f_t , o que acaba fornecendo uma direção mais natural para a atualização dos parâmetros. Outra característica relevante desse mecanismo de atualização está no fato de que ao realizar a atualização do vetor de parâmetros variantes no tempo dessa forma, contempla-se toda a estrutura de informação intrínseca na densidade preditiva das observações e não apenas momentos de primeira e segunda ordem, uma vez que o *score* é função da densidade completa dos dados.

2.2

Especificações do modelo e propriedades

O modelo GAS é definido em Creal *et al.* (2013) da seguinte maneira. Denota-se y_t como um vetor $N \times 1$ referente às variáveis dependentes de interesse, f_t como o vetor dos parâmetros variantes no tempo, x_t um vetor de variáveis exógenas, em todo o tempo t , e θ um vetor de parâmetros estáticos. Defina $Y^t = \{y_1, \dots, y_t\}$, $F^t = \{f_0, \dots, f_t\}$, e $X^t = \{x_0, \dots, x_t\}$. O conjunto de informação disponível em t consiste de $\{f_t, \mathcal{F}_t\}$ onde

$$\mathcal{F}_t = \{Y^{t-1}, F^{t-1}, X^t\}, \text{ para } t = 1, \dots, n.$$

Sendo que y_t é assumido ser gerado a partir da seguinte densidade preditiva

$$y_t \sim p(y_t | f_t, \mathcal{F}_t, \theta). \quad (2-3)$$

O vetor dos parâmetros variantes no tempo da densidade preditiva do modelo é atualizado pela seguinte equação de atualização autoregressiva

$$f_{t+1} = \omega + \sum_{i=1}^p A_i s_{t-i+1} + \sum_{j=1}^q B_j f_{t-j+1}, \quad (2-4)$$

onde ω é um vetor de constantes; A_i e B_j são matrizes que possuem dimensões apropriadas para $i = 1, \dots, p$ e $j = 1, \dots, q$, enquanto o *score* s_t é representado

por uma função apropriada dos dados passados, $s_t = (y_t, f_t, \mathcal{F}_t; \theta)$. Os coeficientes desconhecidos em (2-4) são funções de θ , isto é $\omega = \omega(\theta)$, $A_i = A_i(\theta)$ e $B_i = B_i(\theta)$ para $i = 1, \dots, p$ e $j = 1, \dots, q$. Onde os coeficientes das matrizes $B_1, \dots, B_q$ determinam a persistência do vetor f_t ao longo do tempo, i.e, contém os parâmetros autoregressivos.

Esquemáticamente, o mecanismo de atualização apresentado em (2-4) funciona da seguinte forma. Ao realizar uma nova observação y_t , o parâmetro variante no tempo f_t é atualizado para o próximo período $t + 1$ baseado na equação (2-4) com

$$s_t = S_t \cdot \nabla_t, \quad \nabla_t = \frac{\partial \ln p(y_t | f_t, \mathcal{F}_t; \theta)}{\partial f_t}, \quad S_t = S(t, f_t, \mathcal{F}_t; \theta), \quad (2-5)$$

onde $S(\cdot)$ possui dimensões apropriadas e é denominada de matriz de ponderação. Dessa forma, a atualização de f_t em (2-4) fica em função do vetor *score*(2-5) e do próprio vetor de parâmetros f_t . Em resumo, para uma densidade arbitrária, as equações (2-3)-(2-5) definem o modelo GAS de ordens p e q , que denotaremos nesse trabalho como GAS(p,q).

É importante salientar que, via a escolha da matriz de ponderação S_t , um nível adicional de flexibilidade é introduzido ao modelo. Isto porque, diferentes especificações de S_t resultam em modelos GAS totalmente distintos. Os autores em Creal *et al.* (2013) apontam para o inverso da informação de Fisher como uma especificação intuitiva dessa matriz, isto é

$$S_t = \mathcal{I}_{t|t-1}^{-1}, \quad \mathcal{I}_{t|t-1} = -E_{t-1} \left[\frac{\partial^2 \ln p(y_t | f_t, \mathcal{F}_t; \theta)}{\partial^2 f_t f_t'} \right], \quad (2-6)$$

tal que $Var(s_t) = \mathcal{I}_{t|t-1}^{-1}$. Ao especificar S_t dessa forma, a recursão (2-4) pode ser interpretada como um algoritmo de Gauss-Newton para a estimação de f_t ao longo do tempo.

Além disso, os autores demonstram que o arcabouço dos modelos GAS baseado no inverso da matriz de informação de Fisher, contém como casos particulares muitos outros modelos já consolidados na literatura de séries temporais, incluindo o modelo autoregressivo de duração condicional (ACD) para a distribuição exponencial de Engle e Russel (1998) e o modelo de erro multiplicativo para a distribuição gama de Engle e Gallo (2006).

Vale ressaltar que, pelo fato de ∇_t ser o *score* da densidade com relação a f_t , segue imediatamente que $E_{t-1}[s_t] = 0$ e s_t forma uma sequência de diferenças martingais, e assim não há previsibilidade linear ou não linear na sua média condicional. O processo pode ser reescrito como um processo de

diferenças martingais de médias móveis finito. No caso do modelo GAS(1,1), tem-se

$$f_t = \omega + As_{t-1} + Bf_{t-1} \quad (2-7)$$

$$(1 - BL)f_t = \omega + As_{t-1} \quad (2-8)$$

$$f_t = \Psi(B)(\omega + As_{t-1}), \text{ onde } \Psi(B) = \frac{1}{1 - BL} \quad (2-9)$$

$$\text{como } \frac{1}{1 - BL} = \sum_{i=1}^{\infty} B^i L^i \text{ e supondo } |B| < 1 \quad (2-10)$$

$$f_t = \omega \sum_{i=1}^{\infty} B^i + \sum_{i=1}^{\infty} B^{i-1} As_{t-i}. \quad (2-11)$$

Se a variância condicional de s_t é constante ao longo do tempo, então o processo f_t é estacionário na covariância, quando as raízes de B estão dentro do círculo unitário. Porém geralmente, com o inverso da informação de Fisher como matriz de ponderação, i.e., $S_t = \mathcal{I}_{t|t-1}^{-1}$ a variância condicional não é constante e formular condições de estacionariedade se tornam difíceis. Em Creal *et al.* (2013) os autores resolveram esse fato usando como matriz de ponderação $\mathcal{I}_{t|t-1}^{-1/2}$. Porém, visando ficar próximo ao arcabouço do modelo GARCH multivariado, os autores mantiveram $S_t = \mathcal{I}_{t|t-1}^{-1}$, estratégia essa que será mantida nessa dissertação.

2.3

Parametrização

Uma propriedade interessante dos modelos GAS está na maneira natural que esse se adapta a diferentes parametrizações do vetor f_t . A resposta a uma transformação neste é direta, i.e., assumindo a parametrização $\tilde{f}_t = h(f_t)$ preferível a f_t , supondo uma função arbitrária $h(\cdot)$ estritamente crescente, o vetor de parâmetros variantes no tempo se altera. Essa transformação é útil quando o parâmetro de interesse da distribuição preditiva está sujeito a algum tipo de restrição, como por exemplo o coeficiente de correlação que está entre (-1,1), ou a variância condicional que só pode assumir valores positivos. Ao reparametrizar o problema, o modelo opera na Equação (2-4) com o vetor transformado de parâmetros irrestritos, $\tilde{f}_t$.

Com respeito ao vetor *score* de $\tilde{f}_t$, assuma que $\dot{h}_t = \partial h(f_t) / f'_t$, e conforme ressaltado em Creal *et al.* (2013), através da aplicação da regra da cadeia é possível demonstrar que,

$$\tilde{\nabla}_t = \left(\dot{h}_t \right)^{-1} \times \nabla_t. \quad (2-12)$$

Já a matriz de Informação de Fisher, caso for escolhida a sua inversa como matriz de ponderação do *score*, o esquema de atualização do parâmetro variante no tempo, o *score* s_t , se transforma em $\tilde{s}_t = \dot{h}_t s_t$. Ao passo que se for utilizada a raiz quadrada da (pseudo)-inversa da matriz de Informação de Fisher ao invés da anterior, não haveria mudanças no *score*, i.e., $\tilde{s}_t = s_t$ (Creal *et al.*, 2013).

Será contextualizada essa questão com um exemplo retirado de Creal *et al.* (2013), o qual supõe um modelo GARCH com inovações normais. Afim de estabelecer que f_t assuma apenas valores positivos, ao invés de definir $f_t = \sigma_t^2$, o parâmetro variante no tempo é reparametrizado como $\tilde{f}_t = \log(\sigma_t^2)$. Dessa forma, o modelo deve ser expressado por $y_t = \exp(\tilde{f}_t)\epsilon_t$, e a nova recursão para $\tilde{f}_t$ se torna

$$f_{t+1} = \omega + A_1[\exp(-\tilde{f}_t)y_t^2 - 1] + B_1\tilde{f}_t. \quad (2-13)$$

2.4

Estimação por Máxima Verossimilhança

Uma das principais vantagens ao optar pela aplicação de modelos guiados por observações está na facilidade da avaliação da verossimilhança, visto que nesse arcabouço a mesma possuirá forma analítica. Supondo uma série temporal $y_1, \dots, y_t$, o estimador de máxima verossimilhança pode ser expresso por

$$\hat{\theta} = \underset{\theta}{argmax} \sum_{t=1}^n \ell_t, \quad (2-14)$$

onde $\ell_t = \ln p(y_t | f_t, \mathcal{F}_t, \theta)$ para uma dada realização de y_t .

Dessa forma, a estimação do vetor de parâmetros estáticos dos modelos GAS é relativamente simples. Para tal, basta a implementação do mecanismo de atualização apresentado pela Equação (2-4). Uma vez implementada, basta avaliar o logaritmo da verossimilhança ℓ_t para um particular valor θ^* de θ .

Para avaliar a significância dos parâmetros estimados pelo modelo, é necessário o computo dos erros padrões e as estatísticas t com base na inversa da matriz Hessiana do logaritmo da verossimilhança avaliada no ótimo encontrado (Creal *et al.*, 2013). Isto é, sob certas condições de regularidade, o estimador de máxima verossimilhança $\hat{\theta}$ de θ é consistente e converge em distribuição para

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, H^{-1}), \quad (2-15)$$

onde $H = \lim_{n \rightarrow \infty} E[(\partial l / \partial \theta)(\partial l / \partial \theta')] / n$ e $\ell = \sum_{t=1}^n \ell_t$.

Para provar tal resultado, os autores em Creal *et al.* (2013) se utilizam de resultados indiretos, uma vez que obter tal prova é complexa para os modelos GAS. Tais resultados dissimulados seriam obtidos via os modelos aos quais o modelo GAS se reduz. Por exemplo, para provar a consistência e a normalidade assintótica do estimador de máxima verossimilhança do modelo GAS com densidade preditiva de Poisson, os autores se utilizam de um resultado já provado para os modelos de contagem de Poisson, apresentado em Davis *et al.* (2003).

3

Modelo Beta SARIMA

Outro aspecto interessante dos modelos GAS está na possibilidade de propor dinâmicas distintas para a atualização do vetor de parâmetros variantes no tempo na Equação (2-4). É possível, por exemplo generalizar a evolução de f_t incluindo outros efeitos não lineares tais como mudança de regime ou até mesmo a inclusão de efeitos de memória longa similar a postulada pelos modelos ARFIMA ou FIGARCH (Creal *et al.*, 2013). Nesse capítulo, será apresentado o modelo proposto em Matos (2013) que introduz uma dinâmica sazonal para a atualização de f_t .

3.1

Especificação

Uma propriedade inerente ao uso dos modelos GAS está na escolha da densidade preditiva a ser empregada. Sabendo que os fatores de capacidade (FC) das usinas eólicas assumem valores entre $[0, 100]$, é obrigatória a especificação de uma densidade preditiva que acomode em seu suporte esses valores. Pensando nisso, o autor de Matos (2013) ajustou uma distribuição Beta com suporte entre $[0, k]$ afim de modelar essas variáveis. Tal distribuição é uma modificação da distribuição Beta com a seguinte forma,

$$Y \sim \text{Beta}(\beta, \alpha), \quad y \in [0, k], \quad (3-1)$$

com a função densidade de Y expressa por

$$p(y_t|\beta, \alpha) = \frac{1}{k^{\beta+\alpha-1}} \frac{\Gamma(\beta+\alpha)}{\Gamma(\beta)\Gamma(\alpha)} y_t^{\beta-1} (k-y_t)^{\alpha-1}, \quad y_t \in [0, k], \quad \beta, \alpha > 0 \quad (3-2)$$

cujos dois primeiros momentos condicionais são dados por

$$E(y_t) = k \frac{\beta}{\beta + \alpha} \quad e \quad V(y_t) = k^2 \frac{\beta\alpha}{(\beta + \alpha)^2(\beta + \alpha + 1)}. \quad (3-3)$$

Ao aplicar o logaritmo na função de verossimilhança afim de obter os elementos da matriz de informação de Fisher, assumindo ambos parâmetros

de forma da Beta variantes no tempo teremos

$$\ln p(y|\beta, \alpha) = -\ln(k)(\beta + \alpha - 1) + \ln \Gamma(\beta + \alpha) - \ln \Gamma(\beta) \\ - \ln \Gamma(\alpha) + \beta \ln(y) - \ln(y) + \alpha(k - y) - \ln(k - y). \quad (3-4)$$

com as derivadas parciais em relação aos parâmetros da densidade preditiva Beta expressas por

$$\frac{\partial \ln p(y|\beta, \alpha)}{\partial \beta} = \ln y - [\ln k + \psi_1(\beta) - \psi_1(\beta + \alpha)] \quad (3-5)$$

$$\frac{\partial \ln p(y|\beta, \alpha)}{\partial \alpha} = \ln(k - y) - [\ln k + \psi_1(\alpha) - \psi_1(\beta + \alpha)] \quad (3-6)$$

onde ψ_i representa a i-ésima derivada do logaritmo da função gama expressa por $\psi_i(x) = \frac{d}{dx} \ln \Gamma(x)$.

Ao final, ao aplicarmos a derivada segunda no logaritmo da verossimilhança com relação aos parâmetros da Beta, obtém-se os elementos da matriz de informação de Fisher,

$$\mathcal{I}(\beta, \alpha) = \begin{bmatrix} \psi_2(\beta) - \psi_2(\beta + \alpha) & -\psi_2(\beta + \alpha) \\ -\psi_2(\beta + \alpha) & \psi_2(\alpha) - \psi_2(\beta + \alpha) \end{bmatrix}. \quad (3-7)$$

3.2

Modelo Beta univariado

À semelhança do que foi proposto em Matos (2013), serão derivados modelos GAS com densidade preditiva Beta, introduzindo uma dinâmica sazonal no processo de atualização de f_t , similar à proposta por *Box e Jenkins*, denominada de SARIMA, para modelar a média condicional. Para isso, basta tomar os coeficientes de A_i e B_j não nulos para múltiplos de 12 e em suas proximidades. Essa ideia será adotada em seguida para a modelagem das séries de fator de capacidade eólico, em consequência da sazonalidade presente nos dados dos fatores de capacidade.

Ao utilizar tal dinâmica, o primeiro ponto a gerar discussão fica em torno do número de parâmetros do modelo. Primando pela parcimônia deste e além disso, visando evitar possíveis problemas na otimização dos parâmetros, o arcabouço do modelo em questão contará apenas com um dos parâmetros da densidade preditiva variante no tempo. Isso porque, conforme colocado por Matos (2013), em um caso com os dois parâmetros variando no tempo, o número de parâmetros a ser estimado aumentaria muito. Logo, optou-se pela

especificação variante no tempo $f_t = \beta_t$, fazendo com que qualquer momento condicional dessa densidade preditiva seja variante no tempo, por construção.

Contudo, como o parâmetro de forma da densidade preditiva Beta assume apenas valores positivos, $\beta_t > 0$, é necessário a aplicação da metodologia de parametrização apresentada na Sessão 2.3. Para tal, aplica-se uma função $h(\cdot)$ que restringe o parâmetro β_t a assumir apenas valores positivos. Para satisfazer tal propriedade, análogo ao exemplo da reparametrização da variância condicional nos modelos GARCH, a função logarítmica é julgada mais adequada, logo $\tilde{f}_t = h(\beta)$.

Dessa forma, a nova especificação do modelo GAS em questão com $f_t = \ln(\beta_t)$, seria

$$\begin{aligned} \dot{h}_t &= \frac{1}{\beta_t} \\ \tilde{\nabla}_t &= \beta_t \{ \ln y_t - [\ln k + \psi_1(\beta_t) - \psi_1(\beta_t + \alpha)] \} \end{aligned} \quad (3-8)$$

$$\begin{aligned} \tilde{\mathcal{I}}_{t|t-1} &= \{ \beta_t^2 [\psi_2(\beta_t + \alpha) - \psi_2(\beta_t)] \}^{-1} \\ \tilde{s}_t &= \frac{\ln y_t - [\ln k + \psi_1(\beta_t) - \psi_1(\beta_t + \alpha)]}{\beta_t^{1-2d} [\psi_2(\beta_t) - \psi_2(\beta_t + \alpha)]^{1-d}}, \quad d = 0, 1/2 \text{ ou } 1 \end{aligned} \quad (3-9)$$

onde o *score* com $d = 0$ é referente ao uso da inversa da matriz de informação de Fisher como esquema de ponderação, $d = 1/2$ é referente à raiz quadrada da inversa da informação de Fisher e $d = 1$ a matriz identidade.

3.2.1

Condições Iniciais

Para implementar a estrutura de dinâmica sazonal em f_t , é fundamental obter um conjunto de valores iniciais expressos por $f_{0,1-q} = \{f_0, f_{-1}, \dots, f_{1-q}\}$ para iniciar a otimização. À semelhança do que foi proposto em Matos (2013), tais parâmetros serão estimados a partir da segmentação da série original em 12 séries, onde serão separados os valores mensais, i.e., $\{y_{jan}, y_{fev}, \dots, y_{dez}\}$. A partir disso, estima-se via máxima verossimilhança os parâmetros de forma da distribuição Beta referente a cada uma das 12 séries, observando a reparametrização imposta em $f_t = \log(\beta_t)$, obtendo o conjunto $f_{0,1-q}$.

A partir dessa informação, será possível calcular $s_{0,1-p} = \{s_0, s_{-1}, \dots, s_{1-p}\}$, em razão de que, por construção o *score* é função de f_t . Para isso serão retiradas as 12 primeiras observações da série, formando

o vetor $Y_{inicial}$ e em posse deste e mais $f_{0,1-q}$, obtém-se $s_{0,1-p}$. Vale ressaltar que a verossimilhança e os resíduos serão computados apenas a partir da 13ª observação.

Quanto às condições iniciais dos parâmetros necessárias para dar início ao procedimento de busca do ótimo pelo algoritmo de quasi Newton, são gerados 20 conjuntos de condições iniciais. Esses conjuntos contêm as seguintes informações. Para os parâmetros ω , A_i 's e B_j 's geram-se valores aleatórios a partir de uma distribuição uniforme com suporte entre $[0,0.01]$. Já para o parâmetro α , a primeira coisa a ser feita é estimar os parâmetros de uma distribuição Beta para a série inteira via máxima verossimilhança. Após a estimação do parâmetro, gera-se um valor aleatório a partir de uma distribuição normal com média igual a $\hat{\alpha}_{MV}$, estimado anteriormente, e desvio padrão 0,5.

Com relação a otimização dos parâmetros, utilizamos a mesma heurística do autor apresentada em Matos (2013). A partir dos conjuntos de condições iniciais gerados, inicia a otimização usando o método de *Nelder-Mead*, que se apresentou mais rigoroso nas aplicações empíricas. Os valores dos 12 parâmetros resultantes dessa otimização servirão como condições iniciais para a otimização pelo método BFGS, e esses posteriormente para o método de *Nelder-Mead* e assim sucessivamente, até a diferença (*gap*) entre as soluções ser menor que um valor de tolerância postulado¹.

3.3

Diagnósticos

O diagnóstico para avaliar se o modelo GAS captura a estrutura de dependência intrínseca nos dados é baseado na análise dos testes para os resíduos quantílicos proposta em Kalliovirta (2009). Tais resíduos são julgados adequados para esse caso, uma vez que estão inseridas no arcabouço dos modelos GAS questões como não linearidade e não normalidade. De acordo com Dunn e Smyth (1996), o resíduo quantílico teórico é definido como

$$R_{t,\theta} = \Phi^{-1}(F(y_t|f_t, \mathcal{F}_t, \theta)) \quad (3-10)$$

e o resíduo quantílico observado é expresso por

$$r_{t,\hat{\theta}} = \Phi^{-1}(F(y_t|f_t, \mathcal{F}_t, \hat{\theta})) \quad (3-11)$$

onde $\Phi^{-1}(\cdot)$ é a acumulada da função de distribuição (F.D.) de uma normal padrão, $F(y_t|f_t, \mathcal{F}_t, \theta)$ é a F.D. condicional do processo y_t conforme (2-3), e $\hat{\theta}$

¹Atribuímos $gap = 0.1$

é uma estimativa de θ .

Em Kalliovirta (2009), são propostos testes para avaliar a questão da normalidade, heterocedasticidade condicional e independência serial para esse tipo de resíduo. Mas como apontado em Matos (2013), tais testes são difíceis de serem especificados para o arcabouço em questão, pois o cômputo de relações recursivas para o cálculo de derivadas em relação ao vetor θ não é direto. Sendo assim, pela dificuldade mencionada, os testes convencionais serão utilizados, i.e., Jarque-Bera para a avaliação da normalidade dos resíduos e Ljung-Box para autocorrelação serial e heterocedasticidade condicional.

3.3.1

Propriedades dos resíduos quantílicos

Nas aplicações do Capítulo 6, esses resíduos apresentaram conclusões similares às obtidas com os resíduos de Pearson. Tal conclusão pode ser explicada pelo fato de que, em se tratando de variáveis aleatórias *i.i.d.* em um ambiente Gaussiano, o resíduo quantílico se reduz ao de Pearson. Essa conclusão pode ser observada ao se utilizar ou um modelo linear padrão, ou um modelo autoregressivo linear ou não linear com erros normalmente distribuídos. Nesse sentido, os resíduos quantílicos são uma generalização natural dos resíduos de Pearson.

Nas aplicações que serão apresentadas, houve rejeição da hipótese de normalidade, porém um estudo de simulação foi conduzido com o objetivo de buscar possíveis problemas para essa distorção. Basicamente, como o tamanho de amostra da aplicação é limitada, $T = 360$, suspeitou-se que a normalidade desses resíduos seria rejeitada pelo fato de ser uma propriedade assintótica, não se sustentando para pequenos tamanhos de amostra.

Dessa forma, o estudo foi conduzido da seguinte maneira. Inicialmente a partir de um modelo Beta GAS(1,1) com suporte entre $[0, k]$, geram-se 5 processos para cada tamanho $T = \{100, 500, 1000\}$, fixando os parâmetros $A = -0.5$, $B = 0.72$, $\omega = 0.01$, $\alpha = 7.04$ e $f_0 = 5$. Em posse das séries geradas $\left[\{y_t^{(i)}\}_{t=1}^T \right]_{i=1}^5$, estimam-se os parâmetros de θ via máxima verossimilhança, incluindo o parâmetro f_0 , utilizando o mesmo modelo Beta GAS(1,1). Os resíduos quantílicos produzidos, tanto pela geração quanto pela otimização foram avaliados, não rejeitando em nenhum dos casos a normalidade e independência serial via Jarque Bera e Ljung Box. À vista disso, infere-se que o modelo estava bem especificado nas aplicações do Capítulo 6, e que a rejeição da normalidade dos resíduos quantílicos é devida a algumas observações mal ajustadas, o que posteriormente foi confirmado.

3.4

Previsão fora da amostra

Com relação à previsão fora da amostra dos modelos GAS, essa partilha de algumas similaridades com seu análogo, o modelo GARCH. Em se tratando de um modelo não linear e não gaussiano, as previsões de $\hat{y}_{t+k|t}$, para $k \geq 2$ são obtidas por simulação, posto que a forma analítica da distribuição de previsão para mais de um passo à frente não é conhecida. Para exemplificar tal fato, considere a avaliação da distribuição condicional 2 passos à frente,

$$p(y_{t+2}|f_t, \mathcal{F}_t, \theta) = \int_0^\infty p(y_{t+2}|f_{t+1}, \mathcal{F}_{t+1}, \theta)p(y_{t+1}|f_t, \mathcal{F}_t, \theta)dy_{t+1}. \quad (3-12)$$

Embora a distribuição condicional de $p(y_{t+1}|f_t, \mathcal{F}_t, \theta)$ seja conhecida, por especificação não há garantias de que a distribuição condicional oriunda do produto entre $p(y_{t+2}|f_{t+1}, \mathcal{F}_{t+1}, \theta)p(y_{t+1}|f_t, \mathcal{F}_t, \theta)$ também o será. Resultado disso é que a integral expressa pela Equação (3-12) não terá solução analítica, necessitando de simulação Monte Carlo para sua solução, sendo produzida conforme o Algoritmo 1 apresentado em seguida. No algoritmo, p e q , são as respectivas ordens selecionadas para o modelo GAS(p,q).

Dados: $f_{t,q} = \{f_t, f_{t-1}, \dots, f_{t-q}\}$, $s_{t,p} = \{s_t, s_{t-1}, \dots, s_{t-p}\}$, $\hat{\theta}$

Entrada: Gere m trajetórias k -passos à frente.

```

1   $k \leftarrow 1$ ;
2   $m \leftarrow 1$ ;
3  para  $m \leftarrow 1$  até  $m \leftarrow 2000$  faça
4       $k \leftarrow 1$ ;
5      para  $k \leftarrow 1$  até  $k = K$  faça
6          Com base em  $\hat{\theta}$ , gere 1 valor da densidade preditiva de  $\hat{y}_{t+k|t}^{(m)}$ ;
7          Calcule  $f_{t+k-1,q}$  e  $s_{t+k-1,p}$ ;
8          Atualize  $f_{t+k,q}$  e  $s_{t+k,p}$ ;
9      fim
10     se  $m=1000$  e  $k=K$  então
11          $\hat{y}_{t+k|t}$  é obtido pela média das  $m$  trajetórias calculadas a cada
            instante de tempo  $k$ ;
12         Os respectivos intervalos de confiança são obtidos pelos
            quantis  $(\alpha/2)$  e  $(1 - \alpha/2)$  destas  $m$  trajetórias
13     senão
14          $m=m+1$ ;
15     fim
16 fim
```

Algoritmo 1: Previsão k -passos à frente do modelo GAS via simulação.

4

Copulas e Dependência

Nesse capítulo será conceituada a parte referente a modelagem da estrutura de dependência de dados multivariados usando o conceito de cópulas. Tal conceito se materializa relevante devido ao objetivo dessa dissertação de efetuar a modelagem conjunta dos FC das usinas eólicas em questão no Capítulo 6.

4.1

Introdução à teoria de cópulas

A base teórica da teoria de cópulas é fundamentada pelo Teorema de Sklar (1959). Em seu trabalho, Sklar postula a cópula como uma metodologia aplicável para unir as funções de distribuição marginais, formando uma função de distribuição conjunta. Visando a simplicidade de notação e maior facilidade de compreensão, ilustraremos abaixo o caso bivariado.

Teorema de Sklar Seja F uma função densidade conjunta em $\mathbb{R}^2$ com marginais $F_i, i = 1, 2$. Então existe uma função $C_F : \mathbb{R}^2 \rightarrow \mathbb{R}$, denominada cópula, tal que $\forall x = (x_1, x_2) \in \mathbb{R}^2$:

$$F(x_1, x_2) = C_F(F_1(x_1), F_2(x_2)). \quad (4-1)$$

Quando as F_i são contínuas, C_F é única, já quando F não é contínua, apesar de existir a representação de F pela cópula, esta não é única.

A equação (4-1) pode ser invertida da seguinte maneira. Seja F a função de distribuição conjunta contínua do vetor aleatório $\mathbf{X} = (X_1, X_2)$. Sejam F_1, F_2 as distribuições marginais, e definimos a transformação integral de probabilidade (PIT) para a distribuição *Uniforme*(0, 1) como $F_i(X_i) = U_i, i = 1, 2$. A cópula pertinente a F é definida como

$$C_F(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)). \quad (4-2)$$

Para prosseguir com a definição da função densidade gerada a partir da cópula, é necessária a compreensão do seguinte teorema:

Teorema 2.2.7 de Nelsen (1999) Seja C uma cópula. Para todo u_2 em $\mathbf{I} = [0, 1]$, a derivada parcial $\partial C / \partial u_1$ existe para quase todo u_1 , e para tais u_2 e u_1 ,

$$0 \leq \frac{\partial C(u_1, u_2)}{\partial u_1} \leq 1. \quad (4-3)$$

Similarmente, para todo u_1 em $\mathbf{I}$, a derivada parcial $\partial C / \partial u_2$ existe para quase todo u_2 , e para tais u_1 e u_2 ,

$$0 \leq \frac{\partial C(u_1, u_2)}{\partial u_2} \leq 1.$$

Concluindo, com base nos resultados da Equação (4-1) e do teorema acima, obtemos a densidade multivariada de $\mathbf{X}$ via

$$f(x_1, x_2) = c(F_1(x_1), F_2(x_2)) \times \prod_{i=1}^2 f_i(x_i) \quad (4-4)$$

onde

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}.$$

Assim, pela Equação (4-4) fica evidente que a densidade conjunta de $f(x_1, x_2)$ é composta de duas partes, uma que descreve a estrutura de dependência e outra que descreve o comportamento marginal de cada variável. Logo, a cópula pode ser interpretada como a estrutura de dependência pertinente ao do vetor aleatório $\mathbf{X}$.

Ao contrário das limitações apresentadas pelo coeficiente de correlação linear de Pearson em Embrechts *et al.* (2005), a cópula captura diversos tipos de dependência, incluindo de cauda, entre as variáveis aleatórias, sendo invariante a transformações monótonas crescentes, como por exemplo a transformação logarítmica já mencionada. A prova é demonstrada no Apêndice A. Além disso, a cópula é indiferente ao suporte das distribuições marginais, por exemplo, é possível definir uma estrutura de dependência entre uma distribuição Beta com suporte $[0, 1]$ e uma Normal que assume valores entre $[-\infty, +\infty]$.

Algumas cópulas são de grande importância teórica e prática, como por exemplo as cópulas produto, onde a distribuição conjunta de duas variáveis aleatórias independentes é representada pelo produto das marginas,

$C_{\perp}(u_1, u_2) = u_1 \times u_2$, e as de Frechét que relacionam algum tipo de dependência perfeita. As cópulas Arquimedianas e outros modelos de cópulas também são de grande importância teórica, porém não serão aqui abordadas pois não são o foco deste trabalho. Outras propriedades de Cópulas como ordenação e permutação também podem ser encontradas em Mendes (2004) e Embrechts *et al.* (2005).

4.2

Cópulas Condicionais

A extensão da aplicação da metodologia de cópulas para o arcabouço de séries temporais possui duas vertentes distintas (Patton, 2009). A primeira é empregada em séries temporais multivariadas, a qual foca na modelagem da distribuição conjunta do vetor aleatório $\mathbf{Y}_t = [Y_{1t}, Y_{2t}, \dots, Y_{nt}]'$, condicional a algum conjunto de informação $\mathcal{F}_t$. Já na segunda aplicação em séries temporais, estima-se uma cópula a partir da sequência de observações de uma série temporal univariada. Pode-se, por exemplo, modelar a distribuição conjunta de $[Y_t, Y_{t+1}, \dots, Y_{t+n}]'$. Esta última não será discutida nesse trabalho, uma vez que com base no objetivo da análise do Capítulo 6, será implementada a primeira vertente da metodologia para séries temporais multivariadas.

4.2.1

Modelos de cópulas para séries temporais multivariadas

As aplicações de cópulas em séries temporais multivariadas são realizadas estendendo o teorema de Sklar à distribuição condicional de $\mathbf{Y}_t$, condicionada ao conjunto de informações passadas $\mathcal{F}_{t-1}$.

Em Patton (2006), o autor define uma cópula condicional como uma distribuição multivariada de variáveis (possivelmente correlacionadas) que possuem distribuição *Uniforme* $(0, 1)$ condicionadas à $\mathcal{F}_{t-1}$. Com essa definição, é possível estender o teorema de Sklar para o caso de processos estocásticos como

$$F_t(\mathbf{y}_t | \mathcal{F}_{t-1}) = C_t(F_{1,t}(y_{1t} | \mathcal{F}_{t-1}), F_{2,t}(y_{2t} | \mathcal{F}_{t-1}), \dots, F_{n,t}(y_{nt} | \mathcal{F}_{t-1}) | \mathcal{F}_{t-1}) \quad (4-5)$$

$$\forall \mathbf{y} \in \mathbb{R}^n$$

onde $Y_i | \mathcal{F}_{t-1} \sim F_{i,t}$ e $\mathbf{C}_t$ é a cópula condicional de $\mathbf{Y}_t$ dado $\mathcal{F}_{t-1}$.

É importante notar que o mesmo conjunto de informações $\mathcal{F}_{t-1}$ deve ser usado em cada uma das distribuições marginais e também para a cópula, de maneira que a função distribuição resultante seja a distribuição condicional conjunta multivariada. Esse fato é relevante na construção de modelos de

densidade condicional usando a teoria de cópulas. Ao não utilizar o mesmo $\mathcal{F}_{t-1}$ para $\mathbf{C}_t$ e $F_{i,t} \forall i = 1, \dots, n$, a função distribuição resultante $F(\cdot|\cdot)$ não é em geral uma distribuição conjunta com as distribuições condicionais marginais especificadas. Tal resultado é provado em Patton (2009), demonstrando que ao não utilizar todo o conjunto de informações disponível ao estimar os modelos marginais, o lado direito da equação (4-5) não é uma distribuição condicional conjunta válida.

Porém, existem casos especiais nos quais certas variáveis impactam apenas em um certo subconjunto de variáveis e não outras. Por exemplo, pode ser que cada variável dependa apenas de seus próprios valores defasados, não incluindo informação das demais. É postulado um subconjunto $\mathcal{F}_{i,t-1}$ de $\mathcal{F}_{t-1}$, tal que $Y_{it}|\mathcal{F}_{i,t-1} \stackrel{D}{=} Y_{it}|\mathcal{F}_{t-1}$. Dessa forma é possível construir cada modelo marginal usando apenas $\mathcal{F}_{i,t-1}$, que será diferente para cada marginal e posteriormente utilizar $\mathcal{F}_{t-1}$ para a cópula, obtendo uma distribuição condicional conjunta válida.

É importante realizar algum teste antes de reduzir o conjunto de informação. Em Patton (2009), o autor estuda a modelagem da distribuição condicional conjunta entre duas variáveis e utiliza o teste de causalidade de Granger para justificar que as respectivas marginais dependem apenas de seus próprios valores defasados.

Visto que no arcabouço de modelos GAS, estacionariedade das séries temporais não é uma condição estritamente necessária, o teste de causalidade de Granger não pode ser aplicado. Dessa forma, é proposta nessa dissertação o uso do teste de razão de verossimilhança (RV), retirado de Casella e Berger (2001), para a avaliação da independência dos subconjuntos $\mathcal{F}_{i,t-1}$ de $\mathcal{F}_{t-1}$. O teste consiste na aplicação dos seguintes passos elucidados a seguir.

1. Estime os modelos marginais condicionados a $\mathcal{F}_{i,t-1}$ para $i = 1, 2, \dots, n$, salvando o valor da verossimilhança obtido com $\hat{\theta}_i$ como $V^{(1)}$.
2. Estime os modelos marginais condicionados a $\mathcal{F}_{t-1}$, usando de 1 a 12 *lags* das demais variáveis do conjunto de informação e salve o valor da verossimilhança obtido com $\hat{\theta}_i^*$ como $V^{(2)}$.
3. Calcule a estatística de teste,

$$\Lambda = -2 * \ln \left(\frac{V^{(2)}}{V^{(1)}} \right) \quad (4-6)$$

Assintoticamente, $\Lambda \sim \chi_m^2$, onde m é a diferença entre o número de parâmetros estimados em $\hat{\theta}_i^*$ e $\hat{\theta}_i$. A hipótese de que $\mathcal{F}_{i,t-1}$ é independente

de $\mathcal{F}_{t-1}$ para $i = 1, 2, \dots, n$ é rejeitada caso $\Lambda > \chi_m^2$.

É importante salientar que, em se tratando de modelos GAS, na modelagem de cada marginal condicionada a $\mathcal{F}_{t-1}$, essa informação é adicionada ao cálculo do *score* através da introdução de variáveis exógenas (e_t) no mecanismo de atualização do parâmetro variante no tempo. Usando como exemplo o modelo Beta GAS SARIMA (p,q) do Capítulo 3, para cada uma das n marginais em $\mathcal{F}_{t-1}$,

$$\beta_t = \exp \{f_t + e_t\}, \text{ onde} \quad (4-7)$$

$$e_t = \sum_{i=1}^n \sum_{j=1}^q \phi_{ij} y_{i,t-j+1}, \quad (4-8)$$

excluindo o y_{it} associado à distribuição de y_t . Posteriormente para o estudo de caso apresentado no Capítulo 6, as marginais serão condicionadas apenas em $\mathcal{F}_{i,t-1}$, para $i = 1, 2, 3$ dado que na aplicação do teste de RV não houve rejeição da hipótese nula em nenhum dos três casos.

4.2.2

Estimação do modelo de cópula para séries temporais

Grande parte das aplicações de cópula em séries temporais multivariada se baseia no caso onde o vetor de parâmetros desconhecido é dividido em duas partes, uma contendo os elementos referentes às distribuições marginais e a outra contendo os elementos relacionados à cópula. Nessa situação, uma opção natural é a estimação via máxima verossimilhança multi-estágio. Esse método é denominado de IFM¹, sendo proposto pelos autores em Xu (1996) e Joe (1997). Em Joe (2005), o autor afirma que a perda de qualidade do modelo obtido ao usar tal metodologia em dois passos é mínima e não justifica o custo computacional da modelagem em uma única etapa. Vale ressaltar que a questão principal é que, se estimar simultaneamente os parâmetros das marginais e da cópula, a convergência do algoritmo fica mais difícil.

Basicamente, a metodologia envolve estimar os parâmetros de cada uma das distribuições marginais separadamente via máxima verossimilhança univariada e, em uma segunda etapa, estimar os parâmetros da cópula, condicional as estimativas dos parâmetros das marginais.

Para modelos multivariados obtidos via cópulas, em geral não há uma expressão analítica para a verossimilhança e técnicas numéricas são necessárias para sua estimação. Para a estimação via máxima verossimilhança, o número

¹Inference for Margins

de parâmetros aumenta a medida que o número de dimensões do vetor de séries temporais m aumenta. Conforme concluído pelo autor em Joe (2005), essas são algumas das motivações para a utilização do método multi-estágio.

Neste mesmo artigo, o autor define a metodologia da seguinte maneira. Considerando o modelo de cópula paramétrico $\mathbf{C}_t$ para o vetor de variáveis aleatórias $\mathbf{Y}$, a função densidade de probabilidade conjunta é tal que $f(\cdot|\alpha_1, \alpha_2, \dots, \alpha_m, \delta)$, e é assumido que F_j possui densidade f_j para $j = 1, \dots, m$. Considere a partição do vetor de parâmetros desconhecidos $\Psi = (\alpha_1, \alpha_2, \dots, \alpha_m, \delta)$ para as m marginais e para o vetor de parâmetros da cópula (δ). Dessa forma $\mathbf{Y}$ possui densidade conjunta dada por

$$f(\mathbf{y}_t|\alpha_1, \dots, \alpha_m, \delta) = c(F_1(y_{1t}|\alpha_1), \dots, F_m(y_{mt}|\alpha_m)|\delta) \prod_{j=1}^m f_j(y_{jt}|\alpha_j). \quad (4-9)$$

Para uma amostra de tamanho n , com o vetor de variáveis aleatórias $\mathbf{y}$, são consideradas as m funções de log-verossimilhança para as marginais univariadas,

$$L_j(\alpha_j) = \sum_{i=1}^m \log f_j(y_{ijt}|\alpha_j) \quad j = 1, \dots, m \text{ e } i = 1, \dots, n \quad (4-10)$$

Conjecturando a função log-verossimilhança da distribuição conjunta como,

$$L(\delta, \alpha_1, \dots, \alpha_m) = \sum_{i=1}^m \log f(\mathbf{y}_{it}|\alpha_1, \dots, \alpha_m, \delta). \quad (4-11)$$

Segregando o processo em dois estágios, seriam feitas inicialmente as m otimizações das verossimilhanças univariadas separadas. No segundo estágio, seria feita a otimização da verossimilhança multivariada, condicionada ao vetor de parâmetros obtido no primeiro estágio da modelagem. Conforme resumo de Joe (2005), a metodologia se baseia no seguinte fluxo,

1. Maximize as funções de log-verossimilhança L_j das m marginais univariadas separadamente, obtendo as estimativas de $\hat{\alpha}_1, \dots, \hat{\alpha}_m$
2. Maximize a função $L_j(\delta, \hat{\alpha}_1, \dots, \hat{\alpha}_m)$ para obter $\hat{\delta}$.

Esse procedimento é muito mais simples do que estimar todos os parâmetros simultaneamente na Equação (4-11). Um comparativo desse resultado é apresentado em Joe (2005).

4.2.3

Estimação dos erros dos parâmetros da cópula

Enquanto que a estimação é simplificada ao segmentar a verossimilhança em dois estágios, inferências nos parâmetros da cópula resultantes da estimação são mais difíceis. Isso pelo fato de que o erro de estimação das distribuições marginais devem também ser levados em consideração ao se estimar os erros padrão associados aos parâmetros das cópulas. Consequentemente, não basta tomar o inverso da matriz Hessiana avaliada no ótimo como um estimador da matriz de covariância assintótica dos estimadores dos parâmetros de cópula, pois esse procedimento ignora o erro padrão das estimativas dos parâmetros das marginais.

Maneiras alternativas para calcular esse erro padrão são baseadas em técnicas de *bootstrap* por blocos, como por exemplo o *bootstrap* paramétrico postulado em Politis e Romano (1994). Tal técnica é realizada supondo que a série seja estacionária, e é frequentemente utilizada em modelos para séries de retornos financeiros Avdulaj e Barunik (2015). O passo a passo para o cálculo alternativo é descrito a seguir e foi retirado da página 22 de Patton (2012).

1. Use um *bootstrap* por blocos para gerar uma amostra da série de tamanho T .
2. Estime o modelo usando a mesma técnica multi-estágio aplicada a série de fato.
3. Repita os dois passos anteriores B vezes (ex: $B = 1000$).
4. Calcule os quantis $\alpha/2$ e $1 - \alpha/2$ da distribuição de $\{\hat{\theta}\}_{i=1}^B$ de maneira a obter o intervalo de confiança $1 - \alpha$ dos parâmetros. Vale ressaltar que os parâmetros da cópula e das marginais estão contidos em $\{\hat{\theta}\}$.

Apesar do procedimento descrito acima ser mais adequado, os erros das estimativas dos parâmetros da cópula apresentada no Capítulo 6 serão calculados sem utilizar a informação das marginais. Isso porque implementar o procedimento no arcabouço da nossa aplicação não é trivial, uma vez que as séries de fator de capacidade são sazonais. Basta ter a consciência de que o valor dos erros dos parâmetros da cópula no Capítulo 6 serão sub avaliados.

5

Modelo *t-Student* multivariado

Nesse capítulo será apresentada a especificação do modelo definido em Creal *et al.* (2011), com a especificação pertinente à aplicação utilizada na dissertação.

5.1

Densidade *t-Student* multivariada

O modelo proposto utiliza a especificação GAS apresentada no Capítulo 2, com fins de introduzir uma dinâmica temporal para a atualização da matriz de covariância (Σ_t). Por estar configurado no arcabouço de uma distribuição *t-Student* com ν graus de liberdade, este se vale das mesmas características pertinentes à distribuição, como caudas pesadas e simetria. Admitindo também como caso especial o modelo Normal Multivariado, quando $\nu^{-1} \rightarrow 0$, conforme apresentado no Apêndice B.

Defina o vetor de observações $y_t \in \mathbb{R}^k$ seguindo uma distribuição *t-Student* com ν graus de liberdade e matriz de covariância Σ_t . A média de y_t , μ_t , será assumida nula por simplicidade de notação. Assuma que $\nu > 2$, tal que a matriz de variância exista.

A densidade preditiva de y_t é dada por

$$p(y_t|\Sigma_t; \nu) = \frac{\Gamma\left(\frac{\nu+k}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) [(\nu-2)\pi]^{k/2} |\Sigma_t|^{1/2}} \left[1 + \frac{y_t' \Sigma_t^{-1} y_t}{(\nu-2)}\right]. \quad (5-1)$$

A matriz de covariância é decomposta como

$$\Sigma_t = D_t R_t D_t, \quad R_t = \Delta^{-1} Q_t \Delta^{-1}, \quad \Delta_t = \text{diag}(Q_t)^{1/2}, \quad (5-2)$$

onde D_t é uma matriz diagonal que contém os desvios padrões de y_t . Em nosso contexto as variâncias são assumidas conhecidas, logo $D_t = I$, posto que em uma especificação de cópula apenas as correlações são tomadas como variantes no tempo. Além disso Q_t é uma matriz simétrica, positiva definida e Δ_t uma matriz diagonal, cujos elementos não nulos são iguais a raiz quadrada dos

elementos da diagonal principal de Q_t . A matriz R_t aplica uma parametrização na matriz Q_t que garante uma matriz Σ_t simétrica, positiva definida, com elementos entre $[-1,1]$ nos elementos fora da diagonal principal.

Essa decomposição é similar à utilizada no modelo DCC de Engle (2002A) e Engle e Sheppard (2001). A única diferença fica por conta da dinâmica utilizada para atualizar R_t e D_t . Enquanto que nos modelos DCC os elementos da matriz de covariância são atualizados pelo produto externo do vetor de inovações padronizadas, o modelo com atualização GAS usa apenas a informação do *score* da densidade preditiva (5-1) como mecanismo de atualização para os elementos contidos na matriz de covariância.

Para introduzir o desenvolvimento da função de atualização GAS para a densidade (5-1) com especificação de cópula, assimila-se a matriz de covariância Σ_t em (5-2) como uma função $\Sigma_t(f_t)$ do fator variante no tempo f_t . Antes de contextualizar esse desenvolvimento, é essencial introduzir algumas definições matriciais, uma vez que este modelo está inserido em um contexto multivariado. Esse será o tema da Sessão 5.2.

5.2

Notações matriciais e definições

A presente sessão tem como objetivo introduzir ao leitor operadores fundamentais para a definição da função do mecanismo de atualização GAS utilizado nessa dissertação. Em Abadir e Magnus (2005) é oferecido um tratamento completo de cálculo matricial útil em econometria. Além disso em Creal *et al.* (2011) são exibidas definições utilizadas em outro escopo além de cópulas, com as variâncias da matriz de covariância também variantes no tempo. Para as aplicações que serão descritas posteriormente, as implementações computacionais destas foram feitas utilizando a biblioteca *matrixcalc* no software R.

O produto de Kronecker entre duas matrizes A e B será denotado por $A \otimes B$. No caso em que $A = B$, adotaremos $A_{\otimes}$. O operador $vec(A)$ vetoriza a matriz A em um vetor coluna, enquanto que $vech(A)$ vetoriza a parte triangular inferior dessa matriz. O operador $\oplus$ é operacionalizado como

$$A \oplus B = (A \otimes B) + (B \otimes A)$$

e as matrizes de duplicação $\mathcal{D}_k$ e eliminação $\mathcal{B}_k$ são definidas como

$$\mathcal{D}_k vech(A) = vec(A), \quad \mathcal{B}_k vec(A) = vech(A),$$

para uma matriz A simétrica $k \times k$ e com $\mathcal{B}_k = (\mathcal{D}_k' \mathcal{D}_k)^{-1} \mathcal{D}_k'$. Essa matriz de

duplicação consiste de uma matriz de seleção composta de apenas zeros e uns, tendo mais linhas que colunas, enquanto que a matriz de eliminação possui mais colunas que linhas.

A próxima sessão fundamenta o esquema de atualização GAS para a matriz de covariância em um contexto de cópulas variantes no tempo.

5.3

Atualização de Σ_t via GAS

Os resultados aqui expostos são aplicáveis a um modelo de cópula *t-Student* com matriz de correlação variante no tempo e variância unitária. Em razão das variâncias serem conhecidas, a estrutura de variação temporal estocástica de Σ_t é capturada unicamente pela dinâmica de Q_t . Como Q_t é simétrica, será adotado o vetor de parâmetros variantes no tempo como sendo $f_t = \text{vech}(Q_t)$ com dimensões $[k(k+1)/2] \times 1$.

A matriz de informação de Fisher para a parametrização (5-2) é singular e conseqüentemente na equação (2-6) é estritamente necessário substituir $\mathcal{I}_{t|t-1}^{-1}$ pela pseudo-inversa de Moore-Penrose, caso contrário não será possível definir uma matriz inversa em alguns casos. Assim, a formulação do *score* para atualização dos elementos da matriz de covariância Σ_t variantes no tempo, é realizada com base no Teorema 5.3.1 retirado de Creal *et al.* (2011).

Teorema 5.3.1 *Para a densidade (5-1) e assumindo $D_t = I$ em (5-2) com $f_t = \text{vech}(Q_t)$ em (2-4), os componentes do score s_t , ∇_t em (2-5) e $\mathcal{I}_{t|t-1}^{-1}$ em 2-6, são dados por*

$$\nabla_t = \frac{\partial \ln p(y_t | \Sigma_t; \nu)}{\partial f_t} = \frac{1}{2} \Psi'_t \mathcal{D}'_k \Sigma_{t\otimes}^{-1} [w_t y_{t\otimes} - \text{vec}(\Sigma_t)] \quad (5-3)$$

$$\mathcal{I}_{t|t-1} = E[\nabla_t \nabla'_t] = \frac{1}{4} \Psi'_t \mathcal{D}'_k J'_{t\otimes} [gG - \text{vec}(I)\text{vec}(I)'] J_{t\otimes} \mathcal{D}_k \Psi_t \quad (5-4)$$

com $\Psi_t = \partial \text{vech}(\Sigma_t) / \partial f'_t$, o escalar $w_t = (\nu + k) / (\nu - 2 + y'_t \Sigma_t^{-1} y_t)$, com matriz J_t definida por $\Sigma_t^{-1} = J'_t J_t$ que pode ser obtida a partir de qualquer procedimento de decomposição de matriz. O escalar $g = (\nu + k) / (\nu + 2 + k)$ e os elementos da matriz G , de dimensão $k \times k$, são dados por

$$G[(i-1) \cdot k + \ell, (j-1) \cdot k + m] = \delta_{ij} \delta_{\ell m} + \delta_{i\ell} \delta_{jm} + \delta_{im} \delta_{j\ell} \quad (5-5)$$

com $i, j, \ell, m = 1, \dots, k$ e assumindo δ_{ij} igual a 1 se $i = j$ e zero caso contrário.

Este teorema é a base para incorporar o mecanismo de atualização GAS nos elementos da matriz de covariância. Referencia-se o artigo supracitado,

para uma melhor compreensão acerca das especificações alternativas da matriz Σ_t além da utilizada nessa dissertação. Para a especificação de cópula variante no tempo, a decomposição da matriz Σ_t é dada por

$$\Psi_t = \mathcal{B}_k D_{t \otimes} \Delta_{t \otimes}^{-1} [\mathcal{D}_k - (\Delta_t \oplus Q_t) \Delta_{t \otimes}^{-1} W_{\Delta_t} \mathcal{S}_{\Delta}] \mathcal{S}_Q \quad (5-6)$$

onde $\mathcal{S}_Q = I$ dado que $f_t = \text{vech}(Q_t)$. A matriz W_{Δ_t} é construída inicialmente tomando-a como uma matriz diagonal de dimensão $k^2 \times k^2$ com os elementos na diagonal principal iguais à $0.5 \text{vec}(\Delta_t^{-1})$, e posteriormente eliminam-se as colunas que contenham apenas zeros. Já a matriz $\mathcal{S}_{\Delta}$ operacionaliza como uma matriz de seleção, contendo apenas zeros e uns, representada por $\text{diag}(\Delta_t^2) = \mathcal{S}_{\Delta} \text{vech}(Q_t)$.

Além disso, conforme colocado em Creal *et al.* (2011), para transformar o modelo GAS *t-Student*(p,q) em uma especificação de cópula variante no tempo, a densidade (5-1), irá operar no seguinte vetor, de dimensão $k \times 1$,

$$\tilde{y}'_t = [F_{\nu}^{-1}(u_{1t}), F_{\nu}^{-1}(u_{2t}), \dots, F_{\nu}^{-1}(u_{kt})]'. \quad (5-7)$$

onde $F_{\nu}^{-1}(\cdot)$ é o inverso da F.D. de uma *t-Student* univariada.

Em Creal *et al.* (2011), os autores enumeram as seguintes quatro propriedades ao implementar o Teorema 5.3.1.

1. As dinâmicas dos elementos do vetor f_t serão guiadas pelos desvios do produto do vetor ponderado $w_t y_t y'_t$ com relação a matriz de covariância Σ_t . Para se ter uma ideia da sensibilidade desse escalar, em um contexto de uma distribuição normal, o valor dos pesos w_t colapsam para 1 e o mecanismo de atualização se reduz ao do modelo GARCH multivariado, conforme definido no Teorema 5.3.1;
2. O Teorema é genérico, uma vez que este é similar tanto para a especificação de correlação e/ou volatilidades, a única diferença estará na especificação de Ψ_t e f_t ;
3. O termo w_t introduz uma característica única ao modelo, robustez em termos de possíveis *outliers*. Esse parâmetro automaticamente leva em consideração os valores extremos porque esse atrela uma ponderação menor a y_t se o valor de $y'_t \Sigma_t^{-1} y_t$ for alto. Isto é, caso as séries dos termos de erro assumam algum valor extremo, w_t diminui a quantidade $\tilde{y}'_t \Sigma_t^{-1} \tilde{y}_t$ na atualização do *score* ∇_t melhorando o ajuste dos dados, i.e., não contamina o restante da atualização dos valores de Σ_t com um choque permanente nos elementos de s_t . Dessa forma, indo de acordo com a

conclusão dos autores do artigo, é importante especificar para a função de verossimilhança do termo de erro uma densidade preditiva que contemple caudas pesadas. Ainda, devido principalmente a presença desse termo de ponderação, o modelo GAS apresentado não se reduz ao modelo DCC de Engle (2002A).

4. Uma característica particular do *score*, s_t , para a densidade preditiva *t-Student* é que esta forma uma sequência de diferença martingal e age como uma inovação para o modelo. Isso segue das propriedades do vetor *score*, apresentado em 2-5.

5.4

Especificação de Cópula variante no tempo

Conforme apresentado na Sessão 4.1, as seguintes explicações serão contextualizadas em um arcabouço bivariado para facilitar a notação, posto que a extensão deste para um caso multivariado é trivial. O primeiro passo para a estimação dos parâmetros da cópula condicional dada por

$$F(y_{1t}, y_{2t} | \mathcal{F}_{t-1}) = C_{f_t}[F(y_{1t} | \mathcal{F}_{t-1}), F(y_{2t} | \mathcal{F}_{t-1}) | \mathcal{F}_{t-1}] \quad (5-8)$$

é calcular as pseudo-observações, ou PIT¹, $\hat{u}_t$ e $\hat{v}_t$ posto que o cálculo dessas é função dos parâmetros estimados das marginais. Tal flexibilidade é uma das vantagens do modelo com densidade preditiva *t-Student*, que será aqui denotado por modelo GAS *t-Student*(p,q), pois trata as distribuições marginais de cada série separadamente da estrutura de dependência conjunta. Isso permite que o problema de otimização dos parâmetros seja quebrado em dois estágios, tornando-o mais tratável. Conforme visto na Sessão 4.2.2.

O primeiro estágio diz respeito a estimação dos parâmetros das densidades marginais, associados a cada uma das séries de fator de capacidade. Assuma que ambas sejam modeladas individualmente por um modelo Beta GAS SARIMA(p,q)²,

$$(y_{jt} | \mathcal{F}_{t-1}) \sim \text{Beta}[k_j, \beta_{jt}, \alpha_j | \mathcal{F}_{t-1}] \quad \forall j = \{1, 2\} \quad (5-9)$$

onde pela reparametrização aplicada no vetor de parâmetros variantes no tempo, $\beta_{jt} = \exp(f_{jt})$. Após a estimação dos parâmetros das marginais,

¹Em inglês, *probability integral transform (PIT)*.

²O modelo referenciado foi escolhido apenas para ficar próximo ao arcabouço que será aplicado na dissertação, não obstante poderia ser qualquer outro.

definem-se as PIT's como,

$$\hat{u}_t = F(y_{1t}|\mathcal{F}_{t-1}) \quad (5-10)$$

$$\hat{v}_t = F(y_{2t}|\mathcal{F}_{t-1}). \quad (5-11)$$

Assumindo que as densidades em (5-9) sejam adequadas para modelar as séries de FC individualmente, então espera-se que $\hat{u}_t$ e $\hat{v}_t$ sejam duas variáveis aleatórias *i.i.d.* com distribuição *Uniforme*[0,1]. Isso já é de alguma maneira testado ao julgarmos adequados os modelos marginais via diagnósticos padrões nas séries de resíduos quantílicos produzidos. Em Patton (2002), o autor indica alguns testes para julgar a independência entre essas variáveis com maior rigor. Na aplicação do Capítulo 6 será proposta a aplicação de um teste BDS para investigar tal propriedade.

Via a informação disponibilizada pelas pseudo-observações (PIT), parte-se para o segundo estágio do problema de otimização com o objetivo de estimar os parâmetros da cópula condicional. De (5-8), segue que

$$F(y_{1t}, y_{2t}|\mathcal{F}_{t-1}) = C_{f_t}(\hat{u}_t, \hat{v}_t) \quad (5-12)$$

onde $f_t = f(f_{t-1}, \theta)$, com θ sendo estimado via máxima verossimilhança nesse segundo estágio da otimização.

5.5

Previsão do modelo GAS com densidade preditiva $t_{(\nu, \Sigma_t)}$

Seguindo a lógica de modelagem via cópulas, ao permitir que a densidade (5-1) opere no vetor $(\hat{u}_t, \hat{v}_t)$ via teorema de Sklar, espera-se que a previsão resultante resulte em cenários capazes de reproduzir uma estrutura de dependência mais apurada. Para tal, é considerado um conjunto de informações mais rico das variáveis aleatórias envolvidas, conjunto esse que supõem relações entre as variáveis, positivas ou negativas, tal que a modelagem individual de cada marginal não seja condicionada apenas a sua respectiva informação passada.

Uma vez estimados os parâmetros da densidade conjunta via (5-12), a geração de pares correlatados de $(\hat{u}_t, \hat{v}_t)$ com a estrutura de dependência implicada pela cópula escolhida é realizada segundo o algoritmo *Conditional Sampling* (CS), detalhado em Cherubini *et al.* (2004) e apresentado no Apêndice C. Tal algoritmo gera pares de observações $(\hat{u}_t^*, \hat{v}_t^*)$ que sejam distribuídas uniformemente entre [0,1] com função de distribuição conjunta dada pela cópula, C_{f_t} .

Como $C_{f_t} \sim t_{(\nu, \Sigma_t)}$ é uma cópula elíptica, então o processo de geração é

mais direto, pois já é conhecida a forma analítica da distribuição conjunta, a qual já possui algoritmos conhecidos para a sua geração. Para as implementações computacionais no nosso trabalho foi utilizada a biblioteca *mvtnorm* no software R para a geração de variáveis aleatórias com distribuição *t-Student* multivariada. Os passos dessa implementação são apresentados a seguir.

1. Em posse de $(\hat{u}_t, \hat{v}_t)$, estime via máxima verossimilhança, o vetor de parâmetros estáticos da cópula $\hat{\theta}$.
2. A partir do resultado do passo anterior, atualize a matriz de correlação $\hat{\Sigma}_t$ via mecanismo GAS com $\hat{\omega}, \hat{A}, \hat{B}, \hat{\nu}$;
3. Utilize um gerador de números aleatórios para gerar valores a partir de uma distribuição $t_{(\hat{\nu}, \hat{\Sigma}_t)}$ com média nula.
4. Com os valores gerados no passo anterior, aplique a cada valor do par gerado a metodologia de PIT, descrita anteriormente, supondo uma distribuição *t-Student* com $\hat{\nu}$ graus de liberdade.

Em seguida, parte-se para a previsão dos valores futuros condicionais de $Y_{t+k} = \{y_{1,t+k}, y_{2,t+k}\}$, acrescentando a estrutura de dependência imposta como função daquela descrita na Sessão 3.4 para o modelo GAS univariado. Ao leitor não familiarizado com tal metodologia, é indicado reler a explicação anterior antes de avançar. A previsão é obtida a partir do procedimento apresentado a seguir.

1. Usando apenas as previsões com a metodologia univariada, obtenha as previsões individuais de $y_{t+k}^{(m)}$, explicadas na Sessão 3.4, como também as PIT's geradas a partir dessas previsões, que serão denotadas por $u_{t+k}^{(m)}$, onde m representa o número de caminhos e k o número de passos a frente.
2. De posse dessas previsões, aplique nas m trajetórias a metodologia de CS supracitada, gerando $u_{t+k}^{*(m)}$ e $v_{t+k}^{*(m)}$. Dessa forma as previsões condicionais K passos a frente $y_{t+k}^{*(m)}$ são obtidos pelo seguinte algoritmo:
 - (a) Para $k \geq 1$, a forma analítica da distribuição preditiva um passo à frente ainda é conhecida. Para tanto, basta aplicar o quantil

$F^{-1}(\cdot)$ da densidade preditiva pertinente aos dados,

$$y_{1,t+1}^{*(m)} = F^{-1}(u_{t+1}^{*(m)} | \mathcal{F}_{t-1}) \quad (5-13)$$

$$y_{2,t+1}^{*(m)} = F^{-1}(v_{t+1}^{*(m)} | \mathcal{F}_{t-1}) \quad (5-14)$$

- (b) Para $k \geq 2$, a forma analítica da distribuição de previsão não é conhecida e assim, deve ser obtida via simulação Monte Carlo. Porém, como o problema está inserido em um arcabouço de cópula paramétrica, a densidade da marginal já é conhecida. Logo, em posse dos valores de $y_{t+k}^{(m)}$ e $u_{t+k}^{*(m)}$ obtêm-se o quantil $u_{t+k}^{*(m)}$ da distribuição de $y_{t+k}^{(m)}$.

Figura 5.1: Ilustração gráfica dos valores condicionais de $y_{t+k}^{*(m)}$.

Obtendo dessa forma a distribuição de $y_{t+k}^{*(m)}$. Então para obter o valor da previsão pontual a cada passo k , calcula-se a média das m trajetórias a cada instante de tempo k . Já para os intervalos de confiança $(1 - \alpha)$, obtêm-se os quantis $\alpha/2$ e $1 - \alpha/2$.

Para esclarecer o passo-a-passo da modelagem apresentada até aqui, é apresentado um resumo no diagrama de blocos exposto na Figura 5.2.

5.6

Estimação e propriedades estatísticas

Suponha um conjunto de n realizações do vetor de observações y_t para $t = 1, \dots, n$ com média nula. A função log da verossimilhança do modelo *t-Student* multivariado é dado por

Figura 5.2: Diagrama de blocos para a modelagem incluindo a estrutura de dependência especificada pelo modelo $GAS t\text{-}Student(p,q)$.

$$\mathcal{L} = \sum_{t=1}^n \left\{ \log \left[\Gamma \left(\frac{\nu + k}{2} \right) \right] - \log \left[\Gamma \left(\frac{\nu}{2} \right) \right] - \frac{1}{2} \log |\Sigma_t| - \frac{k}{2} \log [(\nu - 2)\pi] - \frac{\nu + k}{2} \log \left[1 + \frac{y_t' \Sigma_t^{-1} y_t}{\nu - 2} \right] \right\} \quad (5-15)$$

onde a variação temporal de Σ_t é determinada pelas equações de atualização GAS para $vech(Q_t)$, conforme equações (2-4) e (5-2). Os coeficientes desconhecidos são incorporados ao vetor de parâmetros estáticos θ e sua estimação é baseada na maximização da função de verossimilhança (5-15) via o procedimento padrão de otimização quasi-Newton.

Com relação a otimização, no arcabouço GAS, é fácil de impor restrições adicionais no espaço paramétrico primando pela parcimônia e facilidade de estimação dos parâmetros desconhecidos do modelo. Visto que um modelo mais parcimonioso leva a um problema de otimização de dimensões mais moderadas, facilitando sua solução. Um exemplo dessa técnica é a normalização

mencionada em Creal *et al.* (2011) para estimar os elementos de ω sem a necessidade de incluí-los no vetor de parâmetros desconhecidos. Para isso, basta definir, após a estimação dos elementos de A e B ,

$$w = (I - B_1 - \dots - B_q) \cdot \text{vech}(\bar{Q}) \quad (5-16)$$

com $\bar{Q}$ sendo a matriz de correlação amostral incondicional. Empiricamente, ao incluir essa normalização, os resultados em termos de MAE e MSE foram piores nos modelos simulados, o que resultou no abandono dessa técnica e a inclusão de ω no vetor de parâmetros estáticos θ .

Além dessa normalização, também é mencionada a técnica de reduzir o espaço paramétrico restringindo as matrizes $A_1 = a_1 I$ e $B_1 = b_1 I$. Dessa forma, apenas dois parâmetros irão controlar a dinâmica das equações de atualização GAS para os elementos da matriz Σ_t . Tal idéia é retirada dos modelos DCC para tornar a estimação dos parâmetros desconhecidos via máxima verossimilhança menos árdua (Creal *et al.*, 2011). Essa técnica apresentou bons resultados e foi mantida na estimação do modelo do Capítulo 6.

Além disso, é impossível impor restrições em A e B que garantam que o vetor de parâmetros variantes no tempo permaneça no domínio especificado, i.e., correlação entre $[-1,1]$. Esse problema é contornado adotando a estratégia de reparametrização, descrita na Sessão 2.3, restringindo os elementos da matriz Σ_t da seguinte forma,

$$\Sigma_{(i,j),t} = \frac{1 - \exp(-f_{j,t})}{1 + \exp(-f_{j,t})} \quad \forall i, j \in \{1, \dots, k\}, i \neq j \quad (5-17)$$

onde $f_t = \text{vech}(Q_t)$.

Para a estimação do grau de liberdade ν , um algoritmo iterativo seria necessário, onde ν fosse estimado simultaneamente com os demais parâmetros contidos em θ , uma vez que a densidade preditiva opera no vetor (5-7). Para a aplicação do Capítulo 6, será utilizado um processo denominado de verossimilhança perfilada, no qual será fixado a priori o número do grau de liberdade $\nu = 4, 5, \dots$, de maneira a evitar a estimação desse parâmetro. Para cada um desses valores estima-se o modelo GAS *t-Student*(p,q), e é escolhido o ν cujo valor maximize o valor da verossimilhança. O custo de utilizar essa metodologia é que não necessariamente será escolhido o conjunto de parâmetros ótimos.

5.6.1

Otimização com restrições nos parâmetros estáticos

Diferentemente do modelo Beta GAS SARIMA(p,q) especificado no Capítulo 3, na qual a otimização dos parâmetros estáticos era irrestrita, uma estrutura de restrições deve ser imposta a alguns desses parâmetros dentro da otimização no modelo GAS *t-Student*(p,q). Para contextualizar tal necessidade, os autores em Creal *et al.* (2011) utilizam uma analogia com o modelo GARCH(1,1) para mostrar a razão do uso dessa estrutura em α e β . Tal modelo é um caso particular do modelo GAS(1,1) se for atribuído $f_t = \sigma_t^2$, com atualização dada por $f_{t+1} = \omega + \alpha_1 y_t^2 + \beta_1 f_t$, com $\alpha_1 = A_1$ e $\beta_1 = B_1 - A_1$. Seguindo a ideia do que é imposto aos modelos GARCH para manter a variância positiva em todo o instante de tempo t e o processo ser estacionário na covariância, é necessário que $\alpha_1 \geq 0$ e $\beta_1 \geq 0$, bem como $\alpha_1 + \beta_1 < 1$.

Um conjunto de restrições análogo ao do modelo GARCH(1,1) é também adotado no modelo GAS(1,1), mas com $1 > B_1 \geq A_1 \geq 0$. Porém, conforme já mencionado, almeja-se nessa dissertação se manter próximo do modelo GARCH multivariado. Para isso foi definido que $S_t = \mathcal{I}_{t|t-1}^{-1}$, sendo assim, conforme Creal *et al.* (2011), formular condições de estacionariedade não é uma tarefa trivial. A restrição de que $I - zB_1 - \dots - z^q B_q$ esteja fora do círculo unitário não é necessariamente uma condição necessária para estacionariedade na covariância.

Visando atender a restrição de que $1 > B_1 \geq A_1 \geq 0$, é proposta nessa dissertação uma combinação convexa entre os parâmetros A e B de tal sorte que atendam essa restrição. Para especificar a função em questão, inicialmente atenta-se ao fato de que os elementos da matriz B devem se restringir ao intervalo $[0, 1] \forall x \in \mathbb{R}^k$. Dessa forma, utiliza-se que

$$B = [\Psi \cdot A_1 + (1 - \Psi)] \cdot I \quad (5-18)$$

onde $\Psi \in [0, 1]$ e I é uma matriz identidade de dimensão $k(k+1)/2$, sendo k o número de variáveis envolvidas no problema.

No que se refere ao parâmetro ω , por falta de uma sensibilidade sobre qual a amplitude de valores este poderia assumir, optamos por estimá-lo livre de restrições. Em se tratando de um intercepto na atualização autoregressiva, este acompanhou bem a dinâmica do parâmetro $\rho_{t|t-1}$, conforme pode-se comprovar pelos gráficos na Figura 5.3.

O algoritmo *BFGS* foi utilizado para a otimização. A heurística utilizada para a otimização dos parâmetros no modelo Beta GAS SARIMA(p,q),

intercalando a otimização via *BFGS* e *Nelder-Mead*, foi abandonada em consequência de que as soluções encontradas eram muito próximas, diminuindo significativamente o tempo computacional. Posto que o método *BFGS* se baseia no vetor de parâmetros $\lambda_n \in \mathbb{R}$ para avaliar a solução ótima, supondo o vetor $\lambda = \{\lambda_1, \lambda_2, \lambda_3\}$, as funções restritivas impostas para cada um dos parâmetros θ são dados por

$$A = \frac{\exp(\lambda_1)}{1 + \exp(\lambda_1)} \cdot I \quad (5-19)$$

$$B = [\Psi \cdot A_1 + (1 - \Psi)] \cdot I \quad (5-20)$$

$$\Psi = \frac{\exp(\lambda_2)}{1 + \exp(\lambda_2)} \cdot I \quad (5-21)$$

Com relação a ω , conforme colocado em Creal *et al.* (2011), para não gerar problemas de identificação, atribui-se 1 aos elementos em ω que correspondem aos elementos da diagonal de Q_t . Na aplicação do Capítulo 6, onde se modelam 3 variáveis ($k = 3$), a otimização desse parâmetro é feita como

$$\omega = (1, \lambda_3, \lambda_3, \lambda_3, 1, \lambda_3, \lambda_3, \lambda_3, 1)' \quad (5-22)$$

5.6.2

Significância dos parâmetros

É de suma importância alertar para uma particularidade referente ao cálculo dos erros padrões dos parâmetros com essa estrutura de restrições. Já que foram impostas restrições sobre os parâmetros de θ , deve-se utilizar também uma estrutura reparametrizada para o cálculo dos erros padrões.

Esse cálculo é efetuado com base no inverso da matriz Hessiana avaliada no ótimo, ou matriz de Informação de Fisher. Como essa última depende da parametrização do problema, no caso multiparamétrico sendo $\theta = (\theta_i, \dots, \theta_s)$, sob reparametrização teríamos

$$\theta_i = h_i(\xi_1, \dots, \xi_s), \quad i = 1, \dots, s, \quad (5-23)$$

e sendo J a matriz Jacobiano, essa é definida por,

$$J = \left\| \frac{\partial \theta_j}{\partial \xi_i} \right\| \quad (5-24)$$

Desse modo, $\mathcal{I}^*(\xi) = \|\mathcal{I}_{ij}^*(\xi)\|$, é definida como a matriz de informação

de $(\xi_1, \dots, \xi_s)$. Pelo uso da regra da cadeia, obteríamos

$$\mathcal{I}^*(\xi) = J \cdot \mathcal{I} \cdot J', \quad (5-25)$$

onde $\mathcal{I}$ é a matriz de informação de Fisher na escala original.

Para maiores detalhes dessa prova, referenciamos a página 125 de Lehmann e Casella (2003), onde a mesma é apresentada.

No caso dessa dissertação, como J é diagonal, sua forma é expressa por

$$\text{Diag}(J) = \left[\frac{\exp(\hat{\lambda}_1)}{(\exp(\hat{\lambda}_1) + 1)^2}, \frac{\exp(\hat{\lambda}_2)}{(\exp(\hat{\lambda}_2) + 1)^2} \cdot (\hat{A} - 1), 1 \right], \quad (5-26)$$

onde $\hat{A}$, é o elemento contido na diagonal principal da matriz A contida em $\hat{\theta}$ do modelo GAS *t-Student*(p,q).

5.6.3

Valores e condições Iniciais

Similar ao que foi esclarecido na Sessão 3.2.1, uma das principais dificuldades na estimação dos parâmetros dos modelos GAS está na especificação de condições iniciais para estes na otimização, bem como de valores iniciais para o vetor de parâmetros f_t . As heurísticas empregadas devem ser próximas àquelas aplicadas no modelo Beta GAS SARIMA(p,q).

Para a obtenção de valores iniciais de f_t , pode-se pensar na inclusão dos elementos de $f_j (j < 1)$ no vetor de parâmetros estáticos e estimá-los juntos aos demais. Porém, como essa otimização está inserida em um contexto multivariado, caso fossem inseridas muitas variáveis, aumentaria exacerbadamente o número de parâmetros a serem estimados no modelo. À vista disso, em Creal *et al.* (2011) os autores propõem o uso da média incondicional como valores iniciais, i.e,

$$(I_k - B_1 - B_2 - \dots - B_q) \cdot \omega. \quad (5-27)$$

Esta é facilmente implementável e empiricamente mostrou-se bem ajustada à dinâmica de f_t . É importante ressaltar que assim como no modelo Beta GAS SARIMA(p,q), essas observações não contribuirão para a verossimilhança do modelo multivariado estimado via cópulas

Referente às condições iniciais, também são gerados 20 conjuntos de valores para cada parâmetro contido em $\theta = \{A, B, \omega\}$. Para os parâmetros

A, B e ω foram gerados 20 valores com distribuição uniforme com suporte entre -1 e 1.

5.7

Simulação

Com o intuito de auferir os resultados da implementação computacional, foi conduzido um estudo limitado de Monte Carlo, similar ao proposto pelos autores em Engle (2002A), e Creal *et al.* (2011). Contudo é proposta uma abordagem alternativa a apresentada em Creal *et al.* (2011). Será estudado apenas o desempenho do modelo apresentado, comparações com o modelo DCC de Engle (2002A) e Engle e Sheppard (2001) já foram feitas no artigo referenciado e não serão relevantes para o objetivo desse trabalho.

O estudo foi delineado da seguinte forma. Inicialmente geram-se 1000 observações de uma distribuição *t-Student* bivariada com variância unitária e coeficiente de correlação ρ_t . O padrão temporal do coeficiente ρ_t obedece as seguintes funções:

1. Constante: $\rho_t = 0.9$;
2. Seno: $\rho_t = 0.5 + 0.4\cos(2\pi t/200)$;
3. Seno de menor frequência: $\rho_t = 0.5 + 0.4\cos(2\pi t/20)$;
4. Passo: $\rho_t = 0.9 - 0.5(t > T/2)$;
5. Modelo: $\rho_t = \exp(h_t)/(1 + \exp(h_t))$, onde $h_t = -0.4(1 - 0.99) + 0.99h_{t-1} + 0.14\eta_t$, $\eta_t \sim N(0, 1)$.

Após a escolha de um dos padrões acima, este é utilizado no processo gerador de dados (PGD) conforme a seguir:

$$y_t \sim p(y_t|\Sigma_t; \nu), \quad \Sigma_t = R_t = \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix}, \nu = 5, \quad (5-28)$$

onde $p(y_t|\Sigma_t; \nu)$ representa a densidade *t-Student* bivariada expressa pela equação (5-1). Para o estudo de Monte Carlo, são simuladas 1000 séries de tamanho $T = \{300, 1000\}$. A escolha dos tamanhos não foi arbitrária; o primeiro valor foi atribuído pelo fato de que as séries no Capítulo 6 possuem aproximadamente esse tamanho. Já o segundo foi estabelecido com o objetivo de permanecer próximo ao estudo de Creal *et al.* (2011), investigando propriedades assintóticas dos EMV.

Em posse das séries $\left[\{Y_{1t}^{(i)}, Y_{2t}^{(i)}\}_{t=1}^T \right]_{i=1}^{1000} \forall T \in \{300, 1000\}$, é estimado o modelo *t-Student* multivariado com mecanismo de atualização GAS(1,1), onde

a dinâmica dos elementos contidos na matriz de correlação serão atualizados segundo mecanismo (2-4). Ao estimar o modelo para cada um dos 1000 pares de séries simuladas, são calculados $\{\hat{\omega}^{(i)}, \hat{A}^{(i)}, \hat{B}^{(i)}, \hat{\nu}^{(i)}\}_{i=1}^{1000}$. Dessa forma, é obtido o coeficiente de correlação condicional $\hat{\rho}_{t|t-1}^{(i)}$ atualizado via mecanismo GAS.

À vista disso, afim de conferir os resultados, para cada $\hat{\rho}_{t|t-1}^{(i)}$, são calculadas duas medidas de aderência, o erro médio absoluto (MAE) e o erro quadrático médio (MSE) expressos por:

$$MAE = \frac{1}{n} \sum_{t=1}^n |\hat{\rho}_t - \rho_t| \quad MSE = \frac{1}{n} \sum_{t=1}^n (\hat{\rho}_t - \rho_t)^2. \quad (5-29)$$

Os resultados para os 5 PGD's estão expostos na Tabela 5.1.

Tabela 5.1: Medidas de aderência do estudo de simulação.

PGD	T=300		T=1000	
	MAE	MSE	MAE	MSE
1	0,041	0,003	0,042	0,002
2	0,314	0,146	0,293	0,125
3	0,269	0,098	0,271	0,098
4	0,303	0,144	0,286	0,134
5	0,305	0,129	0,300	0,128

A vista dos resultados, tanto para $T = 300$ quanto para $T = 1000$, o MAE e o MSE são próximos. Tal fato fornece insumos para concluir que, embora os autores em Creal *et al.* (2011) tenham utilizado essa metodologia para séries de retornos financeiros, as quais possuem uma vasta disponibilidade de dados, é viável a aplicação desse modelo à séries com um tamanho de amostra limitado.

A título de ilustração, referencia-se a Figura 5.3 para apresentar os padrões gráficos de cada um dos PGD's para $T = 300$ e $T = 1000$, junto com uma das 1000 séries de $\hat{\rho}_t$ estimadas escolhidas arbitrariamente.

Figura 5.3: Gráficos dos cinco PGD simulados com $T=300$ e $T=1000$.

6

Uso dos modelos GAS para simular e prever históricos de FC de usinas eólicas

Nesse capítulo iremos apresentar um estudo de caso, utilizando os modelos previamente apresentados, aplicado ao setor elétrico brasileiro. Mais especificamente, serão modelados os FC de três usinas eólicas localizadas no Nordeste brasileiro.

6.1

Introdução

As usinas eólicas brasileiras em operação possuem uma série histórica relativamente curta pelo fato de estarem operando há, no máximo, 2 anos. Esse fato será contornado via utilização de um modelo de extensão de históricos de FC. Basicamente, esse modelo se baseia na metodologia de regressão linear múltipla para reconstruir os dados "faltantes", onde se utiliza inputs físicos pertinentes a estimação da variável em questão, o FC. Esses inputs são obtidos a partir de bancos de dados de reanálises globais internacionais que fornecem dados gerados aplicando modelos físicos a observações de satélites. Mais detalhes sobre esses modelos podem ser encontrados em Garcia e Street (2013).

Através do modelo de extensão, foram escolhidas 3 séries de FC de usinas localizadas no nordeste Brasileiro que apresentassem regimes eólicos similares, de forma a justificar o uso de modelagem multivariada via cópulas. As usinas escolhidas foram, Rio do Fogo (RF), Icaraizinho (IC) e Enancel (EN), que apresentam a seguinte matriz de distância conforme Tabela 6.1 e distribuição geográfica, conforme Figura 6.1.

Tabela 6.1: Matriz de distância entre as usinas eólicas.

Distância (km)	EN	IC	RF
EN	-	321	277
IC	321	-	590
RF	277	590	-

É importante salientar que, os dados que servem de input para o modelo de extensão de históricos são produzidos por observações de satélites. Logo usinas muito próximas umas das outras podem acabar produzindo séries muito similares de FC. Isso porque as características geográficas pertinentes as regiões possuem uma forte similaridade, e assim, por construção, as séries geradas possuem uma forte correlação.

Figura 6.1: Distribuição geográfica das três usinas. *Fonte: Google Maps.*

Segundo o Atlas do Potencial Eólico Brasileiro de Amarante *et al.* (2001), as três usinas estão localizadas na região geográfica denominada de *Zona Litorânea Norte-Nordeste*. Essa região geográfica possui as seguintes características, retiradas do atlas referenciado:

A Zona Litorânea Norte-Nordeste é definida como a faixa costeira com cerca de 100km de largura, que se estende entre o extremo norte da costa do Amapá e o Cabo de São Roque, no Rio Grande do Norte. Nessa região, os ventos são controlados primariamente pelos alísios de leste e brisas terrestres e marinhas. Essa combinação das brisas diurnas com os alísios de leste resulta em ventos médios anuais entre 5m/s e 7,5m/s na parte norte dessa região (litorais do Amapá e Pará) e entre 6m/s a 9m/s em sua parte sul, que abrange os litorais do Maranhão, Piauí, Ceará e Rio Grande do Norte. As velocidades são maiores na parte sul devido a dois principais fatores: (1) os ventos alísios geralmente tornam-se mais fortes à medida que se afastam da Depressão Equatorial; (2) as brisas marinhas são significativamente acentuadas ao sul dessa região em razão dos menores índices de vegetação e de umidade do

solo, fazendo que a superfície do solo atinja temperaturas mais elevadas durante as horas de sol e, conseqüentemente, acentuando o contraste de temperaturas terra-mar e as brisas marinhas resultantes. As maiores velocidades médias anuais de vento ao longo dessa região estão ao norte do Cabo de São Roque, abrangendo os litorais do Rio Grande do Norte e Ceará, onde a circulação de brisas marinhas é especialmente intensa e alinhada com os ventos alísios de leste-sudeste. Adicionalmente, ocorrem áreas em que os ventos são acentuados por bloqueios ao escoamento causados por montanhas na parte continental. Entretanto, o vento médio anual decresce rapidamente à medida que se desloca da costa para o interior, devido ao aumento de atrito e rugosidade de superfície e ao enfraquecimento da contribuição das brisas marinhas.

Conforme texto supracitado, a particular região em que estão inseridas as três usinas possui duas grandes vantagens no que se refere à produção eólica, (i) maior velocidade e (ii) localização livre de bloqueios. No que se refere ao primeiro ponto, conforme mencionado os ventos médios anuais ficam em torno de 6m/s a 9m/s, uma vez que essas três usinas estão sitiadas ao norte do Cabo de São Roque. Isso terá um impacto direto na modelagem em questão, dado que o fator de capacidade eólico, que é dado em % da potência instalada, é função da geração da usina que é conseqüentemente função da velocidade. Define-se o FC da usina j como

$$FC_t^{(j)} = \frac{Ger_t^{(j)}}{Pot^{(j)}} \cdot 100, \quad j = 1, 2, 3 \quad (6-1)$$

onde $Ger_t^{(j)}$ representa a geração da usina em MWh e $Pot^{(j)}$ a Potência Instalada da usina j em MW .

No que se refere ao segundo ponto, as usinas estão localizadas em locais elevados, livres de obstruções, o que maximiza ainda mais a velocidade do vento e conseqüentemente a sua geração. A Figura 6.2, também retirada de Amarante *et al.* (2001), é um mapa de relevo que basicamente aponta a região em vermelho como as regiões mais elevadas em termos de relevo.

Figura 6.2: Mosaico de imagens de satélite, sobreposto ao modelo de relevo retirado de Amarante *et al.* (2001).

Pelo mapa, conclui-se que as três usinas escolhidas partilham da mesma similaridade tanto no relevo, quanto na velocidade média anual destacada acima. Tal informação é um subsídio adicional ao se cogitar uma análise via cópulas, uma vez que possíveis relações entre o fator de capacidade das usinas não seriam espúrias e sim que possíveis relações são conjecturadas com base em suas similaridades geográficas.

6.2 Análise descritiva

Nessa sessão apresentamos uma análise descritiva das séries de FC das três usinas. Para cada usina, são disponibilizados 372 dados de FC mensal, obtidos através da aplicação do modelo de extensão de histórico já mencionado de Garcia e Street (2013). As séries têm início em janeiro de 1981, estendendo-se até dezembro de 2011. O último ano de dados, de cada uma das três séries, foi excluído da modelagem de forma a permitir uma posterior avaliação *out-of-sample* do modelo.

Inicialmente, são explorados, os gráficos da evolução do FC ao longo do tempo e o Box Plot mensal das séries, expostos pela Figura 6.3. Ambos os gráficos expõem o perfil sazonal das séries. Tal perfil é conhecido e possui um fato estilizado com a vazão, denominado de perfil de complementariedade, descrito a seguir. Quando chove muito, há pouco vento, logo as hidroelétricas figuram como principais fontes de energia alternativa. Já quando há poucas chuvas apesar das hidroelétricas ainda representarem grande parte da matriz energética, a energia eólica aparece como forte candidata a complementar essa fonte, pois os ventos estão no seu regime máximo.

Figura 6.3: Gráfico das séries de FC para as usinas de Rio do Fogo, Icaraizinho e Enacel, resultantes do modelo de extensão de histórico com seus respectivos Box-Plots mensais entre janeiro de 1981 a dezembro de 2011.

Com base no exposta na Figura 6.3, o perfil sazonal mencionado fica nítido. Pelos Box Plots, percebe-se que ao longo do ano, a série de FC eólico nas três usinas pode apresentar dois regimes. Na primeira metade do ano, de janeiro a junho, há uma forte redução na produção eólica, enquanto que na segunda metade do ano, entre os meses de julho a dezembro, há um aumento expressivo na geração eólica.

Outra questão interessante de verificar seria a não estacionariedade das

três séries temporais. Dentro do arcabouço dos modelos GAS, essa não é uma condição necessária para a modelagem e não há necessidade de obter a estacionariedade da série a priori. Apenas a título de ilustração, pelo teste da raiz unitária de Phillips–Perron, houve rejeição da hipótese nula de não estacionariedade nos três casos com p-valor menor que 0,01. Observar que esse teste, a priori, somente é estatisticamente válido para séries não sazonais. A não normalidade das mesmas também foi investigada pelo teste de Jarque-Bera e os resultados foram similares, havendo rejeição da hipótese nula de normalidade nos três casos com p-valor menor que 0,01.

A primeira conclusão obtida com base na Figura 6.3, valida a utilização do modelo Beta GAS SARIMA(p,q), uma vez que o domínio das observações está contida dentro do suporte da densidade preditiva e os dados apresentam sazonalidade. Além disso, como aponta a Figura 6.4, os dados possuem uma forte associação positiva, i.e., quando há uma forte produção em uma usina é provável que também o haja nas demais, sendo interessante quantificar essa informação e introduzi-la na análise via o uso de cópulas.

Figura 6.4: Gráfico de dispersão entre as três usinas.

6.3

Modelagem do FC das usinas via Beta GAS SARIMA(p,q)

O foco dessa sessão é a modelagem dos FC das três usinas, $\{y_{RF}, y_{IC}, y_{EN}\}$, via modelo Beta GAS SARIMA(p,q), conforme apresentado na Sessão 3. Como já foi referenciado, o suporte da densidade preditiva Beta é $[0, k^{(j)}]$, sendo adotado como $k^{(j)}$ o valor inteiro mais próximo do valor máximo da série de FC da usina j . Assim foram especificados $k^{(RF)} = 53$, $k^{(IC)} = 85$ e $k^{(EN)} = 69$, respectivamente.

As ordens p e q do modelo foram estabelecidas escolhendo a ordem que resultassem em resíduos descorrelatados. A modelagem iniciava a partir da introdução de variáveis *dummies* para acomodar observações mal ajustadas. Com o intuito de não alongar o texto, parte-se do ponto onde as *dummies*, para tratamento de resíduos atípicos, já foram introduzidas, dado que a normalidade dos resíduos só foi obtida após a introdução das mesmas. Posteriormente foram efetuados os diagnósticos padrões nos resíduos quantílicos: teste de Jarque Bera para normalidade e Ljung-Box até o *lag* 30 para autocorrelação e efeito ARCH. Os modelos que passassem nos diagnósticos de resíduos e apresentassem os melhores valores em termos de AIC, BIC e log da verossimilhança foram escolhidos. A Tabela 6.2 apresenta os valores dos critérios de informação e o log da verossimilhança dos modelos selecionados fixando a ordem $p = 12$ e $q = 12$.

Tabela 6.2: Logarítmo da verossimilhança e critérios de informação para os modelos Beta GAS SARIMA(p,q).

Estatística	RF	IC	EN
Log	-1030,2	-1148,5	-1119,2
AIC	2090,4	2323,0	2268,4
BIC	2148,2	2390,6	2326,2

Além disso, os valores do vetor de parâmetros estáticos θ , estimados via máxima verossimilhança, são apresentados na Tabela 6.3.

Na sequência, são apresentados os diagnósticos, já mencionados, utilizando os resíduos quantílicos associados a cada um dos modelos univariados. Quanto a normalidade, a introdução de variáveis *dummies* como variáveis exógenas no modelo Beta GAS SARIMA(p,q) se mostrou eficaz, não havendo em nenhum dos três casos a rejeição da hipótese nula de normalidade no teste Jarque Bera, atestando que a densidade preditiva beta é pertinente à modelagem do FC eólico. Além disso, a título de ilustração, na Figura 6.5 são apresentados os QQ-plots e os histogramas das três séries de resíduos quantílicos. Uma

Tabela 6.3: Estimativas de máxima verossimilhança e erros padrão das estimativas - modelos Beta univariados para as séries de FC das usinas RF, IC e EN para o período de janeiro de 1981 à dezembro de 2010.

Parâmetro	RF		IC		EN	
	Estim.	E.P.	Estim.	E.P.	Estim.	E.P.
A_1	0,551	0,006	0,483	0,060	0,481	0,047
A_2	0,231	0,004	0,463	0,022	-0,175	0,051
A_3	-0,147	0,007	0,179	0,043	0,321	0,047
A_{11}	0,008	0,007	0,325	0,066	0,211	0,034
A_{12}	0,110	0,004	0,196	0,059	-0,196	0,047
B_1	-0,080	0,002	0,077	0,069	0,848	0,043
B_2	0,441	0,006	-0,258	0,056	-0,659	0,034
B_3	-0,369	0,011	-0,125	0,049	0,120	0,019
B_{11}	0,207	0,013	-0,060	0,052	0,190	0,010
B_{12}	0,325	0,011	0,604	0,091	0,285	0,042
ω	3,139	0,019	6,352	0,336	2,227	0,209
α	9,408	0,042	10,275	0,688	8,937	0,747
$D_{t=21}^{(Set,1983)}$	-	-	2,741	0,329	-	-
$D_{t=22}^{(Out,1983)}$	-	-	2,932	0,252	2,341	0,348
$D_{t=129}^{(Set,1992)}$	-	-	3,168	0,327	1,082	0,366
$D_{t=308}^{(Ago,2007)}$	1,771	0,346	-	-	-	-
$D_{t=309}^{(Set,2007)}$	2,384	0,369	-	-	-	-
$D_{t=345}^{(Set,2010)}$	3,004	0,277	-	-	-	-
$D_{t=346}^{(Out,2010)}$	-	-	-1,266	0,356	-1,237	0,386

análise apurada nos mesmos revela que os três histogramas apresentam uma leve assimetria e que apenas o QQ-plot dos resíduos da usina de Rio do Fogo apresenta uma observação mal ajustada, não resultando em maiores problemas no que se refere à normalidade.

Porém, em se tratando de autocorrelação e efeitos ARCH, houveram algumas violações na FAC e FAC², conforme Figura 6.6.

Pelo comportamento apresentado na FAC dos resíduos quantílicos na Figura 6.6, em conjunto com os resultados apresentados na Tabela 6.4, observa-se uma estrutura de dependência na FAC dos resíduos quantílicos das usinas de IC e EN e na FAC² de IC. Uma investigação nas séries de resíduos ainda apontavam alguns *outliers* que poderiam ser os responsáveis pela presença desse comportamento. Entretanto mesmo após a inclusão de variáveis *dummies* adicionais nesses *outliers* remanecentes, as estruturas de dependência se mantiveram, piorando os resultados referentes a normalidade das séries de resíduos. Dessa forma, como os resultados com o conjunto de parâmetros apresentados na Tabela 6.3, foram satisfatórios do ponto de vista de previsibilidade, conforme será apresentado na Sessão 6.5, mantiveram-se esses como opção final.

Ainda, conforme levantado no Capítulo 4, antes de dar prosseguimento a modelagem multivariada na próxima sessão, é de suma importância realizar o teste de RV. Através desse teste, é definido qual será o conjunto de informação condicionante para a construção das pseudo-observações das marginais $\mathbf{u}_t =$

Figura 6.5: Histogramas e QQ-plots para dos resíduos quantílicos produzidos pelos modelos Beta GAS SARIMA(p,q) aplicados nas séries de FC das três usinas eólicas.

Tabela 6.4: P-valores para os testes de normalidade, autocorrelação e efeitos ARCH dos resíduos quantílicos produzidos pelos modelos univariados para as séries de FC das usinas RF, IC e EN.

Teste	RF	IC	EN
Normalidade	0,242	0,8535	0,259
Autocorrelação	0,174	0,000	0,000
ARCH	0,703	0,035	0,905

Figura 6.6: FAC e FAC^2 das séries de resíduos quantílicos associados aos FC eólicos das usinas RF, IC e EN.

(u_{1t}, u_{2t}, u_{3t}) para $t = 13, 2, \dots, 360$. Com o intuito de simplificar a notação, a partir desse momento, será adotada a seguinte nomenclatura: u_{1t} se refere as pseudo-observações de RF, u_{2t} para IC e u_{3t} para EN.

Após a realização do referido teste, conforme o passo a passo descrito na Sessão 4.2.1: estima-se o modelo marginal i condicionado apenas ao subconjunto de informação $\mathcal{F}_{i,t-1}$ e posteriormente ao conjunto de informação $\mathcal{F}_{t-1}$ com os lags 1, 2, 3, 11 e 12 das demais variáveis tomados como não nulos. Os resultados obtidos indicam que não houve rejeição da hipótese nula, de que cada variável pode ser condicionada apenas a seu respectivo subconjunto de informações, em nenhum dos três casos, conforme Tabela 6.5.

Tal resultado dá subsídio para concluir que, com 95% de significância, não há necessidade da inclusão de todo o conjunto de informação para a estimação dos parâmetros das três marginais. Condicionar cada marginal apenas a seu respectivo conjunto de informação passada para estimar seus respectivos parâmetros já é suficiente.

Tabela 6.5: Teste da Razão de Verossimilhança (RV).

Usina	nº de parâmetros em $\mathcal{F}_{t-1}$	nº de parâmetros em $\mathcal{F}_{i,t-1}$	Λ	p-valor
RF	25	15	1,787	0,997
IC	26	16	1,740	0,998
EN	25	15	2,092	0,995

6.4

Modelagem da estrutura de dependência via GAS t -Student(p,q)

O modelo Beta GAS SARIMA(p,q) fornece uma descrição adequada acerca do comportamento mensal do FC individual de cada uma das três usinas em questão. Entretanto, é possível que esses FC possuam algum tipo de dependência estocástica. Sendo esse o caso, devemos modelar essa dependência e introduzi-la na geração de cenários de cada série individual.

Inicialmente, aproveitando o resultado da sessão anterior acerca do teste da RV, serão construídas as pseudo-observações, $\mathbf{u}_t$, a partir das marginais condicionadas apenas a seu respectivo conjunto de informação $\mathcal{F}_{i,t-1}$ para $i = 1, 2, 3$. Vale ressaltar que as variáveis *dummies* estimadas, também fazem parte desse conjunto e irão alterar os valores de $\mathbf{u}_t$, uma vez que a PIT da marginal i é dada por

$$u_{it} = F_{Beta}(y_{it}|f_{i,t}, \mathcal{F}_{i,t}, \theta_i) \text{ para } i = 1, 2, 3 \text{ e } t = 13, \dots, 360. \quad (6-2)$$

Conforme os pressupostos para a modelagem de cópula, as variáveis em $\mathbf{u}_t$ devem ser uniformes e individualmente devem ser *i.i.d.* Para avaliar o primeiro ponto, foram conduzidos dois testes de aderência à distribuição uniforme, Kolmogorov-Smirnov e Anderson-Darling, onde $H_0 : u_{it} \sim U[0, 1]$ para $i = 1, 2, 3$. Em nenhum dos três casos houve rejeição da hipótese nula, concluindo que nesse aspecto o ajuste das marginais foi adequado conforme Tabela 6.6.

Tabela 6.6: P-valor dos testes de aderência à distribuição uniforme das variáveis PIT.

Teste	RF	IC	EN
KS	0,353	0,869	0,285
AD	0,343	0,768	0,312

O segundo ponto já foi de alguma maneira testado ao aplicarmos o teste

¹ Começando em $t = 13$ porque os resíduos quantílicos não são definidos para $t = 1, \dots, 12$.

de autocorrelação nos resíduos quantílicos produzidos pelas marginais. Porém com o intuito de dar maior subsídio a essa decisão, outro teste é realizado. As variáveis PIT foram transformadas em variáveis Gaussianas $x_{it} = \Phi^{-1}(u_{it})$, onde Φ^{-1} denota o inverso da F.D. de uma normal padrão, e posteriormente é conduzido um teste BDS, que tem como hipótese nula (H_0) a classificação da variável aleatória como sendo *i.i.d.* e hipótese alternativa (H_1) a presença de dependência linear ou não linear na variável aleatória. O teste é denotado por $C(m, \epsilon)$ com $m = \{2, 3, 4, 5\}$ e $\epsilon = \{0.5, 1, 1.5, 2\}\hat{\sigma}_i$, onde $\hat{\sigma}_i$ é o desvio padrão da série de resíduos quantílicos.

Tabela 6.7: Teste BDS na variável x_{1t} referente à usina RF.

$C(m, \epsilon)$	2	3	4	5
$0.5\hat{\sigma}_1$	0,589	0,717	0,927	0,754
$1\hat{\sigma}_1$	0,240	0,932	0,825	0,829
$1.5\hat{\sigma}_1$	0,223	0,643	0,824	0,473
$2\hat{\sigma}_1$	0,457	0,292	0,674	0,277

Tabela 6.8: Teste BDS na variável x_{2t} referente à usina IC.

$C(m, \epsilon)$	2	3	4	5
$0.5\hat{\sigma}_2$	0,313	0,123	0,096	0,012
$1\hat{\sigma}_2$	0,388	0,274	0,2653	0,060
$1.5\hat{\sigma}_2$	0,029	0,222	0,213	0,064
$2\hat{\sigma}_2$	0,041	0,223	0,255	0,126

Tabela 6.9: Teste BDS na variável x_{3t} referente à usina EN.

$C(m, \epsilon)$	2	3	4	5
$0.5\hat{\sigma}_3$	0,393	0,261	0,396	0,370
$1\hat{\sigma}_3$	0,611	0,294	0,385	0,395
$1.5\hat{\sigma}_3$	0,885	0,256	0,334	0,415
$2\hat{\sigma}_3$	0,895	0,342	0,4281	0,510

Conforme os resultados apresentados nas Tabelas 6.7, 6.8 e 6.9, é possível concluir que as PIT estão bem ajustadas e que de fato não podemos rejeitar a hipótese de que sejam variáveis aleatórias *i.i.d.*, indicando que os modelos marginais foram corretamente especificados. O próximo passo é quantificar o nível de concordância entre as variáveis aleatórias $\mathbf{u}_t$. Serão utilizados os

coeficientes de correlação não-paramétricos de Spearman (ρ^S) e o de Kendall (ρ^τ), que se baseiam nos postos das observações. Os valores desses coeficientes servem como medida de concordância entre as variáveis. Caso as mesmas não possuem dependência conjunta, não faria sentido uma modelagem por cópula para geração de cenários integrados. Bastaria levantar as respectivas distribuições de FC individualmente para cada usina. O resultado desses coeficientes se mostram significativos em todos os casos (p-valor<0,001), apontando para uma dependência positiva entre as três variáveis $\mathbf{u}_t$. As Tabelas 6.10 e 6.11 apresentam os respectivos estimadores dos coeficientes.

ρ^S	EN	IC	RF
EN	-	0,702	0,594
IC	0,702	-	0,575
RF	0,594	0,575	-

Tabela 6.10: Matriz dos coeficientes de correlação de Spearmann.

ρ^τ	EN	IC	RF
EN	-	0,515	0,423
IC	0,515	-	0,407
RF	0,423	0,407	-

Tabela 6.11: Matriz dos coeficientes de correlação de Kendall.

Tendo ciência de que as variáveis $\mathbf{u}_t$ possuem associação umas com as outras, e que além disso se movimentam no mesmo sentido, o próximo passo é postular qual a distribuição conjunta apropriada para as pseudo-observações $\mathbf{u}_t$. Conforme mencionado em Mendes (2004), apesar das simulações através de cópulas contornarem o difícil problema da especificação da distribuição multivariada de uma carteira, a identificação da cópula pertinente à distribuição conjunta dos dados reais não é tarefa fácil ou óbvia.

No contexto dessa dissertação, o principal objetivo é estudar a especificação do modelo de Creal *et al.* (2011) no arcabouço de cópulas. À vista disso não foi ajustada nenhum outro tipo de cópula fora da classe da família de distribuições elípticas. Serão adotadas a cópula *t-Student* e a cópula Gaussiana como distribuição conjunta das pseudo-observações $\mathbf{u}_t$.

Inicialmente, efetuando uma inspeção visual na Figura 6.7, é passível inferir que, nos diagramas de dispersão entre as variáveis u_1 x u_3 e u_2 x u_3 parece haver uma concentração de eventos nos quadrantes superiores direito e quadrantes inferiores esquerdo, e além disso há uma associação positiva entre essas variáveis. Já o diagrama de dispersão entre u_1 x u_2 , apenas indica uma dependência positiva, não havendo uma concentração tão relevante de eventos nos quadrantes mencionados anteriormente. Tal característica pode indicar que a cópula Gaussiana seja adequada para caracterizar a dependência entre

os FC dessas duas usinas. Uma característica similar às três combinações de distribuições bivariadas é a simetria.

Figura 6.7: Diagrama de dispersão das pseudo-observações $\mathbf{u}_t$, produzidas pelos modelos univariados duas a duas. Na primeira linha estão os diagramas de dispersão entre as pseudo-observações de EN e IC na esquerda e o diagrama de dispersão entre as pseudo-observações de EN e RF na direita. Na segunda linha encontra-se o diagrama de dispersão entre as pseudo-observações de IC e RF.

Adotou-se a visualização da estrutura de dependência duas a duas para facilitar a visualização de características pertinentes à distribuição multivariada adequada para caracterizar a dependência entre as usinas, tais como simetria e concentração de eventos nos quadrantes. A visualização dos mesmos para um caso multivariado, com dimensão maior do que 3, não seria possível. A Figura 6.8 apresenta o gráfico de dispersão para $\mathbf{u}_t$ no arcabouço trivariado. É importante notar que as conclusões obtidas anteriormente, para a estrutura de dependência das pseudo-observações duas a duas, já se tornam mais árduas de se interpretarem.

Dando continuidade à estimação do modelo apresentado no Capítulo 5, será adotada a mesma nomenclatura de Creal *et al.* (2011) para o modelo de cópula *t-Student*, modelo *t-GAS*. Neste modelo, o grau de liberdade ν é fixado em $\nu = \{4, 5, \dots\}$ e o restante dos parâmetros contidos em θ são estimados utilizando a função de verossimilhança da *t-Student* multivariada,

Figura 6.8: Diagrama de dispersão 3D das pseudo-observações produzidas pelos modelos Univariados.

dada pela Equação (5-15). Esse procedimento é conhecido como estimação da verossimilhança perfilada e conforme já ressaltado na Sessão 5.6, tal artifício foi adotado de maneira a evitar a construção de um algoritmo iterativo para a estimação de ν simultaneamente com os demais parâmetros.

Adotando os mesmos conceitos já explicados na Sessão 5.6, acerca dos procedimentos que devem ser tomados para a inicialização da otimização dos parâmetros no modelo GAS t -Student(p, q), é estimado o referido modelo t -GAS para o vetor de parâmetros $\tilde{\mathbf{y}}_t$,

$$\tilde{\mathbf{y}}'_t = [F_\nu^{-1}(u_{1t}), F_\nu^{-1}(u_{2t}), F_\nu^{-1}(u_{3t})]' \text{ para } t = 13, \dots, 360. \quad (6-3)$$

As ordens p e q desse modelo foram ambas fixadas em 1. Definindo as matrizes A e B como diagonais, i.e., $A = a \cdot I$ e $B = b \cdot I$, e selecionando dentre todos aqueles modelos estimados, aquele que apresente os melhores critérios de informação, obtêm-se os resultados da Tabela 6.12.

Optando pelo modelo com melhores critérios de informação, é selecionado o modelo t -GAS com $\hat{\nu} = 340$, sendo que esse é o valor de ν que maximiza o log da verossimilhança. Além disso para $\nu > 340$, a otimização não encontra solução viável. Foram testados outros valores de ν que não foram expostos na Tabela 6.12 com o intuito de não estender muito esse resultado pois não eram relevantes. Porém é importante salientar que uma densidade t -Student com $\hat{\nu} = 340$ tende a se aproximar de uma normal padrão. Para elucidar esse fato,

Tabela 6.12: Verossimilhança e critérios de Informação para o modelo GAS t -Student(p,q).

Modelo	ν	Log	AIC	BIC
t -GAS	4	-1505,834	3017,668	3029,225
	5	-1433,943	2873,886	2885,443
	6	-1390,405	2786,811	2798,367
	7	-1361,141	2728,282	2739,839
	8	-1340,128	2686,256	2697,813
	9	-1324,326	2654,652	2666,209
	10	-1312,024	2630,048	2641,604
	11	-1302,181	2621,918	2610,361
	12	-1294,135	2594,271	2605,827
	13	-1287,443	2580,886	2592,443
	14	-1281,797	2569,583	2581,139
	15	-1276,972	2559,944	2571,534
	16	-1272,789	2551,578	2563,135
	17	-1269,142	2544,285	2555,841
	18	-1265,947	2537,894	2549,451
	19	-1263,089	2532,177	2543,734
	20	-1260,552	2527,103	2538,660
	30	-1244,981	2495,962	2507,518
	40	-1237,560	2481,120	2492,676
	50	-1233,255	2472,509	2484,066
	60	-1230,417	2466,835	2478,391
	70	-1228,427	2462,854	2474,411
	100	-1224,908	2455,816	2467,373
	150	-1222,268	2450,535	2462,092
	200	-1220,914	2447,827	2459,384
	330	-1219,419	2444,838	2456,395
	340	-1219,353	2444,706	2456,262

geramos 1000 valores de uma normal padrão e 1000 valores de uma t -Student com $\hat{\nu} = 340$, e apresentamos suas respectivas densidades na Figura 6.9.

Como já comentado anteriormente e apresentado no Apêndice B, o modelo GAS t -Student(p,q) se reduz a um modelo normal multivariado quando $\nu^{-1} \rightarrow 0$. Fato que é comprovado pelas densidades da Figura 6.9. Um grau de liberdade alto para a densidade t -Student já era um resultado esperado, uma vez que o conjunto de dados é mensal, reduzindo dessa forma os efeitos de caudas pesadas.

O gráfico da Figura 6.10 apresenta a dinâmica do coeficiente de ponderação w_t , apresentado no Teorema 5.3.1, ao longo do tempo para o respectivo conjunto de dados com $\hat{\nu} = 340$. De acordo com Creal *et al.* (2011), ao considerar a densidade preditiva uma normal multivariada ($\nu^{-1} \rightarrow 0$), o coeficiente de ponderação tomaria a forma $w \equiv 1$, o que é observado empiricamente.

Figura 6.9: Comparação entre a densidade de uma normal padrão e uma densidade *t-Student* com $\nu = 340$.

Com relação à otimização, como era de se esperar, a restrição descrita anteriormente, na Sessão 5.6.1, de que $1 > B_1 \geq A_1 \geq 0$, é atendida conforme Tabela 6.13. Como já explicitado na Sessão 4.2.3, o cálculo dos erros padrão dos estimadores dos parâmetros da cópula não levam em consideração os erros padrão dos estimados dos parâmetros das marginais, sendo portanto subavaliados.

Atenta-se para a estimativa do parâmetro A , o qual introduz a informação vinculada ao *score* nos elementos da matriz Σ_t . Caso $A = 0 \cdot I$, a dinâmica do coeficiente de correlação entre as variáveis pode ser tomada como constante ao invés de variante no tempo, conforme mencionado em Patton (2009). Pelo resultado da Tabela 6.13, como o valor do parâmetro A estimado é relativamente alto e ainda se mostrou significativo, não pode ser descartada uma análise com o coeficiente de correlação variante no tempo.

Tabela 6.13: Estimativas de máxima verossimilhança e erro padrão das estimativas - GAS *t-Student*(p,q).

Parâmetro	<i>t-GAS</i>	
	Estim.	E.P.
ω	1,589	0,063
A_1	0,354	0,017
B_1	0,425	0,083

As Figuras 6.11, 6.12 e 6.13 apresentam as correlações $\hat{\rho}_{t|t-1}^{(i,j)}$ variantes no tempo atualizadas pelo modelo GAS *t-Student*(p,q), duas a duas, entre $\tilde{y}_{it}$ e $\tilde{y}_{jt}$ contidas em $\tilde{\mathbf{y}}_t$ para $(i, j) = 1, 2, 3$ e $i \neq j$. A linha pontilhada constante

Figura 6.10: Dinâmica de w_t ao longo do tempo no modelo de cópula t -Student com $\hat{\nu} = 340$.

nos gráficos representa a correlação incondicional entre as variáveis contidas em $\tilde{\mathbf{y}}_{it}$.

Como era de se esperar, a correlação entre as variáveis se manteve alta, em grande parte do tempo $t = 13, \dots, 360$. Vale lembrar que as três usinas partilhavam de características geográficas similares e que o coeficiente de correlação incondicional entre as marginais era alto, logo era de se esperar um resultado que indicasse uma correlação significativa entre as pseudo-observações ao longo do tempo.

Figura 6.11: Elementos de $\hat{\Sigma}_t$ atualizados pelo mecanismo GAS entre $\tilde{y}_{1t}$ e $\tilde{y}_{2t}$, que representam as usinas de RF e IC respectivamente.

Figura 6.12: Elementos de $\hat{\Sigma}_t$ atualizados pelo mecanismo GAS entre $\tilde{y}_{1t}$ e $\tilde{y}_{3t}$, que representam as usinas de RF e EN respectivamente.

Figura 6.13: Elementos de $\hat{\Sigma}_t$ atualizados pelo mecanismo GAS entre $\tilde{y}_{2t}$ e $\tilde{y}_{3t}$, que representam as usinas de IC e EN respectivamente.

O comentário a ser feito acerca das Figuras 6.11, 6.12 e 6.13 é referente a captura do nível da correlação incondicional entre as três variáveis, pelo modelo *t-GAS*. Neste modelo, as estruturas de correlação apresentaram um padrão semelhante, sendo que para os cruzamentos entre $\tilde{y}_{1t} \times \tilde{y}_{2t}$ e $\tilde{y}_{1t} \times \tilde{y}_{3t}$, o nível da correlação condicional ficou acima da correlação incondicional em grande parte do tempo. Já para $\tilde{y}_{2t} \times \tilde{y}_{3t}$, o contrário foi observado, sendo o nível da correlação condicional menor que a incondicional.

Vale lembrar que os dados na análise são mensais e dessa forma, não há muita variabilidade nos mesmos. Em outros contextos, como em dados de retorno financeiro, conforme descrito em Embrechts *et al.* (2005), a medida

em que é aumentado o intervalo de medição dos dados para semanal, mensal, trimestral ou anual, os agrupamentos de volatilidade, fato estilizado para séries de retorno financeiro, tendem a diminuir e os retornos tendem a se tornar menos dependentes com densidades apresentando caudas mais leves. Esse fato pode ser um dos responsáveis pela falta de variações bruscas nos coeficientes de correlação fazendo com que a dinâmica de dependência entre essas variáveis não se altere bruscamente, uma vez que estão se movendo sempre na mesma direção, conforme concluído pelos coeficientes de concordância de Kendall e Spearmann.

Pelos resultados apresentados nas Tabelas 6.12 e 6.13, o modelo t -GAS se mostra adequado para a modelagem da estrutura de dependência contida em $\tilde{\mathbf{y}}_t$, sendo que não houve rejeição da hipótese nula para nenhuma das cópulas postuladas, gaussiana e t -Student com $\nu = 340$. Entretanto a evidência é mais forte para a t -Student, embora com $\hat{\nu} = 340$, a densidade tende a se aproximar de uma Gaussiana.

Na próxima Sessão, será feita uma previsão utilizando os modelos marginais com e sem a introdução da estrutura de dependência modelada pela cópula. A partir dessas informações é possível realizar uma análise de risco na distribuição de renda dos contratos vinculados aos perfis de geração das três usinas.

6.5 Previsão Condicional

Tendo em mente os conceitos de cópula condicional, é importante iniciar essa sessão ressaltando que qualquer melhora nos critérios de avaliação da previsão da média condicional k passos à frente é meramente devido a uma flutuação estatística. Para ilustrar esse conceito, considere o seguinte processo retirado de Patton (2012),

$$Y_{it} = \mu_i(\mathbf{Z}_{t-1}) + \sigma_i(\mathbf{Z}_{t-1})\varepsilon_{it}, \text{ para } i = 1, 2, \text{ onde } \mathbf{Z}_{t-1} \in \mathcal{F}_{t-1} \quad (6-4)$$

$$\varepsilon_{it}|\mathcal{F}_{t-1} \sim F_i(0, 1) \quad \forall t. \quad (6-5)$$

Dessa forma, como a informação modelada via o uso de cópulas é introduzida nas inovações do modelo de série temporal ε_{it} , a média condicional se manterá inalterada para a simulação de cenários em todo o instante de tempo t . No arcabouço dessa dissertação, conforme elucidado no Capítulo 5.5, a partir da informação da cópula, serão calculados os quantis da densidade preditiva k passos à frente que trazem a informação da estrutura de dependência. Dessa forma, a média condicional também não sofrerá alterações.

Será adotada a nomenclatura *independente*, para as previsões realizadas individualmente pelos modelos univariados Beta GAS SARIMA(p,q), não contemplando a informação de dependência modelada pelo modelo GAS *t-Student*(p,q). Quando contemplada essa informação de dependência, será atribuída a nomenclatura de Beta *t-GAS*.

Como apresentado nas sessões 3.4 e 5.5, a previsão dessa classe de modelos é feita via simulação Monte Carlo. Seguindo o passo a passo apresentado nas sessões referenciadas com $k = 1, \dots, 12$ e $m = 1, \dots, 2000$, é construída para cada instante de tempo k uma densidade de previsão.

Além dos modelos GAS, é proposto o modelo *VARX* de Street *et al.* (2012) referenciado na introdução para a modelagem conjunta das três séries de FC. Os resultados obtidos a partir do uso desse modelo servirão como *benchmark* para as previsões dos modelos GAS. Porém, para manter uma base de comparação justa, para a estimação desse modelo nessa dissertação, não foram introduzidas variáveis exógenas, ao contrário do artigo supracitado que inclui como variáveis exógenas as vazões dos rios.

Para avaliar a capacidade preditiva 12 passos à frente geradas pela metodologia, foram excluída as 12 últimas observações da amostra para posterior comparação. Esse banco de dados corresponde aos FC das usinas de RF, IC e EN entre janeiro a dezembro de 2011, e assim, $k = 1, 2, \dots, 12$. A comparação é feita utilizando as seguintes medidas de aderência, RMSE (*Root Mean Square Error*), MAE (*Mean Absolute Error*) e o pseudo R^2 , calculados da seguinte maneira:

$$RMSE = \sqrt{\frac{1}{12} \sum_{k=1}^{12} (\hat{y}_{t+k|t} - y_{t+k})^2}, \quad k = 1, \dots, 12 \quad (6-6)$$

$$MAE = \frac{1}{12} \sum_{k=1}^{12} |\hat{y}_{t+k|t} - y_{t+k}|, \quad k = 1, \dots, 12 \quad (6-7)$$

$$pseudo R^2 = [corr(\hat{y}_{t+k|t}, y_{t+k})]^2, \quad k = 1, \dots, 12. \quad (6-8)$$

Após aplicar os passos decritos nas Sessões 3.4 e 5.5, obtendo dessa forma as $m = 1, \dots, 2000$ trajetórias para $k = 1, \dots, 12$ passos à frente, calcula-se a média das 2000 trajetórias a cada passo à frente k . Esse será o estimador das previsões $\hat{y}_{t+k|t}$ para o modelo independente e para o modelo Beta *t-*

GAS. Posteriormente são calculadas as medidas de aderência supracitadas e os resultados são apresentados pela Tabela 6.14.

Tabela 6.14: Medidas de aderência das Previsão fora da amostra k -passos frente, levantadas por simulação dos modelos Beta Gas SARIMA(p,q), Beta Gas SARIMA(p,q) com GAS t -Student(p,q) e *VARX*.

Modelo	Medidas	RF	IC	EN
Independente	RMSE	5,639	10,710	5,994
	MAE	5,209	9,244	4,990
	pseudo R2	0,734	0,822	0,915
Beta t -GAS	RMSE	5,979	10,529	6,030
	MAE	5,191	9,049	5,038
	pseudo R2	0,736	0,826	0,914
<i>VARX</i>	RMSE	5,709	7,087	6,581
	MAE	5,270	5,403	5,808
	pseudo R2	0,783	0,912	0,903

Com base nos resultados expostos na Tabela 6.14, é possível afirmar que o modelo GAS sem a modelagem da estrutura de dependência (independente), apresentou uma melhor capacidade preditiva em duas (RF e EN) das três usinas, no que se refere aos critérios de previsão nessa análise *out-of-sample*. Esse resultado vai de acordo com o que foi encontrado em Matos (2013), onde os modelos GAS obtiveram uma capacidade preditiva superior às do modelo *VARX*, inclusive com a introdução de variáveis exógenas (vazão dos rios).

Os gráficos dessas previsões são apresentados na Figura 6.14. É importante salientar que os intervalos de confiança que aparecem nos gráficos são referentes às previsões do modelo Beta GAS SARIMA(p,q). Nesses gráficos, os resultados obtidos via uso da informação modelada pelas cópulas serão omitidos, uma vez que esses são praticamente iguais às previsões independentes. Isso porque, como já comentado, a informação trazida pela cópula não possui influência relevante na média condicional. As densidades de previsão, levantadas por simulação Monte Carlo, para cada instante de tempo $k = 1, 2, \dots, 12$ nessa análise serão insumos importantes para os resultados da próxima Sessão 6.6.

Figura 6.14: Previsões 12 passos à frente dos FC pelos modelos Beta GAS SARIMA(p,q) e VARX para o período de janeiro à dezembro de 2011.

6.6

Análise de risco no Ambiente de Comercialização Livre (ACL)

Nessa sessão será feita uma análise de risco para um contrato de quantidade de um ano, referente ao período de janeiro a dezembro de 2011, no Ambiente de Comercialização Livre (ACL), supondo um portfólio composto pelas três usinas, RF, IC e EN. Além disso também serão estudados os contratos de cada usina individualmente. Não é o objetivo dessa dissertação discutir sobre a comercialização de energia elétrica e sim a de modelar a dependência estocástica entre as séries de FC das três usinas em questão via aplicação do modelo GAS t -Student(p,q). Logo, maiores detalhes e discussões sobre o ACL e outras definições sobre o mercado de energia elétrico brasileiro podem ser encontradas em Freire *et al.* (2015).

Os Fatores de Capacidade (FC) de RF, IC e EN foram modelados pelos modelos marginais Beta GAS SARIMA(p,q), com e sem estrutura de dependência modelada pelo modelo GAS t -Student(p,q). Posteriormente, iremos utilizar simulação Monte Carlo para levantar uma distribuição de renda a partir dessas variáveis. Deve-se inicialmente transformar o FC, que representa a geração em (%) de potência, para geração em (%) do lastro² da usina ($\tilde{G}_t$). Para tal, basta isolar a geração na equação (6-1) e multiplicar pela razão entre potência e lastro como,

$$\tilde{G}_t = \frac{FC_t}{100} \times \frac{Pot}{Lastro}. \quad (6-9)$$

Para tal se utilizam os seguintes valores de lastro³, $Lastro^{RF} = 20,74$, $Lastro^{IC} = 20,76$ e $Lastro^{EN} = 7,93$. Define-se $I = \{1, 2, 3\}$ subconjunto de usinas, $S = \{1, \dots, 2000\}$ os cenários modelados sem a estrutura de dependência da cópula, $C = \{1, \dots, 2000\}$ os cenários modelados com a estrutura de dependência da cópula, e $T = \{1, \dots, 12\}$ o número de períodos, ou passos à frente em que os cenários são avaliados.

Em posse das variáveis $\tilde{G}_{t,i,v}$ com $v \in V$ e $V = \{S, C\}$, a distribuição de renda estocástica de um contrato de quantidade padrão de venda no ACL para uma determinada usina i durante determinado período de tempo t , é construída substituindo essa variável na seguinte equação,

$$\tilde{R}_{t,i,v} = P \times Q \times h_t + (\tilde{G}_{t,i,v} - Q) \times \tilde{\pi}_{t,v} \times h_t \text{ para } t = 1, 2, \dots, 12 \quad (6-10)$$

onde $\tilde{\pi}_{t,v}$ representa os valores do preço de curto prazo de energia (popular-

²Limite máximo que a usina pode contratualmente comercializar.

³Fonte: <http://www.aneel.gov.br/aplicacoes/capacidadebrasil/energiaassegurada.asp>

mente conhecido como preço *spot*) simulados pelo modelo Newave (Pereira e Pinto, 1991) para o ano de 2012, P é o preço do contrato, h_t é o número de horas no período t e Q é a quantidade de energia comercializada de cada usina. No contexto dessa dissertação, Q e P foram tomadas como variáveis determinísticas onde o agente gerador compromete-se a entregar $Q = 1 \text{ MW médios}$ a $P = 100R\$/MWh$.

Agora, a partir dos valores de $\tilde{R}_{t,i,v}$, é construída a distribuição da renda anual de cada um dos conjuntos de cenários $s \subseteq S$ e $c \subseteq C$ para $t = 1, 2, \dots, 12$, trazendo a valor presente o valor da receita atrelada ao contrato a partir da seguinte expressão,

$$\tilde{R}_{i,v} = \sum_{t \in T} \frac{\tilde{R}_{t,i,v}}{(1 + K)^t}, \quad i = 1, 2, 3 \quad (6-11)$$

onde K seria a taxa de juro. Nesse caso será fixado $K = 0$.

Além das três distribuições $\{\tilde{R}_{i,v}\}_{v=1}^{2000}$ definidas, parte-se para composição da distribuição de renda anual supondo um portfólio com as três usinas da análise. Para tal, é necessário somar as gerações das três usinas produzidas nos 2000 cenários, de maneira a substituir na Equação (6-10) a variável $\tilde{G}_{t,i,v}$ por $\sum_{t \in T} \sum_{i \in I} \tilde{G}_{t,i,v}$.

Posteriormente, para obter os valores da renda do portfólio, basta somar os valores das distribuições de renda das usinas individuais, definidos pela Equação (6-10), para cada período t como,

$$\tilde{R}_{t,v}^{Port} = \sum_{t \in T} \sum_{i \in I} \tilde{R}_{t,i,v}. \quad (6-12)$$

Lembrando que a quantidade comercializada na equação (6-10) para o portfólio é a soma das quantidades individuais, logo $Q = 3$.

Em posse das distribuições $\{\tilde{R}_{i,v}\}_{v=1}^{2000}$ e $\{\tilde{R}_{t,v}^{Port}\}_{v=1}^{2000}$ com $V \in \{S, C\}$, a próxima etapa é avaliar as distribuições em termos de risco financeiro. Para tal, julga-se adequado utilizar a métrica do *Conditional Value at Risk* (CVaR) ao invés do *Value at Risk* (VaR), assim como feito em Freire *et al.* (2015) e Street *et al.* (2012). Isso porque o VaR (Jorion, 2006) apenas determina a perda mínima que não será excedida com probabilidade α , onde α representa um determinado nível de significância. Enquanto que o CVaR quantifica a perda esperada nos $(1 - \alpha)\%$ piores casos.

De maneira simplificada, o valor do CVaR representa a média dos $(1 - \alpha)\%$ piores cenários da distribuição de probabilidade da variável aleatória em estudo. Além disso, ao contrário do VaR, o CVaR é uma medida de risco

coerente, satisfazendo a propriedade de subaditividade ⁴ (Artzner *et al.*, 1999). Matematicamente,

$$VaR^\alpha(\tilde{R}) = \inf\{r \mid \Pr(\tilde{R} \leq r) \geq 1 - \alpha\} \quad (6-13)$$

$$CVaR^\alpha(\tilde{R}) = E[\tilde{R} \mid \tilde{R} \leq VaR^\alpha(\tilde{R})]. \quad (6-14)$$

Dando continuidade a análise, calculando o CVaR das distribuições supracitadas, tem-se os seguintes resultados expostos na Tabela 6.15.

Tabela 6.15: CVaR das distribuições de renda vinculados a um contrato de um ano (2011) no ACL via uso dos modelos GAS. Os valores estão em mil R\$.

$CVaR^{95\%}$	Dependente	Independente
$\tilde{R}_{1,v}$	-967,06	-1.051,00
$\tilde{R}_{2,v}$	490,31	316,79
$\tilde{R}_{3,v}$	622,79	467,27
$\tilde{R}_v^{Port}$	1.224,39	1.093,95

Damos ênfase ao resultado do CVaR de $\tilde{R}_v^{Port}$, o qual julga se a modelagem da estrutura de dependência é benéfica. Apesar da periodicidade dos dados de FC ser mensal, o que acaba por reduzir as probabilidades nas caudas da distribuição conjunta, é possível concluir pelos resultados da Tabela 6.15 que é vantajoso modelar a estrutura de dependência. Ao ser considerada na modelagem, a dependência entre os FC de diferentes parques eólicos, obtém-se um aumento de 12% nessa medida de risco, quando comparada ao da distribuição que não contempla a modelagem conjunta. Em termos práticos, isso significa que o investidor irá experimentar um aumento de 12% no valor com o qual ele enxerga seu investimento, ou seja, o valor pelo qual ele fica indiferente em trocar pelo seu investimento. Além disso, o CVaR da distribuição de renda do contrato atrelado a usina de IC, $\tilde{R}_{2,v}$, apresentou a maior diferença quando comparados os resultados dos modelos com e sem dependência, tendo um aumento de 54,7% na métrica de risco ao considerar a modelagem da dependência conjunta. Esses resultados dão subsídios para a metodologia proposta nessa dissertação, via aplicação dos modelos GAS para simular cenários a partir de uma densidade conjunta de FC.

Entretanto, vale destacar que a natureza das séries de FC apresentam algumas limitações, as quais prejudicam a modelagem conjunta via cópulas. Inicialmente, existe uma suavização nas séries temporais de FC, o que reduz

⁴Essa propriedade afirma que o CVaR da renda conjunta entre dois portfólios fundidos é maior ou igual à soma dos CVaR das rendas dos portfólios considerados individualmente.

os efeitos de probabilidade nas caudas da distribuição conjunta, justamente onde a modelagem por cópulas traria o principal benefício na análise de risco. Denomina-se suavização porque um ponto da série temporal de FC é calculado a partir da média da produção horária e posteriormente a média dessas produções diárias para obter a média mensal. Ademais, essas séries apresentam um bom comportamento, i.e., não há eventos extremos com picos de volatilidade. Tal argumento ganha sustentação ao estimar o grau de liberdade da cópula *t-Student*, $\hat{\nu} = 340$ como o valor que minimiza o log verossimilhança, especificando dessa forma uma cópula Gaussiana para a modelagem conjunta da estrutura de dependência dos dados.

6.7

Análise dos cenários condicionais gerados

Ressaltando que o objetivo dessa dissertação não é o de julgar qual o melhor modelo para a simulação conjunta dos FC das três usinas em estudo. Será apresentada nessa sessão uma análise descritiva dos cenários de FC gerados para o ano de 2011, a partir do modelo *VARX* e do modelo Beta GAS SARIMA(p,q) com estrutura de dependência modelada por GAS *t-Student*(p,q), ou Beta *t-GAS*. A aderência dos modelos será julgada visualmente via gráficos de estatísticas descritivas das simulações contra a série histórica. Para tal, segmenta-se a série original em 12 séries, onde serão separados os valores mensais, i.e., $\{y_t^{jan}\}_{t \in T}$, ..., $\{y_t^{dez}\}_{t \in T}$. E o mesmo é feito também para os 2000 cenários dos modelos *VARX* e Beta *t-GAS*.

Será proposto o uso de estatísticas descritivas como os quantis $Q^{(\alpha\%)} \forall \alpha \in \{0.05, 0.10, 0.5, 0.9, 0.95\}$ com o intuito de julgar se a metodologia proposta nessa dissertação, via o uso de cópula com a matriz de correlação sendo atualizada pelo mecanismo GAS e posteriormente calculando o quantil da densidade preditiva k passos à frente com essa informação, possui resultados consistentes com os resultados do modelo *VARX*.

Inicialmente pretendendo julgar o ajuste dessas simulações às séries de dados reais, é proposta uma análise gráfica entre as séries dos quantis mensais das simulações dos modelos contra a série dos quantis mensais da série histórica. A proposta dessa análise é que, se grande parte dos quantis da série histórica estiverem contidos dentro dos quantis dos cenários gerados, há indícios de que os cenários gerados pelos modelos foram capazes de capturar corretamente a estrutura de dependência da série real.

As Figuras 6.15, 6.16 e 6.17 apresentam as séries dos quantis $Q^{(\alpha\%)} \forall \alpha \in \{0.05, 0.10, 0.5, 0.9, 0.95\}$, calculados a partir dos cenários gerados pelos mode-

los, Beta t -GAS e VARX, contra aqueles da série histórica.

A partir dos gráficos contidos nas Figuras 6.15, 6.16 e 6.17, é possível afirmar que ambos os modelos propostos capturaram corretamente o perfil sazonal das séries temporais de FC. Assumindo o modelo VARX como um referencial para geração de cenários de FC conjuntos, é possível concluir que a metodologia proposta nessa dissertação via aplicação do modelo Beta t -GAS a um problema real de geração de cenários integrados de FC, fornece cenários coerentes. Ou seja, esse resultado empírico fornece subsídios para sustentar a metodologia proposta, pois os resultados da mesma vão de encontro com o resultado de um modelo que já está consolidado na literatura para o mesmo contexto.

A partir de uma análise detalhada nos cenários reproduzidos para cada usina, é possível obter as seguintes conclusões. Referente aos cenários produzidos para a usina de RF, o modelo Beta t -GAS subavalia o risco em dois períodos de queda (março e abril) e em três períodos de alta (agosto, setembro e outubro). Entretanto, a dinâmica da mediana dos cenários simulados por este modelo fica muito similar a dinâmica da mediana das séries reais. Já os cenários produzidos pelo modelo VARX possuem uma amplitude interquartilica menor, o que acaba prejudicando os ajustes e causando violações em março, abril, junho, agosto e setembro. Todavia os cenários simulados pelo modelo VARX são capazes de reproduzir muito bem a dinâmica do mês de novembro.

A respeito dos cenários de IC, novamente o modelo Beta t -GAS reproduz cenários que subavaliam o risco em abril e setembro, ao passo que os cenários obtidos pelo modelo VARX para essa usina, apesar de violações não muito significativas em janeiro e fevereiro, capturam bem a dependência. Esse fato vai de encontro com o resultado exposto pela Tabela 6.14 sobre a previsão univariada para a série dessa usina, onde o modelo VARX superou a capacidade preditiva do modelo Beta GAS SARIMA(p,q).

Para finalizar, acerca dos cenários produzidos para a usina de EN, o modelo Beta t -GAS subavalia os períodos de alta no FC. Isto é, de janeiro a outubro, não há violações nos quantis superiores, $\alpha = \{0.9, 0.95\}$, apenas nos meses de abril e março. Ao passo que os cenários produzidos a partir do modelo VARX subavaliam a alta no FC em todo o primeiro semestre, sendo junho o período com uma maior distorção entre os quantis dos cenários produzidos e o da série histórica. Para uma maior quantificação a respeito das violações mencionadas nesse parágrafo, referenciamos as Tabelas 6.16 e 6.17, onde estão expostos os resultados dos quantis calculados.

Ademais, apesar de algumas violações, é possível concluir que os ajustes

Figura 6.15: Análise descritiva dos cenários de FC gerados para a usina de RF no ano de 2011 entre os modelos Beta t -GAS e VARX contra a série histórica.

Figura 6.16: Análise descritiva dos cenários de FC gerados para a usina de IC no ano de 2011 entre os modelos Beta t -GAS e VARX contra a série histórica.

Figura 6.17: Análise descritiva dos cenários de FC gerados para a usina de EN no ano de 2011 entre os modelos Beta t -GAS e VARX contra a série histórica.

de ambos os modelos estão adequados. Isso porque em maior parte do tempo, os quantis das séries históricas de RF, IC e EN estão contidos dentro dos quantis dos cenários simulados a partir dos dois modelos. Vale relembrar que os cenários são produzidos para o ano de 2011 e que estamos utilizando a série histórica apenas como um critério para avaliar a capacidade dos modelos em gerar cenários coerentes com a estrutura de dependência sazonal dos dados de FC. Logo um evento atípico em um mês de alta, ou baixa, na série histórica, pode ocasionar uma violação quando comparada aos quantis dos cenários produzidos pelos modelos.

Os resultados expostos nas Tabelas 6.16 e 6.17 fornecem subsídios adicionais, em favor da metodologia proposta. Tal argumento se baseia no fato de que, os resultados calculados a partir dos cenários de ambos os modelos, possuem resultados muito similares. O que, apesar das particularidades da série e sua periodicidade mensal, sustenta a proposição da nossa metodologia para a modelagem conjunta de séries temporais. Novamente frisando, o objetivo desse estudo não é o de julgar se um modelo possui uma capacidade de geração de cenários superior ao outro, e sim o de sustentar a proposta teórica feita nessa dissertação via a aplicação prática em um problema real.

Tabela 6.16: Estatísticas descritivas das densidades dos FC avaliadas para o primeiro semestre de 2011 para as usinas RF, IC e EN.

		RF			IC			EN					
	Quantil	Série	Beta	t-GAS	VARX	Série	Beta	t-GAS	VARX	Série	Beta	t-GAS	VARX
JAN	5%	19.894	20.269	18.644	18.644	18.150	18.293	8.483	8.483	20.206	16.572	13.616	13.616
	10%	22.079	22.336	20.346	20.346	19.148	21.344	11.207	11.207	22.680	19.693	15.541	15.541
	50%	28.681	29.674	28.101	28.101	28.599	33.065	23.864	23.864	32.644	29.876	24.330	24.330
	90%	35.896	36.831	36.175	36.175	43.469	46.419	39.842	39.842	41.556	40.918	35.469	35.469
	95%	38.323	38.700	38.109	38.109	45.374	49.808	43.967	43.967	46.485	44.099	38.596	38.596
FEV	5%	16.164	14.906	17.088	17.088	8.314	6.839	7.258	7.258	10.920	8.220	13.025	13.025
	10%	18.748	17.635	19.374	19.374	10.936	9.444	9.648	9.648	14.104	10.626	15.489	15.489
	50%	27.358	26.617	28.315	28.315	19.616	20.730	23.424	23.424	27.591	21.674	25.478	25.478
	90%	34.085	35.010	36.275	36.275	35.661	35.404	35.940	35.940	37.140	35.150	35.295	35.295
	95%	35.331	37.119	38.229	38.229	39.626	40.192	38.381	38.381	39.127	38.452	37.765	37.765
MAR	5%	10.071	10.927	7.842	7.842	5.034	4.227	2.639	2.639	6.798	4.249	3.962	3.962
	10%	11.938	13.497	9.702	9.702	5.887	6.387	4.435	4.435	8.705	6.563	5.429	5.429
	50%	22.266	23.064	17.321	17.321	15.270	18.406	17.296	17.296	18.930	16.552	15.205	15.205
	90%	28.774	31.796	26.100	26.100	25.962	34.696	30.899	30.899	27.757	30.839	27.076	27.076
	95%	31.290	33.948	28.707	28.707	27.401	39.545	33.230	33.230	33.801	34.646	29.805	29.805
ABR	5%	7.224	9.605	8.540	8.540	3.030	5.096	1.851	1.851	5.728	3.310	3.569	3.569
	10%	12.021	11.859	10.571	10.571	4.354	7.272	2.922	2.922	7.165	5.321	5.317	5.317
	50%	20.129	21.569	19.265	19.265	15.803	20.444	14.540	14.540	16.130	16.352	15.643	15.643
	90%	31.020	31.237	28.537	28.537	31.017	37.476	29.179	29.179	29.887	29.509	28.205	28.205
	95%	32.532	33.710	31.155	31.155	32.793	41.991	31.600	31.600	31.863	33.246	30.136	30.136
MAI	5%	16.166	10.026	13.824	13.824	8.688	4.816	6.799	6.799	8.162	3.091	6.656	6.656
	10%	16.505	12.768	15.700	15.700	8.900	8.247	9.505	9.505	8.463	4.957	9.303	9.303
	50%	24.050	22.963	23.662	23.662	26.105	24.084	24.500	24.500	21.298	16.254	22.536	22.536
	90%	31.477	32.052	30.657	30.657	36.971	41.376	39.502	39.502	36.312	29.373	34.635	34.635
	95%	31.927	34.331	32.365	32.365	41.042	46.121	42.503	42.503	37.805	33.627	37.086	37.086
JUN	5%	20.480	14.039	16.862	16.862	20.860	11.026	10.447	10.447	11.910	3.609	3.555	3.555
	10%	22.272	17.406	17.871	17.871	21.784	15.348	14.781	14.781	13.148	6.088	4.624	4.624
	50%	29.607	27.838	23.716	23.716	34.270	33.256	34.673	34.673	22.640	17.426	17.254	17.254
	90%	36.082	35.896	31.343	31.343	50.312	50.866	54.904	54.904	38.427	31.444	31.130	31.130
	95%	36.631	38.291	33.633	33.633	53.287	54.820	58.246	58.246	42.172	35.311	32.373	32.373

Tabela 6.17: Estatísticas descritivas das densidades dos FC avaliadas para o segundo semestre de 2011 para as usinas RF, IC e EN.

		RF			IC			EN		
	Quantil	Série	Beta <i>t-GAS</i>	<i>VARX</i>	Série	Beta <i>t-GAS</i>	<i>VARX</i>	Série	Beta <i>t-GAS</i>	<i>VARX</i>
JUL	5%	25.503	21.752	24.312	28.546	25.621	22.055	19.373	7.702	15.388
	10%	27.105	24.020	25.892	34.140	31.181	25.017	20.347	10.789	19.018
	50%	35.829	33.103	33.133	49.105	51.042	40.442	29.289	23.456	34.509
	90%	40.402	40.214	41.569	59.311	64.247	57.877	43.294	37.699	47.260
	95%	41.247	41.843	43.405	61.272	67.268	61.577	43.949	41.424	50.179
AGO	5%	32.528	28.987	28.495	54.965	51.858	54.988	33.493	18.537	28.724
	10%	32.799	31.855	29.809	55.660	56.982	57.678	36.803	23.455	31.301
	50%	39.206	39.363	36.550	66.343	69.071	69.678	44.323	37.337	42.595
	90%	47.067	44.729	44.720	71.435	75.672	78.524	53.519	49.735	54.302
	95%	49.083	45.949	46.829	73.268	77.044	80.034	54.282	52.678	57.315
SET	5%	34.639	31.792	31.461	66.163	60.502	66.369	49.443	35.759	46.400
	10%	35.539	33.859	32.839	67.072	63.793	68.945	50.332	39.920	48.149
	50%	40.226	40.463	39.131	74.177	71.663	78.044	54.902	51.073	56.094
	90%	46.517	45.197	46.019	78.368	77.064	84.744	60.344	58.017	63.843
	95%	49.955	46.187	47.960	82.638	78.258	86.006	63.816	59.639	65.415
OUT	5%	32.323	27.816	32.454	64.352	62.408	60.477	50.845	47.835	50.639
	10%	33.296	30.419	34.148	64.618	64.776	62.055	52.321	50.790	53.075
	50%	40.957	37.752	40.755	73.278	71.974	71.516	59.139	57.737	61.445
	90%	44.261	43.328	47.035	77.474	77.423	80.433	64.449	62.424	68.875
	95%	45.725	44.695	48.592	78.714	78.453	82.997	66.692	63.280	70.441
NOV	5%	33.929	24.466	32.584	55.756	45.034	58.680	51.426	43.978	49.099
	10%	34.902	27.330	33.263	60.344	48.901	60.750	52.194	47.101	49.819
	50%	37.695	35.347	36.670	64.665	60.527	68.788	55.617	55.062	54.116
	90%	40.725	41.878	40.525	70.845	69.181	75.112	59.838	60.427	59.028
	95%	41.036	43.116	41.399	72.076	71.030	76.486	60.520	61.526	60.286
DEZ	5%	27.787	20.519	28.407	31.490	24.352	28.432	34.241	30.223	35.683
	10%	28.080	23.298	29.667	36.572	28.906	31.643	36.331	33.955	38.130
	50%	34.158	31.887	35.235	50.156	45.002	48.805	46.957	44.929	48.569
	90%	38.722	39.065	40.366	58.003	57.161	65.768	52.889	53.647	57.727
	95%	39.230	40.623	41.524	59.967	60.638	69.126	55.495	55.723	59.884

7

Considerações finais

Essa dissertação teve como principal objetivo a apresentação de um modelo multivariado com mecanismo de atualização GAS para a modelagem da estrutura de dependência via cópulas, com uma aplicação em um contexto de séries temporais não gaussianas. Vale lembrar que o arcabouço apresentado aqui é geral, não se restringindo apenas a marginais com densidade preditiva Beta. Essa densidade foi escolhida devido a particularidade ao se trabalhar com modelos GAS, pois é necessário escolher uma densidade preditiva capaz de acomodar em seu suporte os valores assumidos pelas séries modeladas. Conforme já mencionado na sessão 3.1, os FC das usinas estudadas assumem valores entre $[0, 100]$, logo escolhe-se uma Beta $[0, k]$ para a modelagem de cada marginal, indo de acordo com o que foi feito em Matos (2013).

É importante salientar que, a introdução da informação obtida a partir da modelagem da dependência estocástica via cópulas em um contexto de dados sazonais para a classe de modelos GAS é inédita. Para a modelagem da estrutura de dependência entre as usinas foi utilizado o modelo GAS t -Student(p,q) de Creal *et al.* (2011). Conforme apresentado no Capítulo 2, os modelos GAS não possuem termo de erro e por isso, ao contrário dos trabalhos de Patton (2006), Patton (2009) e Patton (2012), nos quais se introduz a informação da distribuição conjunta das variáveis no termo de erro para a simulação de uma carteira de ativos, nessa nova classe de modelos propomos o uso do quantil da densidade preditiva k passos à frente, gerado a partir da matriz de correlação atualizada pelo modelo GAS t -Student(p,q). Uma das principais vantagens em torno do uso dos modelos GAS com essa proposição está na simulação de cenários conjuntos, pois é garantido que os valores gerados sempre estarão contidos dentro do suporte paramétrico da densidade preditiva especificada para o modelo, supondo um arcabouço de cópulas totalmente paramétrico. No estudo de caso do Capítulo 6, a densidade especificada foi Beta, logo os valores de FC eólico simulados serão sempre positivos e menores do que 100, respeitando os limites físicos da variável modelada.

Em geral, ao simular os modelos com uma distribuição conjunta postu-

lada para as inovações do modelo, a utilização de cópulas fornece uma informação sobre o movimento conjunto entre as variáveis. Entretanto conforme já mencionado anteriormente, os dados de FC são mensais. Este valor é obtido a partir de uma média da produção horária e posteriormente a média dessa produção diária, resultando em um único valor mensal de FC. Essa média suaviza os efeitos de cauda, expurgando essa característica na modelagem conjunta. Consequentemente a distribuição dos termos de $F_{\nu}^{-1}(u_{it})$ para $I = \{1, 2, 3\}$ e $T = \{13, \dots, 360\}$ tendem a assumir caudas mais leves, o que é observado empiricamente sendo $\hat{\nu} = 340$ o valor de ν que maximiza o valor do log da verossimilhança. Esse valor de ν aponta para a especificação de uma cópula gaussiana, reduzindo probabilidade nas caudas. Vale ressaltar que o processo de estimação utilizado foi o da verossimilhança perfilada e que este não garante que o conjunto de parâmetros estimados seja o ótimo.

Entretanto, apesar das limitações da natureza da série modelada, ao comparar os CVaR^{95%} da Tabela 6.15, referentes as distribuições de renda de um contrato de um ano no ACL, vinculado a um portfólio das três usinas (RF, IC e EN), a modelagem contemplando a estrutura de dependência fica evidente. O investidor irá experimentar um aumento no seu retorno em todos os casos utilizando o modelo que contempla a modelagem conjunta.

Da mesma forma, é importante destacar a contribuição teórica dessa dissertação. A aplicação de cópula com a matriz de correlação sendo atualizada pelo mecanismo GAS e posteriormente calculando o quantil $\alpha\%$ da densidade preditiva k passos à frente com essa informação, na qual o parâmetro da densidade preditiva também é variante no tempo. Os resultados expostos na sessão 6.7 dão insumos para concluir que, ao aplicar a metodologia proposta a um problema real, a mesma obteve resultados similares ao modelo *VARX* que é utilizado como *benchmark* para a modelagem conjunta de FC eólico, mantendo a estrutura sazonal dos dados. Ainda, é importante salientar que a contribuição teórica dessa dissertação se estende a outros arcabouços, com diferentes densidades preditivas especificadas para as marginais. Por exemplo, seria possível a modelagem multivariada das séries de vazão com periodicidade horária, onde se especificaria uma densidade Gama para as marginais, posto que o suporte dessa densidade acomoda os valores dessas séries.

O objetivo da referida sessão não era o de julgar qual era o melhor modelo, logo fica a critério do encarregado de analisar os dados optar por um dos dois para geração de cenários condicionais, segundo as características e especificidades de cada um. Apenas para dar um embasamento prático, vale mencionar que uma das vantagens do modelo *VARX* que já é discutida

na literatura fica em torno da possibilidade de gerar cenários condicionados aos preços *spot*. Tal característica também é facilmente implementável no arcabouço de modelos GAS, basta incluir como variável exógena no *score* da densidade preditiva as variáveis *dummi* referente as ENAS, o que é feito em Matos (2013). Quanto a vantagem acerca do uso do modelo Beta *t-GAS*, essa já foi mencionada no início desse Capítulo, que é a garantia de que os cenários gerados para o FC serão sempre positivos. Ao passo que no caso dos modelos *VARX*, é necessário trabalhar com variáveis transformadas para não gerar cenários negativos. Para isso, após a simulação dos cenários, aplica-se a transformação inversa para que as séries geradas retornem à escala original.

Referente aos ajustes dos modelos univariados Beta GAS SARIMA(p,q), apesar de algumas violações nos gráficos da função de autocorrelação e autocorrelação elevada ao quadrado, além da rejeição do teste de autocorrelação de Ljung Box em duas das três usinas, o modelo Beta GAS SARIMA(p,q) foi capaz de capturar corretamente a estrutura de dependência individual de cada densidade marginal. Além disso, tal fato é comprovado pelos diagnósticos das previsões 12 passos à frente com o modelo Beta GAS SARIMA(p,q), os quais apresentaram um bom ajuste segundo os critérios estabelecidos.

Para futuras pesquisas, seria interessante dirimir as limitações encontradas nessa dissertação. O primeiro ponto ficaria em torno do cálculo do erro padrão para os parâmetros do modelo de cópula usando a informação das marginais. Em segundo, seria interessante adotar um método mais adequado para a estimação do grau de liberdade ν para o modelo GAS *t-Student*(p,q), de maneira a evitar o uso de verossimilhança perfilada. Finalmente, o terceiro ponto fica em torno da proposição e implementação de um teste estatístico para a comparação dos conjuntos de simulação gerados pelos modelos *VARX* e Beta *t-GAS*.

Algumas aplicações interessantes com esse modelo seriam, por exemplo, a implementação da estrutura de cópula condicional proposta em Neves *et al.* (2014). A partir dessa ferramenta gerencial, seria possível gerar cenários de fator de capacidade eólico condicionando as demais usinas a um determinado valor, i.e., gerar cenários onde dado que uma usina 1 produziu x e a usina 2 produz y , são gerados cenários condicionais a esses valores para a usina 3. Além disso, outra aplicação que seria de grande interesse seria a modelagem em alta dimensão com o modelo GAS *t-Student*(p,q) usando a especificação via fatores, conforme apresentado em Oh e Patton (2012), para a modelagem de toda a matriz do sistema elétrico brasileiro, definindo os fatores como os subsistemas (Sudeste/Centro-Oeste, Sul, Norte e Nordeste).

Referências Bibliográficas

- ABADIR, K. M. e MAGNUS, J. R. **Matrix Algebra**. Cambridge University Press, 2005.
- AMARANTE, O. A. C., BOWER, M., ZACK, J. e DE SÁ, A. L. **Atlas do potencial eólico brasileiro**, Ministerio de Minas e Energia, 2001.
- ARTZNER, P., FREDDY, D., EBER, J. M. e HEATH, D. Coherent Measures of Risk, **Mathematical Finance**, v. 9, n. 3, p. 203–228, 1999.
- AVDULAJ, K. e BARUNIK, J. **Are benefits from oil-stocks diversification gone? a new evidence from a dynamic copulas and high frequency data**. Working paper, 2015
- BEZERRA, B., PORRUA, F., BARROSO, L. A., LINO, P., RALSTON, F. e PEREIRA, M. V. Wind power insertion through energy auctions in Brazil. In **Proceedings of the IEEE PES General Meeting**. 2010
- CASELLA, G. e BERGER, R. L. **Statistical Inference**. Cengage Learning, 2001.
- CHERUBINI, U., LUCIANO, E. e VECCHIATO, W. **Copula methods in finance**. John Wiley & Sons, 2004.
- COX, D. R. Statistical analysis of time series: some recent developments. **Scandinavian Journal of Statistics**, v. 8, p. 93-115, 1981.
- CREAL, D., KOOPMAN, S. J. e LUCAS, A. A General Framework for Observation Driven Time-Varying Parameter Models. **Tinbergen Institute Discussion**, n. 108, 2008.
- CREAL, D., KOOPMAN, S. J. e LUCAS, A. A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations, **Journal of Business & Economic Statistics**, v. 29, n.4, p. 552–563, 2011.

CREAL, D., KOOPMAN, S. J. e LUCAS, A. Generalized Autoregressive Score Models with Applications. **Journal of Applied Econometrics**, v. 28, n.5, p. 777–795, 2013.

DAVIS, R. A., DUNSMUIR, W. T. M. e STREETT S. B. Observation driven models for Poisson counts. **Biometrika**, v. 90, n. 4, p. 777-790, 2003.

DIEBOLD, F. X., HAHN, J. e TAY, A. S. Multivariate density forecast evaluation and calibration in financial risk management: high-frequency returns on foreign exchange. **Review of Economics and Statistics**. v. 81, n. 4, p. 661–673, 1999.

DUNN, P. K. e SMYTH, G. K. Randomized quantile residuals. **Journal of Computational and Graphical Statistics**. v. 5, n. 3, p. 236–244, 1996.

EMBRECHTS, P., FREY, R. e MCNEIL, A. **Quantitative risk management**. Princeton University Press, 2005.

ENGLE, R. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. **Journal of Business & Economic Statistics**, v. 20, n. 3, p. 339-350, 2002A.

ENGLE, R. New frontiers for arch models. **Journal of Applied Econometrics**, v. 17, n. 5, p. 425–446, 2002B.

ENGLE, R. F. e BOLLERSLEV, T. Modelling the persistence of conditional variances. **Econometric reviews**, v. 5, n. 1, p. 1–50, 1986.

ENGLE, R. F. e GALLO, G. M. A multiple indicators model for volatility using intra-daily data. **Journal of Econometrics**, v. 131, n. 1, p. 3–27, 2006.

ENGLE, R. F. e RUSSEL, J. R. Autoregressive conditional duration: a new model for irregularly spaced transaction data. **Econometrica**, v. 66, n. 5, p. 1127-1162, 1998.

ENGLE, R. F. e SHEPPARD, K. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. **Technical report, National Bureau of Economic Research**, 2001.

FREIRE, L., STREET, A., LIMA, D. A. e BARROSO, L. A. **A Hybrid MILP and Benders Decomposition Approach to Find the Nucleolus Quota Allocation for a Renewable Energy Portfolio**. IEEE Trans. Power Syst. Dec 2014.

GARCIA, J. D. e STREET, A. **A methodology for wind power series extension: Data for investments, operation and expansion planning of the power sector**. Working paper, 2013.

HARVEY, A. C. **Dynamic models for volatility and heavy tails: with applications to financial and economic time series**. Cambridge University Press, 2013.

JOE, H. **Multivariate models and multivariate dependence concepts**. CRC Press, 1997.

JOE, H. Asymptotic efficiency of the two-stage estimation method for copula-based models. **Journal of Multivariate Analysis**. v. 94, n. 2, p. 401–419, 2005.

JORION, P. **Value at Risk: The New Benchmark for Managing Financial Risk**. New York: McGraw-Hill, 2006.

KALLIOVIRTA, L. **Diagnostic Tests Based on Quantile Residuals for Nonlinear Time Series Models**. 2009. 165f. Tese de doutorado – Universidade de Helsinki. 2009.

KOOPMAN, S. J., LUCAS, A., SCHARTH, M., BLASQUES, F., JANUS, P. e SCHWAAB, B. Generalized autoregressive score models for time-varying parameters in economics and finance **Research Seminar for National Bank of Serbia**. Belgrade, 6 de Julho de 2012, 2012.

LEHMANN, E. L. e CASELLA, G. **Theory of Point Estimation**. Springer, 2003.

MATOS, G. G. D. **Modelos GAS Aplicados a Séries Temporais de Vazão e Vento**. 2013. 103f. Dissertação (Mestrado em Engenharia Elétrica) - PUC-Rio, Rio de Janeiro. 2013.

MENDES, B. V. M. **Introdução à Análise de Eventos Extremos**. Editora E-papers, 2004.

NELSEN, R. B. **Value at Risk: An introduction to copulas**. Springer, 1999.

NEVES, C., FERNANDES, C. e MELO, E. Forecasting surrender rates using elliptical copulas and financial variables. **North American Actuarial Journal**. v. 18, n. 2, p. 343-362, 2014.

OH, D. H. e PATTON, A. J., **Modelling dependence in high dimensions with factor copulas**. Working paper. 2012.

ORD, J. K., KOEHLER, A. B. e SNYDER, R. D. Estimation and prediction for a class of dynamic nonlinear statistical models. **Journal of the American Statistical Association**, v. 92, n. 440, p. 1621 - 1629, 1997.

PATTON, A. J. **Applications of copula theory in financial econometrics**. 2002. 217f. Tese de doutorado – Universidade da Califórnia. 2002.

PATTON, A. J. Copula methods for forecasting multivariate time series. **Handbook of Economic Forecasting**, v. 2, p. 1 - 77, 2012.

PATTON, A. J. Copula-based models for financial time series. **Handbook of Financial Time Series**. p. 767-785, 2009.

PATTON, A. J. Modelling asymmetric exchange rate dependence. **International economic review**. v. 47, n. 2, p. 527-556, 2006.

PEREIRA, M. e PINTO, L. M. Multistage stochastic optimization applied to energy planning. **Mathematical Programming**. v. 52, n. 1-3, p. 359-375, 1991.

POLITIS, D. N. e ROMANO, J. P., The stationary bootstrap. **Journal of the American Statistical Association**. v. 89, n. 428, p. 1303-1313, 1994.

ROSS, S. M. **Simulation**. Academic Press, 2012.

SKLAR, M. **Fonctions de Répartition à n Dimensions et Leurs Marges**. Publications de Institut de Statistique de l'Université de Paris, n.8, p.229-231, 1959.

STREET, A., LIMA, D. A., VEIGA, A., FANZERES, B., FREIRE, L. e AMARAL, B. Fostering Wind Power Penetration into the Brazilian

Forward-Contract Market. **Power and Energy Society General Meeting, 2012 IEEE** . IEEE, 2012. p. 1-8.

XU, J. J. **Statistical modelling and inference for multivariate and longitudinal discrete response data**. 1996. Tese de doutorado – Universidade British Columbia. 1996.

A

Prova da Cópula ser invariante a transformações crescentes

A prova apresentada a seguir foi retirada das páginas 104 e 105 de Ross (2012). Para provar, considere $s(X_1)$ e $t(X_2)$ onde $s(\cdot)$ e $t(\cdot)$ são duas funções crescentes, então a cópula gerada pelo vetor $s(X_1), t(X_2)$ é igual à cópula gerada por X_1 e X_2 .

Proposição Se s e t são funções crescentes, então

$$C_{s(X_1), t(X_2)}(x_1, x_2) = C_{X_1, X_2}(x_1, x_2).$$

Prova Sendo $F_i, i = 1, 2$ as representações marginais, então a função de distribuição de F_s é

$$\begin{aligned} F_s &= P(X_1 \leq x_1) \\ &= P(X_1 \leq s^{-1}(x_1)) \text{ Porque } s \text{ é uma função crescente} \\ &= F(s^{-1}(x_1)) \end{aligned}$$

Analogamente a função de distribuição de $t(X_2)$, expressa por F_t é dada por

$$F_t = F(t^{-1}(x_2))$$

Por consequência,

$$\begin{aligned} F_s(s(X_1)) &= F(s^{-1}(s(X_1))) = F(X_1) \\ F_t(t(X_2)) &= F(t^{-1}(t(X_2))) = F(X_2) \end{aligned}$$

Provando que

$$\begin{aligned} C_{s(X_1), t(X_2)}(x_1, x_2) &= P(F_s(s(X_1)) \leq x_1, F_t(t(X_2)) \leq x_2) \\ &= P(F_s(X_1) \leq x_1, F_t(X_2) \leq x_2) \\ &= C_{X_1, X_2}(x_1, x_2). \end{aligned}$$

B

Casos especiais da densidade *t-Student* multivariada

Consideramos dois casos especiais da densidade (5-1). O primeiro se restringe ao caso em que $\nu^{-1} \rightarrow 0$, a densidade se reduz à densidade normal multivariada dada por

$$p(y_t|\Sigma_t) = \frac{1}{(2\pi)^{k/2}|\Sigma_t|^{1/2}} \exp\left(-\frac{1}{2}y_t'\Sigma^{-1}y_t\right). \quad (\text{B-1})$$

O segundo caso se restringe quando $k = 1$, a densidade (5-1) se reduz a densidade *t-Student* univariada expressa por,

$$p(y_t|\sigma_t; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi\sigma_t^2}} \left[1 + \frac{y_t^2}{(\nu-2)\sigma_t^2}\right]^{-(\nu+1)/2} \quad (\text{B-2})$$

onde σ_t^2 é um escalar variante no tempo.

C

Conditional Sampling Bivariado

Para realizar tal, se utiliza a distribuição condicional para a variável aleatória V no ponto $U = u$, tal como,

$$c_u(v) = \Pr(V \leq v | U = u). \quad (\text{C-1})$$

A aplicação de tal procedimento, considerando um caso bivariado, é feita seguindo o passo a passo do quadro ilustrado abaixo.

Basicamente sabemos que

$$\begin{aligned} c_u(v) &= \Pr(F_2 \leq v | F_1 = u) = \lim_{\Delta u \rightarrow 0} \frac{C(u + \Delta u, v) - C(u, v)}{\Delta u} = \frac{\partial C}{\partial u} \\ &= C_u(v) \end{aligned}$$

onde $C_u(v)$ é a derivada parcial da cópula. Como é sabido que $c_u(v)$ é uma função não decrescente e existe para quase todo o $v \in [0, 1]$, geram-se os pares (u_t^*, v_t^*) da seguinte maneira:

1. Gera duas uniformes independentes $(u, w) \in [0, 1]$, armazenando u como o primeiro dado do par, i.e., $u = u^*$;
2. Compute a (quasi)-inversa da função $c_u(v)$. Isso irá depender dos parâmetros da cópula e de u . Dessa forma, para obter o segundo valor do par, basta resolver $v = c_u^{-1}(w)$.