

FELIPE FRANCISCO ROMANO

ANÁLISE DA MATRIZ ENERGÉTICA BRASILEIRA
SOB A PERSPECTIVA DE TEORIA DE PORTFÓLIOS

SÃO PAULO
2014

FELIPE FRANCISCO ROMANO

ANÁLISE DA MATRIZ ENERGÉTICA BRASILEIRA
SOB A PERSPECTIVA DE TEORIA DE PORTFÓLIOS

Trabalho de Formatura apresentado à
Escola Politécnica da Universidade de São
Paulo para obtenção do diploma de
Engenharia de Produção

Orientadora: Prof.^a Dr.^a
Celma de Oliveira Ribeiro

SÃO PAULO
2014

Catálogo-na-publicação

Romano, Felipe Francisco

Análise da Matriz Energética Brasileira sob a Perspectiva de Teoria de Portfólios. / F. F. Romano. -- São Paulo, 2014.

93 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1.Otimização de portfólios. I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Produção II.t.

Dedicado à minha mãe.

AGRADECIMENTOS

À minha mãe, cujo legado de carinho e sabedoria ainda me acompanha. Nunca esquecerei as palavras, gestos e ensinamentos que moldaram a minha personalidade e guiaram os meus passos até onde cheguei.

Ao meu pai, pela educação durante toda minha vida, sempre se preocupando em me prover as melhores condições para uma boa formação.

Aos meus irmãos, por sempre estarem ao meu lado e serem uma fonte de inspiração.

À minha namorada, por compreensão, carinho e diversão infintos.

Ao restante da minha família, pelo suporte ao longo da vida.

Aos meus amigos de longa data, com muitos dos quais venho passando inesquecíveis momentos.

Aos mestres Cláudio Possani e Cláudio Arconcher, por despertarem em mim a faísca do prazer eterno pelo conhecimento.

À Profa. Dra. Celma de Oliveira Ribeiro, pela excelente orientação durante o Trabalho de Formatura, através de preciosos conselhos e críticas.

Aos demais professores da Escola Politécnica que vêm compartilhando seus conhecimentos e contribuindo para a minha formação acadêmica e profissional desde 2009.

Aos meus colegas de trabalho no The Boston Consulting Group, pela engrandecedora experiência de trabalhar numa empresa motivadora, respeitosa e desafiante.

“Sem planos, só há promessas e esperanças.

Antes de trabalho duro, porém, planos não passam de boas intenções.”

(Peter Drucker)

RESUMO

Este trabalho de formatura consiste em uma aplicação da Teoria de Portfólios à matriz energética brasileira, com foco no modelo de Média-Variância de Markowitz.

É apresentado, inicialmente, um estudo comparativo de diferentes modelos de otimização de portfólio propostos pela literatura, sob duas métricas: desempenho e rugosidade, em que rugosidade é um indicador de quanto um portfólio varia ao longo do tempo. Em seguida, algoritmos evolutivos são utilizados para otimizar os parâmetros dos modelos estudados e, consequentemente, obter uma melhora substancial de seus desempenhos.

Concluimos que, apesar das diversas propostas de melhoria presentes na literatura acadêmica, como otimização bayesiana e o uso do CVaR como medida de risco, o modelo clássico de Markowitz, mesmo não sendo o mais robusto, mantém a sua posição como modelo com melhor balanço entre desempenho e robustez dentre os analisados. Verificamos também que utilizar algoritmos evolutivos para parametrizar modelos de otimização de portfólio é extremamente eficaz.

Ao aplicar o modelo de Markowitz à problemática da geração de energia no Brasil, nós inicialmente localizamos a atual matriz energética em relação ao plano de risco-retorno e demonstramos através da fronteira eficiente que ela é sub-ótima. Em seguida, propomos uma matriz energética otimizada que concilia uma redução considerável do risco e um aumento significativo do retorno (neste contexto, denominamos “retorno” a energia gerada por unidade monetária paga pelo governo).

Por fim, analisamos a sensibilidade do portfólio ótimo em relação à proporção de energia hidrelétrica exigida, concluindo que o Brasil teria grandes benefícios em diminuir a dependência de recursos hídricos. Ao comparar os nossos resultados com o de outros autores, demonstramos que, sob a perspectiva do nosso estudo, a nossa proposta é mais benéfica para o país.

Palavras-chave: otimização de portfólios, geração de energia, matriz energética, modelo da média-variância, otimização robusta, VaR, CVaR, algoritmos evolutivos, redes neurais

ABSTRACT

This bachelor thesis presents an application of portfolio theory, in particular of Markowitz's mean-variance model, to energy generation in Brazil.

We start with a comparative study of several portfolio optimization models found in the academic literature, according to two criteria: their performances and their roughness, in which roughness is defined as a measure of portfolio composition change over time. We then proceed with the use of evolutionary algorithms to optimize the parameters of these models and, hence, substantially improve their performance.

We conclude that, in spite of several improvement propositions, such as the use of bayesian optimization and the CVaR as a risk measure, the classic Markowitz model, although not the most robust among the analyzed models, keeps its position as the best-performing model. We also verify that evolutionary algorithms are extremely efficient to optimize the parameters of portfolio models.

By applying the mean-variance model to energy generation in Brazil, we demonstrate that Brazil's current energy-generating portfolio is suboptimal and propose an optimized energy portfolio, which conciliates a sizable reduction in risks and a valuable increase in energy production return (in this context, "return" corresponds to the amount of energy generated by each dollar paid by the government).

We then conduct a sensibility analysis on the optimal portfolio, regarding the minimum required ratio of hydroelectric power. We conclude that Brazil would greatly benefit from decreasing its dependence on water resources. Lastly, we compare our results to those of similar studies. We demonstrate that, according to our models, the optimal portfolio we propose is superior to the other approaches.

Keywords: *portfolio optimization, energy generation, energy generation portfolio, mean-variance model, VaR, CVaR, evolutionary algorithms, neural networks*

LISTA DE FIGURAS

Figura 1 – Organização matricial do BCG	19
Figura 2 – Cálculo do VaR (Value-at-Risk).....	30
Figura 3 – Ineficiência do VaR	31
Figura 4 – Ilustração da fronteira eficiente.....	33
Figura 5 – Exemplo de evolução da composição do portfólio ótimo.....	37
Figura 6 – A curva formada pelos dj'	39
Figura 7 – Esquema de um neurônio humano.	41
Figura 8 – Gráfico da função sigmoide ou “logística”	41
Figura 9 – Representação de uma rede neural.....	42
Figura 10 – Arquitetura de uma rede neural.....	43
Figura 11 – Representação MATLAB da rede neural.....	48
Figura 12 – Exemplo da capacidade de predição das redes neurais.....	48
Figura 13 – Interface de utilização desenvolvida	51
Figura 14 – Variação da composição dos portfólios segundo o modelo.....	54
Figura 15 – Desempenho ao longo do tempo segundo os modelos	57
Figura 16 – Desempenho de Markowitz antes da otimização	64
Figura 17 – Desempenho de Markowitz depois da otimização.....	65
Figura 18 – Matriz energética brasileira (2013)	67
Figura 19 – Fronteira eficiente no contexto de geração de energia.....	70
Figura 20 – Retornos nivelados de energia (LROE) de acordo com a BNEF.....	78
Figura 21 – Riscos das fontes de energia de acordo com a BNEF.....	79
Figura 22 – Fronteira eficiente da matriz energética brasileira.....	80
Figura 23 – Fronteira eficiente com restrições de factibilidade	83
Figura 24 – Fronteira eficiente irrestrita e o portfólio MKW2014.....	83
Figura 25 – Sensibilidade da fronteira eficiente à energia hidrelétrica	85
Figura 26 – Comparação com projeções de outros estudos	87

LISTA DE TABELAS

Tabela 1 – Os 12 modelos estudados	50
Tabela 2 – Estudo de rugosidade dos modelos (situação 1)	52
Tabela 3 – Tabela ANOVA para o estudo de rugosidade dos modelos (situação 1).....	53
Tabela 4 – Estudo de rugosidade dos modelos (situação 2)	55
Tabela 5 – Estudo do desempenho dos modelos (situação 1).....	56
Tabela 6 – Tabela ANOVA para o estudo do desempenho dos modelos (situação 1)	56
Tabela 7 – Estudo do desempenho dos modelos (situação 2).....	58
Tabela 8 – Razão desempenho/rugosidade (situação 1)	59
Tabela 9 – Razão desempenho/rugosidade (situação 2)	60
Tabela 10 – Tabela ANOVA para desempenho/rugosidade (situação 1).....	60
Tabela 11 – Parâmetros do algoritmo evolutivo (MATLAB).	63
Tabela 12 – Resultados da otimização para todos os modelos	64
Tabela 13 – Margens das empresas do setor elétrico mais líquidas	76
Tabela 14 – Composição dos portfólios A, B e C.....	81
Tabela 15 – Composição do portfólio MKW2014	84

LISTA DE ABREVIATURAS E SIGLAS

ENERGIA

ANEEL	Agência Nacional de Energia Elétrica
BNEF	<i>Bloomberg New Energy Finance</i>
LCOE	<i>Levelized Cost of Energy</i> Custo Nivelado de Energia
ONS	Operador Nacional do Sistema Elétrico
PDE	Plano Decenal de Expansão da Energia
SIN	Sistema Interligado Nacional

ENGENHARIA E FINANÇAS

CVaR	<i>Conditional Value-at-Risk</i>
NAR	<i>Non-linear Autoregressive Neural Network</i> Rede Neural Não-Linear Auto-Regressiva
NYSE	<i>New York Stock Exchange</i>
SR	<i>Sharpe ratio</i>
TSR	<i>Total Shareholder Return</i>

SUMÁRIO

1. INTRODUÇÃO	19
1.1. Apresentação da Empresa	19
1.2. Programa de Estágio	20
1.3. Formulação do Problema.....	21
1.4. Objetivo do Trabalho	22
1.5. Justificativa do Trabalho	23
1.6. Estrutura do Trabalho	25
2. FUNDAMENTOS TEÓRICOS.....	27
2.1. Pesquisa Operacional	27
2.2. Preço, Retorno e Risco de Ativos.....	28
2.3. Modelo de Markowitz	32
2.4. Otimização Bayesiana	35
2.5. Rugosidade.....	36
2.6. Desempenho	37
2.7. Algoritmos Evolutivos.....	39
2.8. Redes Neurais	40
3. ESTUDO COMPARATIVO DE MODELOS DE OTIMIZAÇÃO DE PORTFÓLIO SELECIONADOS	45
3.1. Metodologia.....	45
3.2. Fatores considerados nos modelos	45
3.3. Modelos estudados	49
3.4. Detalhes Adicionais	51
3.5. Estudo de Rugosidade	52
3.6. Estudo de Desempenho.....	56

3.7.	Estudo da razão Desempenho/Rugosidade-----	58
3.8.	Aprimoramento do modelo de Markowitz através de algoritmos evolutivos ---	62
4.	OTIMIZAÇÃO DA MATRIZ ENERGÉTICA BRASILEIRA	67
4.1.	Mercado de energia no Brasil -----	67
4.2.	Planejamento energético baseado em gestão de portfólios -----	68
4.3.	Modelo de custos utilizado-----	71
4.4.	Metodologia-----	74
4.5.	Análise preliminar dos dados -----	77
4.6.	Resultados -----	80
4.7.	Análise de sensibilidade à energia hidroelétrica -----	84
4.8.	Comparação com outros resultados -----	86
5.	CONCLUSÕES.....	89
	REFERÊNCIAS BIBLIOGRÁFICAS	91
	APÊNDICE.....	93
	Matriz das Covariâncias -----	93

1. INTRODUÇÃO

1.1. Apresentação da Empresa

Fundada em 1963 por Bruce Henderson, o Boston Consulting Group (BCG) está presente no Brasil desde 1997. Com faturamento de 4 bilhões de dólares em 2013, o BCG é a segunda maior empresa do mundo de consultoria estratégica e se organiza matricialmente através de áreas funcionais e áreas de expertise na indústria, conforme apresentado na Figura 1:



Figura 1 – Organização matricial do BCG

Conhecida mundialmente por seus serviços, ferramentas analíticas (matriz BCG de portfólio, curva de experiência, entre outras), publicações e ex-funcionários, o BCG tem como missão ajudar corporações com grande relevância em seus segmentos de atuação a adquirir e sustentar uma vantagem estratégica de longo prazo. Como exemplos de clientes mundiais do BCG, podemos citar Ford, Pfizer e IBM. No âmbito de responsabilidade social, o Boston Consulting Group apoia muitos projetos sociais, como o Instituto Ayrton Senna, o Instituto Natura e a plataforma de inovação social Yunus Social Business. Atualmente, os principais concorrentes do BCG no mercado de consultoria estratégica são McKinsey & Co. e Bain & Co.

Nos rankings *A Melhor Empresa Para se Trabalhar* de 2014, o BCG ficou em 1º na Consulting Magazine e em 3º na Fortune, atrás exclusivamente do Google e de outra empresa de softwares.

1.2. Programa de Estágio

O programa de estágio se organizou em torno de duas tarefas principais:

- A.** Realização de um estudo macroeconômico, comparando o atual contexto do Brasil com o de outros países emergentes, em especial os outros BRICS (Rússia, Índia, China e África do Sul). O foco do estudo foi em produtividade, definida como o produto interno bruto gerado por pessoa no país.

Como resultados principais, pode-se citar que o Brasil possui produtividade baixa, alinhada com a dos países em desenvolvimento, e taxa de crescimento também baixa, similar à dos países desenvolvidos. Esta combinação de características configura um grave problema macroeconômico para o país.

Para melhorar compreender o problema da produtividade, foi realizado um estudo sobre o setor elétrico do país, que motivou a definição do tema deste trabalho.

- B.** Criação de um modelo de análise de Retorno Total para o Investidor (TSR, do inglês *total shareholder return*). O TSR é uma métrica de desempenho amplamente utilizada no mercado de capitais, dada a sua fácil compreensão:

$$\text{TSR} = \frac{\text{Valorização da Ação} + \text{Dividendos}}{\text{Investimento Inicial}}$$

Conquanto seja simples, o TSR é extremamente poderoso, pois é possível decompô-lo em subcomponentes, de forma a analisar as alavancas de geração de valor para cada empresa. Pode-se, por exemplo, decompor o TSR em crescimento de receita, alteração na margem, valorização nos múltiplos¹, aumento do endividamento, diluição das ações da empresa, etc.

¹ Múltiplos são medidas de desempenho de uma empresa, expressas na forma de razões entre duas métricas mais simples. Por exemplo, o múltiplo P/L é muito utilizado, correspondendo à razão entre o Preço de uma ação e o Lucro gerado pela empresa. Neste caso, o múltiplo é uma métrica de quanto o mercado está disposto a pagar por cada real que a empresa gere de lucro.

Através deste modelo, diversas análises setoriais foram feitas, de maneira a comparar quais são as alavancas mais proeminentes de geração de valores entre empresas que compõem a economia brasileira.

As duas atividades desenvolvidas durante o estágio motivaram grandemente a definição do tema deste Trabalho de Formatura. Em particular, a utilização de técnicas do mercado de capitais para análises macroeconômicas é uma temática unificadora entre o estágio e o trabalho.

Apesar desta intrínseca relação, por questões de confidencialidade dos dados dos clientes do BCG, este trabalho não contém absolutamente nenhuma informação detalhada acerca das atividades desenvolvidas na empresa. Não obstante este fato, os resultados deste trabalho final vêm sendo utilizados pela empresa, com relevância crescente ao longo do estágio. Neste capítulo e nos subsequentes, as análises serão tratadas como independentes das atividades desenvolvidas na empresa, sem menções explícitas a nenhum dado confidencial do BCG.

1.3. Formulação do Problema

Compreender as atividades de geração, distribuição e transmissão de energia no Brasil é uma tarefa que envolve conhecimentos multidisciplinares e tange à intersecção entre Engenharia, Economia e Política. Adicionalmente, a relevância crucial deste setor para a vida de todos os brasileiros aumenta substancialmente a sua complexidade, em especial devido à atual crise hídrica no país.

Os modelos de análise de composição de portfólios, muito presentes no mercado de capitais, vão de encontro a esta problemática. Tais modelos são robustos – dada a enorme variação das condições macro e microeconômicas que regem os preços dos ativos –, abrangentes – de acordo com a natureza complexa de grande parte dos ativos financeiros –, eficientes – para responder rapidamente às demandas instáveis do mercado – e, por fim, eficazes – para atender às altas expectativas de desempenho dos seus utilizadores.

Esta conexão não-trivial entre o setor de energia elétrica e o mercado de capitais é exemplificada pela crescente utilização de modelos financeiros nos mercados de carbono e nos mercados “spot” de energia. Os mercados de carbono constituem um ambiente para negociação de créditos de emissão de gases de efeito estufa, cuja institucionalização foi acelerada pelo Protocolo de Quioto. Os mercados spot, por sua vez, são utilizados por empresas geradoras de

energia para comercialização da capacidade que não estiver sendo vendida para o SIN (Sistema Interligado Nacional). Nestes dois mercados, os modelos financeiros vêm sendo utilizados com respeitável sucesso, demonstrando a adaptabilidade destes modelos às complexas questões energéticas de uma empresa, de um país ou de um conjunto de países.

Compreender, analisar e formular melhorias para um setor de tal importância pressupõe a utilização de técnicas avançadas de Engenharia. Neste trabalho, optamos por empreender os conhecimentos de Pesquisa Operacional e de Teoria de Portfólios para estudar o setor energético, por serem multidimensionais, abrangentes e computacionalmente eficientes.

1.4. Objetivo do Trabalho

O presente Trabalho de Formatura objetiva aplicar conhecimentos de Engenharia a uma questão imprescindível para o desenvolvimento do Brasil: a escolha da matriz energética. Como 71% da energia brasileira vem de usinas hidrelétricas (Balanço Energético Nacional, 2014), a matriz brasileira é passível de muitas melhoras; em particular, diversificá-la traria diversas vantagens, tais como mitigação de riscos e aumento da disponibilidade de fontes alternativas. Neste contexto, o problema que este trabalho se propõe a estudar é a priorização das possíveis fontes de energia para o Brasil, através de modelos utilizados no mercado de capitais.

Para realizar este estudo, subdividimos o trabalho em dois objetivos; o primeiro funciona como uma etapa intermediária para a realização do segundo. Primeiramente, realizaremos um estudo comparativo de diferentes modelos de otimização de portfólio propostos pela literatura. Em seguida, já em posse deste estudo, poder-se-á de fato prosseguir com o objetivo principal do trabalho, que é o de estudar e propor melhorias à matriz energética brasileira.

Tivemos de subdividir o objetivo em duas partes subsequentes devido à absoluta escassez de dados sobre energia no Brasil. Este problema é frequentemente encontrado nos estudos realizados sobre o tema. Por exemplo, os autores de um dos artigos mais recentes que analisaram a matriz brasileira sob a perspectiva de portfólios (LOSEKANN *et al.*, 2013) tiveram de utilizar dados de outro artigo (AWERBUCH; SPENCER, 2007) para realizar etapas de cálculo fundamentais, não obstante o fato de os dados provirem de contextos completamente diferentes.

Esta dificuldade motivou o estudo preliminar que comporá a primeira parte deste trabalho. Realizamo-lo para diminuir as consequências da escassez de dados nos resultados finais e, desta

maneira, assegurar a plausibilidade das recomendações que faremos. Este primeiro estudo não será feito no mercado de energia, mas no mercado de capitais. Esta opção decorreu da incomensurável diferença existente entre a disponibilidade e a variedade de dados sobre estes mercados; se no mercado de energia dados são extremamente escassos, no mercado financeiro há uma miríade de informações. Desta maneira, o mercado de capitais atuará como um campo de treinamento para os estudos sobre o mercado de energia.

O primeiro objetivo será alcançado em duas etapas. Inicialmente, como os modelos de otimização estão sujeitos a incertezas nos cálculos de seus parâmetros, estudaremos a influência destas incertezas no desempenho global dos modelos. Em seguida, de acordo com os resultados da primeira etapa, proporemos uma otimização dos parâmetros do modelo de melhor desempenho, através de algoritmos evolutivos.

Finalmente, o segundo objetivo está organizado em três etapas. Primeiramente, levaremos em consideração todas as descobertas do estudo preliminar e aplicaremos a teoria de portfólios à problemática da geração de energia no Brasil. Em seguida, realizamos uma análise de sensibilidade à proporção de energia hidrelétrica no portfólio. Por fim, compararemos os nossos resultados com os de outros estudos similares.

1.5. Justificativa do Trabalho

Ao aliar energia e finanças, dois temas aparentemente desconexos, este Trabalho de Formatura – bem como muitos outros artigos e teses – propõe uma política de geração de energia no país que difere da atual. Desta forma, ele ultrapassa o âmbito da Engenharia ou da Matemática Aplicada e ataca diretamente a situação econômica do país. Considerando-se que a economia do país tem impacto direto sobre a sociedade como um todo, em termos sociais, geográficos e políticos, este trabalho passa então a ter um amplo campo de influência.

A utilização de técnicas de otimização, provenientes da Pesquisa Operacional, na problemática da geração de energia, no Brasil ou em qualquer outro país, garante simultaneamente economia de gastos, maior estabilidade energética e uso mais eficiente da matriz produtora de energia (AWERBUCH, 2000).

Em particular, as técnicas de otimização de portfólios podem ser extremamente relevantes para a questão energética, já que podem ser aplicadas tanto em um contexto puramente

financeiro quanto em ativos de natureza não-financeira. Quando aplicadas a portfólios de investimentos financeiros, elas permitem alocar os recursos de maneira a maximizar o retorno para um dado risco. Pode-se, alternativamente, minimizar o risco para um dado retorno. Como será mostrado nas próximas seções, diversos modelos para realizar esta otimização existem, com suas respectivas vantagens e inconvenientes.

Quando aplicadas a portfólios de outra natureza, a interpretação de variáveis como “risco” e “retorno” costuma depender fortemente do contexto de aplicação. No caso da geração de energia, veremos que o “retorno” corresponde à quantidade de energia gerada por cada unidade monetária investida.

Os conceitos de otimização e diversificação de portfólio têm sido essenciais tanto no desenvolvimento e na compreensão dos mercados financeiros quanto nas análises de tomada de decisão (KOLM *et al.*, 2013). O maior avanço neste campo ocorreu em 1952, com o advento do Modelo da Média-Variância, conhecido como Teoria Moderna de Portfólios (MARKOWITZ, 1952), que vem sendo há mais de 60 anos o pilar de sustentação unificador de quase todos os modelos de otimização de portfólio. Este modelo forneceu uma resposta consistente à questão fundamental de alocação de investimentos: como um investidor deve alocar os seus recursos entre todas as opções possíveis? Primeiramente, Markowitz quantificou retorno e risco de um ativo, através das estatísticas retorno esperado e desvio-padrão. Em seguida, ele sugeriu que estas medidas deveriam ser consideradas conjuntamente, de maneira que o critério de escolha seja dependente da relação entre estas duas variáveis.

Diversos modelos foram propostos posteriormente, utilizando abordagens matemáticas mais avançadas. Notavelmente, Rockafellar e Uryasev (2000) propuseram um modelo de otimização utilizando o CVaR (*Conditional Value-at-Risk*) que é em geral considerado uma evolução do modelo de Markowitz. Outras evoluções do modelo incluem o uso de otimização bayesiana (MOČKUS, 1972).

Quanto à geração de energia no Brasil, sabemos que o país desfruta de uma grande riqueza fluvial. Historicamente, este fato influenciou o país a adotar uma matriz energética fortemente focada em hidroelétricas. No entanto, outras opções existem e, como veremos ao longo do Trabalho, podem ser mais interessantes em muitos cenários. Como ocorreu em todos os países, a atual combinação de formas de geração de energia no Brasil é fruto de uma confluência de fatores: interesses políticos, ocupação territorial, clima e vegetação, desenvolvimento

econômico, entre outros. A atual crise hídrica no país reatesta a importância deste tema, demonstrando que uma matriz energética fortemente dependente de água pode ser uma escolha demasiadamente arriscada e imprudente para uma nação das proporções do Brasil.

É natural e previsível que o “ótimo” que encontraremos matematicamente dificilmente conseguirá ser integralmente posto em prática. Evidentemente, isto não diminui a sua importância como orientação para futuras escolhas de investimentos – neste sentido, o ótimo que encontraremos atua como um alvo ideal a ser alcançado.

1.6. Estrutura do Trabalho

O primeiro capítulo apresenta a empresa do estágio e a formulação do problema, bem como o objetivo e a justificativo do trabalho. O capítulo seguinte apresenta, concomitantemente, os fundamentos teóricos que atuarão como pilares de sustentação do trabalho e a revisão bibliográfica. Decidimos permear os tópicos teóricos com a revisão bibliográfica para diminuir a compartimentalização do trabalho.

Inicia-se a parte original do trabalho no Capítulo 3. Primeiramente, são apresentados os fatores que definem os modelos que serão estudados. Tais fatores seguem uma categorização, visando a facilitar as análises posteriores. Ainda no Capítulo 3, é realizado um estudo comparativo dos modelos estudados, que correspondem tanto a modelos clássicos quanto a modelos modernos da Literatura Acadêmica. Este estudo focará em duas métricas; a saber, desempenho ao longo do tempo e variação da composição do portfólio ao longo do tempo.

Ainda no terceiro capítulo, já em posse do modelo com melhor desempenho, far-se-á uso de Algoritmos Evolutivos para a otimização dos parâmetros do modelo, objetivando melhorá-lo segundo as métricas de desempenho estabelecidas.

O Capítulo 4 apresenta a aplicação do modelo de otimização de portfólio usado no mercado financeiro à matriz energética brasileira. Em seguida, são discutidos os resultados e as consequências, em termos de políticas governamentais. O foco é dado nos conceitos por trás da aplicação deste algoritmo e nas adaptações realizadas, de tal forma a esclarecer as vantagens e os inconvenientes de tal abordagem.

Por fim, o Capítulo 5 apresenta as conclusões.

2. FUNDAMENTOS TEÓRICOS

2.1. Pesquisa Operacional

Apesar de se fundamentar em conhecimentos muito anteriores, o campo de *pesquisa operacional*, também conhecido como *ciência da gestão*, tornou-se uma disciplina independente no fim da década de 1930. De acordo com van Hoesve (2005), grande parte da motivação para esta disciplina veio das operações militares durante a 2ª Guerra Mundial, daí o nome de “operacional”. Apesar da dificuldade de encontrar uma definição que englobe todos os ramos de pesquisa, o *Institute for Operations Research and the Management Sciences* usa a seguinte definição:

“Pesquisa Operacional e as Ciências da Gestão são as disciplinas profissionais que lidam com a aplicação de tecnologia da informação para tomadas de decisão conscientes.”
(Tradução livre.)

Na prática, Pesquisa Operacional designa um conjunto de algoritmos e técnicas, como programação linear, programação inteira, programação não-linear e algoritmos evolutivos. Estes algoritmos são usados para avaliar linhas de ação alternativas e encontrar as soluções que melhor servem aos objetivos de indivíduos ou organizações. Dada a sua força econômica comprovada, pesquisa operacional foi – e ainda é – fruto de grandes desenvolvimentos.

Historicamente, muitas das técnicas de pesquisa operacional têm se provado extremamente úteis para resolver problemas de cunho prático, como no campo da logística ou das finanças. Atualmente, com o apoio de recursos computacionais de crescente capacidade de processamento, a pesquisa operacional permite utilizar enormes quantidades de dados para analisar problemas extremamente complexos, sempre visando a garantir que a solução encontrada seja a mais interessante para o usuário, de acordo com o modelo utilizado.

Além de servir aos propósitos de empresas, a pesquisa operacional também tem sido utilizada de maneira promissora por governos, sobretudo dos EUA. No Brasil, esta prática ainda não é bem disseminada, porém este Trabalho de Formatura será um exemplo concreto do poder da utilização destas técnicas para o bem público.

2.2. Preço, Retorno e Risco de Ativos

Antes de apresentarmos o modelo de Markowitz, que será a base deste trabalho, definiremos alguns conceitos que serão muito usados posteriormente².

2.2.1 Preço

Numa primeira abordagem, consideraremos ativos negociados no mercado de ações para exemplificar os conceitos importantes para otimização de portfólios. No caso dos ativos financeiros, chamamos P_t o preço no instante t . Geralmente, toma-se o preço de fechamento do ativo, no fim do dia, semana ou mês. O preço de fechamento é o último preço a que um título foi comercializado no intervalo considerado. Outros preços importantes, e que serão utilizados nas seções subsequentes, são os de alta (*high*) e baixa (*low*) em um determinado período, correspondendo respectivamente ao preço mais alto e ao mais baixo a que o ativo foi negociado. Por enquanto, focaremos apenas no preço de fechamento P_t .

Consideraremos que o universo de investimentos possíveis seja composto de n ativos com retornos r_1, \dots, r_n , onde $\forall i, r_i$ é uma variável aleatória.

2.2.2 Retorno

Para estimar o retorno futuro de um ativo, a maneira mais simples é calcular a média histórica. Portanto, para cada ativo i , o retorno esperado é dado numa primeira abordagem por:

$$E[r_i] = \mu_i$$

Na equação acima, $E[r_i]$ designa a esperança do retorno do ativo i e μ_i é a média aritmética do vetor de retornos do ativo i . Cada componente deste vetor será calculado através da seguinte fórmula, que calcula o retorno entre os instantes $t - 1$ e t :

$$R_t = \ln \frac{P_t}{P_{t-1}}$$

O retorno calculado é logarítmico (natural), pois a distribuição seguida é mais próxima de uma distribuição normal do que se o retorno fosse calculado da maneira tradicional (FABOZZI *et al.*, 2007). Similarmente ao retorno tradicional, definido pela equação abaixo,

² Estrutura baseada em Félix Roudier (2007).

$$R_t^{trad} = \frac{P_t - P_{t-1}}{P_{t-1}}$$

o retorno logarítmico tem a conveniente propriedade de ser negativo quando o preço do ativo caiu e positivo quando este subiu.

Podemos também calcular o retorno esperado de todo um portfólio de investimentos \mathbf{P} , composto por n ativos, em que ω_i é a porcentagem dos recursos alocada a cada ativo:

$$\mu_P = E[r_P] = \sum_{i=1}^n \omega_i E[r_i] = \sum_{i=1}^n \omega_i \mu_i$$

2.2.3 Risco

A. Variância

Existem várias formas de calcular o risco de ativos. A mais clássica é a variância dos retornos dos ativos, por ter sido estabelecida no modelo de Markowitz.

Para uma dado ativo i , o risco é, portanto, dado por:

$$\sigma_i^2 = \text{Var}(r_i) = E[(r_i - E[r_i])^2] = E[(r_i - \mu_i)^2]$$

A covariância entre os ativos i e j é definida como:

$$\sigma_{ij} = \text{Cov}(r_i, r_j) = E[(r_i - E[r_i])(r_j - E[r_j])] = E[(r_i - \mu_i)(r_j - \mu_j)]$$

A correlação entre os ativos i e j é definida como:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$$

Estamos agora em condições de calcular o risco (variância) de um portfólio de investimentos \mathbf{P} , composto por n ações, em que ω_i é a porcentagem dos recursos alocada a cada ativo (GRUBER *et al.*, 2014, p. 56):

$$\sigma_P^2 = E[(r_P - \mu_P)^2] = E\left[\left(\sum_{i=1}^n \omega_i r_i - \sum_{i=1}^n \omega_i \mu_i\right)^2\right] = E\left[\left(\sum_{i=1}^n \omega_i (r_i - \mu_i)\right)^2\right]$$

$$\sigma_p^2 = E \left[\sum_{i=1}^n \omega_i^2 (r_i - \mu_i)^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \omega_i \omega_j (r_i - \mu_i)(r_j - \mu_j) \right]$$

$$\sigma_p^2 = \sum_{i=1}^n (\sigma_i \omega_i)^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \omega_i \omega_j \sigma_{ij}$$

Um dos problemas de se estimar o risco através da variância é que ela não leva em consideração o sinal da variação do preço de um ativo. Um ativo pode ter alto risco se o preço estiver subindo significativamente vários dias seguidos, por mais que, intuitivamente, isto seja uma boa indicação para os investidores. Para corrigir este problema, outras medidas foram criadas, notavelmente o VaR e o CVaR. Outras medidas ainda mais avançadas, como o EVaR (Entropic Value-at-Risk), existem mas não serão abordadas neste Trabalho de Formatura (consultar, por exemplo, AHMADI-JAVID, 2012).

B. VaR (Value-at-Risk)

Intuitivamente, o VaR (Value-at-Risk) é o cálculo da maior porcentagem de perda que um investidor pode ter no período seguinte, com um nível de confiança de α . O VaR possui três componentes: um período de tempo, um nível de confiança e uma porcentagem de perda.

Evidentemente, o intervalo considerado (dia, semana, mês, ano) e o nível de confiança (90%, 95%, 99%) são escolhidos para cada situação analisada. Há diversas maneiras de calcular o VaR. Uma das mais diretas é através do histograma da frequência das faixas de valor em que se encontraram os riscos históricos, tal qual a figura abaixo:

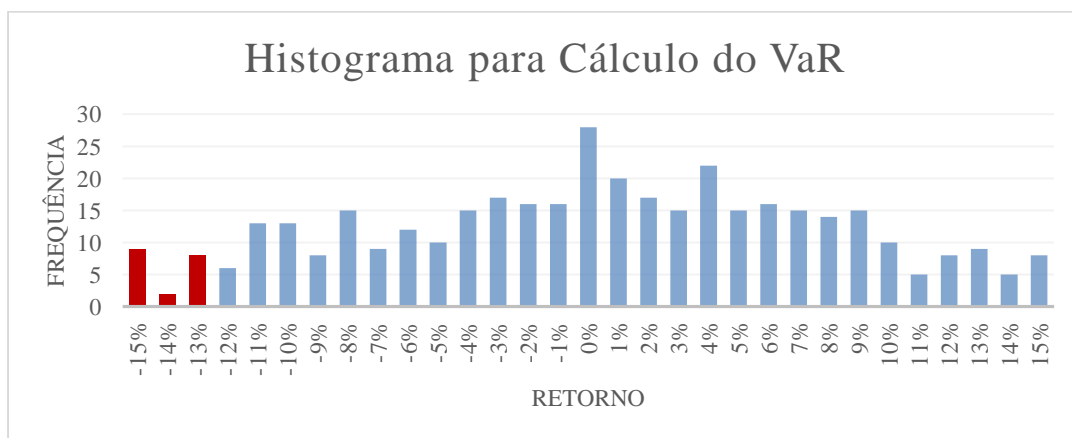


Figura 2 – Cálculo do VaR (Value-at-Risk)

A figura acima é o histograma de 400 números aleatórios, gerados segundo uma distribuição normal ($\mu = 0, \sigma = 0.08$). Para calcular o VaR relativo a um dia com um nível de confiança de 5%, basta encontrar a menor perda que ocorreu entre os 20 menores valores (5%). Este valor corresponde a 13%. Portanto, podemos estimar, com 95% de confiança, que a pior perda diária não vai exceder 13%.

Em conclusão, definimos o VaR com confiança de α como o $(100 - \alpha)$ -ésimo percentil da distribuição de retornos históricos para uma distribuição X como (ROCKAFELLAR, R. T.; URYASEV, S., 2000):

$$VaR_{\alpha} = \min\{z | F_X(z) \geq \alpha\}$$

$$\text{onde } F_X(z) = P\{X \leq z\}$$

C. CVaR (Conditional Value-at-Risk)

O CVaR (Conditional Value-at-Risk) deriva do VaR, numa tentativa de ampliar a sua utilidade e conferir propriedades matemáticas mais interessantes. Intuitivamente, o VaR calcula quão ruim pode ser o desempenho de um portfólio, enquanto o CVaR calcula quanto dinheiro será perdido, caso o portfólio tenha um desempenho ruim.

O CVaR é uma média ponderada entre o VaR e o valor das perdas que excedem o VaR. Para ilustrar a vantagem desta medida, podemos supor que a distribuição das perdas tenha uma cauda (à esquerda) que, apesar de ocorrer com baixa frequência, corresponda a uma perda enorme. Nestes casos, o VaR é incapaz de detectar quão grande será a perda nos piores casos, enquanto que o CVaR não terá esta deficiência. A figura abaixo, adaptada da anterior, ilustra esta situação. Podemos reparar que o tamanho das perdas, nos piores casos, é maior do que no histograma anterior.



Figura 3 – Ineficiência do VaR

Para uma distribuição X e um nível de confiança α , o CVaR é definido como (ROCKAFELLAR, R. T.; URYASEV, S, 2000):

$$CVaR_{\alpha}(X) = \int_{-\infty}^{\infty} z dF_X^{\alpha}(z)$$

Onde:

$$F_X^{\alpha}(z) = \begin{cases} 0, & \text{quando } z < VaR_{\alpha}(X) \\ \frac{F_X(z) - \alpha}{1 - \alpha}, & \text{quando } z \geq VaR_{\alpha}(X) \end{cases}$$

$$F_X(z) = P\{X \leq z\}$$

2.3. Modelo de Markowitz

2.3.1 Histórico

O primeiro modelo importante de otimização de portfólios foi desenvolvido por Harry Max Markowitz, um economista americano que recebeu o *John von Neumann Theory Prize* em 1989 e o *Nobel Memorial Prize in Economic Sciences* em 1990.

Em 1952, Markowitz publicou o artigo clássico “Portfolio Selection”, no *Journal of Finance*. Neste artigo, Markowitz deriva o que é chamado de Teoria da Média-Variância, afirmando que um investidor racional deve sempre maximizar o retorno para um dado nível de risco ou minimizar o risco para um dado retorno.

2.3.2 Fronteira eficiente

A aplicação deste princípio leva à criação de uma “fronteira eficiente”, na qual todos os pontos obedecem à regra supracitada: para um dado retorno, todos os pontos à direita da fronteira representam riscos excessivos; inversamente, para um dado risco, todos os pontos abaixo da fronteira apresentam retornos menores do que o retorno do portfólio da fronteira.

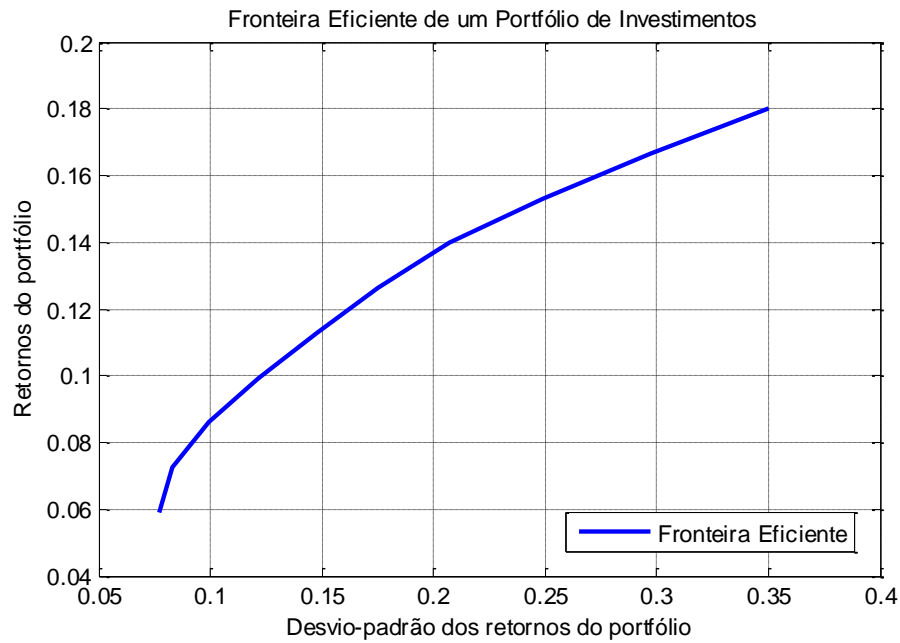


Figura 4 – Ilustração da fronteira eficiente

Neste método, a medida do risco do ativo utilizada é a variância. Utilizamos o desvio-padrão nas representações gráficas para facilitar a comparação com o retorno, já que ambos são expressos nas mesmas unidades.

2.3.3 Formulação matemática

A formulação clássica do modelo de Markowitz é (GRUBER *et al.*, 2014, p. 102):

Minimizar (a variância):

$$\sigma_P^2 = \bar{\omega}^T \Sigma \bar{\omega}$$

Sujeito às restrições:

$$\bar{\omega}^T \bar{\mu} = r_0$$

$$\omega_i \geq 0, \forall i$$

$$\bar{1}^T \bar{\omega} = 1$$

Onde:

σ_P^2 é a variância do portfólio P ;

$\vec{\omega}$ é o vetor dos pesos de cada elemento do portfólio;

$\vec{\mu}$ é o vetor dos retornos esperados de cada elemento do portfólio;

$\vec{1}$ é o vetor $(1, 1, \dots, 1)$;

r_0 é retorno desejado para o portfólio;

Σ é a matriz das covariâncias dos elementos do portfólio.

A matriz das covariâncias Σ é uma matriz simétrica, cuja diagonal é a variância de cada elemento do portfólio e o restante é formado pelas covariâncias dos elementos dois a dois, ou seja, $\sigma_{ij} = \text{Cov}(r_i, r_j)$. Este parâmetro do modelo será objeto de estudo deste artigo, por estar sujeito a incertezas. Pretende-se, portanto, analisar o grau de impacto das incertezas na solução ótima do modelo.

O modelo de Markowitz consiste em, dado um retorno-alvo r_0 , minimizar-se a variância σ_p^2 do portfólio de investimentos.

2.3.4 Solução analítica

No caso elementar apresentado acima, o modelo de Markowitz possui uma solução analítica, dada por (ROUDIER, 2007):

$$\vec{\omega} = \Sigma^{-1}(\vec{\mu} \vec{1}) A^{-1} \begin{pmatrix} r_0 \\ 1 \end{pmatrix}$$

Onde A é definida por:

$$A = \begin{pmatrix} \vec{\mu}^T \Sigma^{-1} \vec{\mu} & \vec{\mu}^T \Sigma^{-1} \vec{1} \\ \vec{\mu}^T \Sigma^{-1} \vec{1} & \vec{1}^T \Sigma^{-1} \vec{1} \end{pmatrix}$$

Nos casos mais complexos, quando da existência de restrições sobre a composição do portfólio ou da consideração dos custos de transação envolvidos, não há soluções analíticas. Nestes casos, é necessário recorrer a algoritmos numéricos, cuja aplicação pode ser custosa computacionalmente. O custo computacional de utilizar estes algoritmos numéricos, especialmente no caso da utilização do CVaR como medida de risco, será recorrente nos capítulos subsequentes.

Mesmo após mais de 60 anos de publicação do artigo de Markowitz e de o modelo estar sujeito a incertezas, devido às estimações realizadas no cálculo do vetor de retornos $\bar{\mu}$ e da matriz de covariâncias Σ , o modelo da Média-Variância ainda é a abordagem padrão para a resolução deste tipo de problema. No entanto, muitos pesquisadores vêm mostrando que o método não tem um bom desempenho, dado que ele põe peso excessivo para ativos com alto retorno esperado, independentemente de possíveis erros nos dados de entrada. Além disso, os portfólios gerados pelo modelo de Markowitz costumam ter um custo de transação muito grande, devido à excessiva realocação dos ativos. Com isto, muitos outros modelos foram estabelecidos, notavelmente a otimização robusta e o modelo de Black-Litterman.

Entre os pesquisadores que lidaram com os problemas do modelo da Média-Variância, encontramos Michael J. Best e Robert R. Grauer (*“On the sensitivity of mean-variance-efficence portfolios to changes in asset means: Some analytical and computacional results”*), bem como Richard C. Green e Burton Hollifield (*“When will mean-variance efficient portfolios be well diversified?”*).

2.4. Otimização Bayesiana

A abordagem clássica de Markowitz é composta de duas etapas: estima-se as distribuições que descrevem o mercado e então realiza-se a otimização, como se as distribuições fossem exatas. No entanto, as distribuições contêm incertezas, tornando a otimização sub-ótima. Adicionalmente, como o processo de otimização é extremamente sensível aos dados de entrada, a sub-otimalidade devida aos erros de estimação pode ser considerável (MEUCCI A., 2005).

Para contornar este problema, otimização robusta incorpora no próprio modelo a existência de incertezas: o investidor escolhe a melhor alocação de portfólio no pior cenário possível dentro de um certo conjunto de incerteza (*uncertainty set*). A otimização bayesiana é, portanto, uma das maneiras de se aumentar a robustez das otimizações realizadas.

Ainda de acordo com Meucci (2005), a formulação do modelo de média-variância, na versão robusta, é:

$$\omega_i = \operatorname{argmax} \left\{ \min_{\mu \in \bar{\theta}_\mu} \{\omega' \mu\} \right\}$$

Sujeito a:

$$\omega \in \mathcal{C}, \max_{\Sigma \in \widehat{\theta}_{\Sigma}} \{\omega' \Sigma \omega\} \leq v_i$$

Onde:

- $\widehat{\theta}_{\Sigma}$ e $\widehat{\theta}_{\mu}$ são, respectivamente, os conjuntos de incerteza para Σ e μ .
- \mathcal{C} é um conjunto de restrições sobre os pesos do portfólio.
- v_i são variâncias-alvo para o retorno do portfólio.

O importante a se notar sobre esta formulação é que há restrições sobre a variância do portfólio, de forma que as incertezas já são levadas em consideração durante a otimização.

2.5. Rugosidade

Os custos de transação podem ser grandes o suficiente para corroer significativamente a lucratividade de um dado portfólio, devido às taxas fixas e variáveis cobradas pelas corretoras de valores e pelas bolsas de valores. Consequentemente, quando calculada a evolução da composição do portfólio ótimo ao longo tempo, temos que, quanto menor é a variação desta composição, melhor para o investidor. Para medir a magnitude das variações, criamos uma métrica, a qual denominaremos de rugosidade R :

$$R = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^{M-1} |\omega_{i,j+1} - \omega_{i,j}|$$

$\omega_{i,j}$ é um número real entre 0 e 1 (ou 0 e 100%), que traduz a porcentagem do ativo i presente no portfólio ótimo do dia j . Na expressão da rugosidade acima, o portfólio é composto por um total de N ações e o período da análise é de M dias. A figura abaixo mostra a evolução dos $\omega_{i,j}$ se aplicarmos o modelo clássico de Markowitz ao longo dos 70 dias de negociação que precedem 15 de maio de 2014, para um portfólio composto pelas 10 ações cujos símbolos estão à direita (bolsa NYSE):

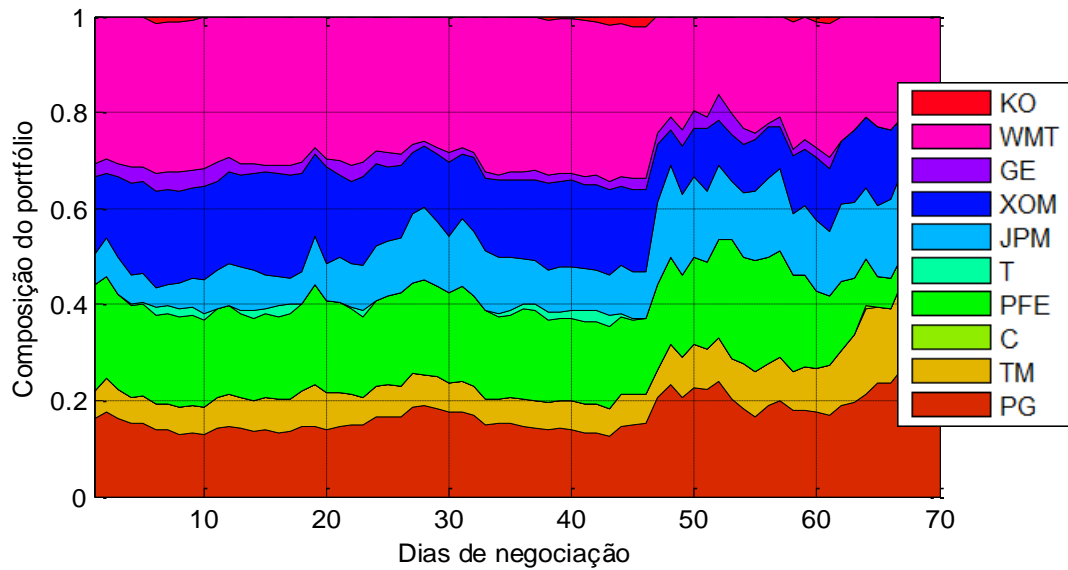


Figura 5 – Exemplo de evolução da composição do portfólio ótimo

A rugosidade é, portanto, uma medida da frequência e da amplitude dos “picos” da figura. Para os 35 primeiros dias de negociação acima, em que os portfólios estiveram relativamente estáveis, $R = 1.04$. Para os 35 dias finais, onde houve mais instabilidade na composição, a medida de rugosidade cresceu 45%, atingindo $R = 1.51$.

2.6. Desempenho

Para o cálculo do desempenho, utilizaremos o retorno tradicional R^{trad} . Fazeremos para poder comparar com outros investimentos, que são expressos desta forma, como o retorno da poupança. No entanto, removeremos por simplicidade o sobrescrito **trad** e adicionaremos os índices **i** e **j**, correspondendo ao ativo **i** e ao dia **j**:

$$R_{i,j} = \frac{P_{i,j} - P_{i,j-1}}{P_{i,j-1}}$$

Na fórmula acima, $P_{i,j}$ corresponde ao preço do ativo **i** no dia **j**.

Para medir o desempenho do modelos, calculamos primeiramente os desempenhos em cada dia. Definimos o desempenho no dia **j**, d_j , como a média dos retornos dos ativos neste dia, ponderados pela parcela de investimento em cada uma:

$$\mathbf{d}_j = \mathbf{1} + \sum_{i=1}^N \omega_{i,j} \cdot \mathbf{R}_{i,j}$$

Em seguida, compomos estes desempenhos ao longo do tempo. Definimos \mathbf{d}'_j como o produto dos desempenhos de todos os dias anteriores até o presente dia, para simular o que aconteceria se estivéssemos de fato investindo seguidamente na bolsa de valores:

$$\mathbf{d}'_j = \left(\prod_{k=1}^j \mathbf{d}_k \right) - \mathbf{1}$$

Na expressão dos \mathbf{d}_j , somamos 1 aos $\mathbf{r}_{i,j}$ para evitar que haja algum \mathbf{d}_j igual a zero, o que tornaria todos os \mathbf{d}'_j subsequentes também iguais a zero, dada o produtório acima. Por fim, a medida de desempenho \mathbf{D} que propomos é a soma de todos os \mathbf{d}'_j :

$$\mathbf{D} = \sum_{j=1}^M \mathbf{d}'_j$$

Como \mathbf{M} é uma constante quando comparamos os modelos, esta medida de desempenho pode ser interpretada como a média dos desempenhos ao longo do tempo (multiplicada por \mathbf{M} , naturalmente). Podemos, alternativamente, visualizar \mathbf{D} como a área sob a curva formada pelos \mathbf{d}'_j ao longo do tempo.

Um exemplo de tal curva, quando utilizamos o modelo de otimização clássico de Markowitz ao longo dos 70 dias de negociação que precedem 15 de maio de 2014, com as mesmos ativos que foram listadas na *Figura 1*, é mostrada abaixo:

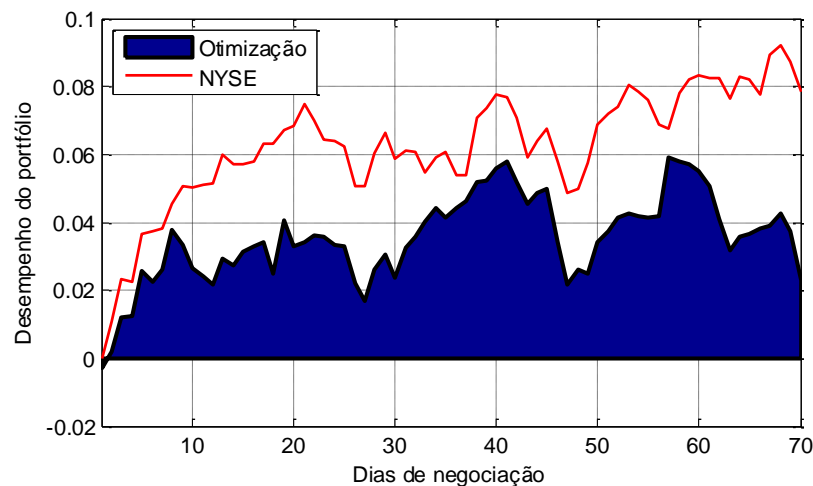


Figura 6 – A curva formada pelos d'_j .

Subtraímos 1 na expressão dos d'_j para garantir que esta integral seja de fato negativa quando o desempenho é predominantemente negativo ao longo do tempo. Se não o fizéssemos, a curva acima estaria deslocada verticalmente, gerando possivelmente medidas de desempenho positivas em casos onde esta seria intuitivamente negativa.

Em detrimento de calcular o desempenho apenas com base no valor final do portfólio, como é comum em investimentos quotidianos, a medida proposta utiliza todos os valores intermediários. Fizemos deliberadamente esta escolha, para poder diferenciar o desempenho de portfólios que possuam o mesmo valor no último dia da análise, porém com trajetórias diferentes ao longo do caminho.

2.7. Algoritmos Evolutivos

Segundo Mitchell (1996), algoritmos evolutivos são mecanismos heurísticos que mimetizam processos biológicos para resolver problemas matemáticos. Pertencentes ao campo de Inteligência Artificial da Ciência da Computação, estes métodos são geralmente usados para gerar soluções úteis – possivelmente ótimas – para problemas de busca ou otimização.

Dentre os processos naturais que são mimetizados matematicamente, encontramos reprodução, herança genética, mutação, seleção e *crossover*. Inicialmente, geramos aleatoriamente um conjunto de possíveis soluções para o problema, que comporão uma “população”. Em seguida, esta população será avaliada segundo uma função de *fitness* (ou aptidão), que determinará quais são as melhores candidatas. A próxima população será então

obtida através dos processos citados acima, combinando as melhores soluções (reprodução), passando características para os “filhos” (herança genética), alterando as soluções aleatoriamente (mutação), etc.

Segundo pontos levantados por Mitchell (1996), Goldberg (1989) e Roudier (2007), há pontos favoráveis e desfavoráveis na utilização de algoritmos evolutivos. Em particular, é um método excelente para avaliar funções descontínuas ou não-lineares, para as quais os métodos convencionais de otimização são ineficazes ou mesmo inaplicáveis, ou funções “caixa-preta”, como é o caso da que estudaremos neste artigo. Por outro lado, estes métodos têm o viés de não garantir a otimalidade da solução fornecida, dada a sua natureza heurística.

2.8. Redes Neurais

2.8.1 História

Esta seção é baseada em Schmidhuber *et al.* (2012). No caso de classificação não-linear, os métodos mais simples de aprendizagem automática e inteligência artificial são pouco eficazes, devido a, por exemplo, problemas de sobreajuste (*overfitting*). É frequentemente o caso da regressão logística, um dos modelos mais simples de classificação

Para contornar este problema nas aplicações mais complexas, diversas alternativas foram propostas ao longo do tempo, tais que regularização ou redução do número de variáveis explanatórias, porém poucas destas alternativas são capazes de aumentar significativamente a eficiência destes métodos.

Uma alternativa mais poderosa para este tipo de problema são redes neurais, também chamadas de redes neurais artificiais. Este método de aprendizagem é uma tentativa de simular o funcionamento do cérebro humano. Apesar de ter sido amplamente popular na década de 80, a popularidade diminuiu próximo do ano 2000. Recentemente, no entanto, houve um aumento significativo do uso de redes neurais, graças ao advento de novas técnicas mais poderosas e computacionalmente eficazes.

2.8.2 Conceito

Como citado acima, redes neurais simulam o funcionamento do cérebro humano – mais especificamente, dos neurônios que compõem o sistema nervoso humano. A figura abaixo indica os principais componentes dos neurônios humanos:

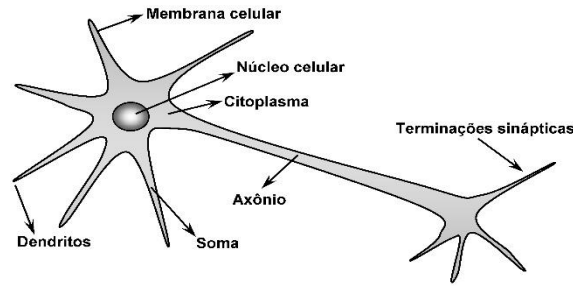


Figura 7 – Esquema de um neurônio humano.

Os dendritos funcionam como fios de entrada de dados (*input wires*) e os axônios como fios de saída de dados (*output wires*). Juntamente com os outros componentes, os dendritos e os axônios permitem que os neurônios funcionem como unidades relativamente autônomas de processamento, que se comunicam com os outros neurônios através do envio de impulsos elétricos. As terminações sinápticas são responsáveis pela transmissão destes impulsos elétricos.

2.8.3 Modelo

Definimos o vetor \mathbf{x} como o vetor entrada e o vetor $\boldsymbol{\theta}$ como um vetor de parâmetros. Por exemplo, para um modelo com 3 variáveis de entrada, \mathbf{x} e $\boldsymbol{\theta}$ seriam:

$$\mathbf{x} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$$

Onde x_0 é sempre igual a 1, comumente chamado de “*bias unit*”. Definimos também a função “sigmoide” ou “logística” $h_{\boldsymbol{\theta}}(\mathbf{x})$ de acordo com a fórmula abaixo:

$$h_{\boldsymbol{\theta}}(\mathbf{x}) = g(\boldsymbol{\theta}^T \mathbf{x}) = \frac{1}{1 + e^{-\boldsymbol{\theta}^T \mathbf{x}}}$$

O comportamento desta função nos permitirá classificar binariamente as entradas. O gráfico abaixo ilustra o comportamento da função:

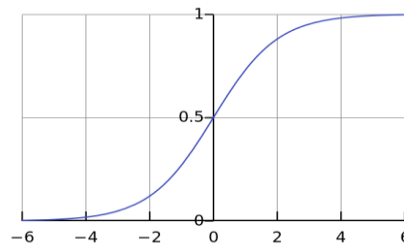


Figura 8 – Gráfico da função sigmoide ou “logística”

Para valores negativos de x , a função retorna valores abaixo de 0.5. Inversamente, para valores positivos de x , a função retorna valores acima de 0.5. Além disso, a função sigmoide possui duas assíntotas horizontais, nas ordenadas **0** e **1**. Esta propriedade nos permite classificar a entrada em duas categorias, uma delas correspondendo aos valores de $h_{\theta}(x)$ abaixo de 0.5 e a outra correspondendo aos valores de $h_{\theta}(x)$ acima de 0.5. Além de classificar os valores de x , a função sigmoide também permite quantificar a classificação, isto é, verificar qual a probabilidade de que cada uma das classificações seja verdadeira.

Suponha que estejamos tentando classificar e-mails em “spam” e “não-spam”. Após parametrizarmos corretamente o modelo, para cada e-mail – correspondente a um certo vetor x – teremos um valor de $h_{\theta}(x)$. Neste caso, podemos definir o modelo de tal forma que os valores de $h_{\theta}(x)$ correspondam à probabilidade de que o e-mail seja spam. Desta maneira, se $h_{\theta}(x)$ for maior do que 0.5, a probabilidade de que este e-mail seja spam é maior do que 50% e portanto classificamo-lo como spam. Senão, classificamo-lo como não-spam.

2.8.4 Representação

Abaixo, vemos uma possível representação de uma rede neural:

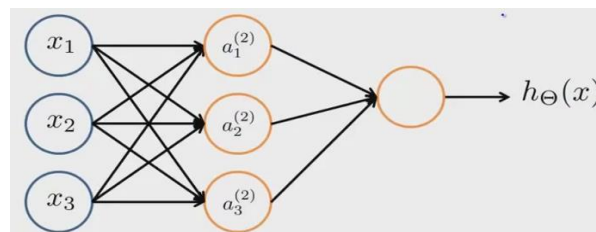


Figura 9 – Representação de uma rede neural

- O vetor x é chamado de “camada de entrada” (*input layer*) ou simplesmente “camada 1”.
- O vetor a corresponde à “camada escondida” (*hidden layer*) ou simplesmente “camada 2”.
- O resultado final, obtido através da função $h_{\theta}(x)$, é chamado de “camada de saída” (*output layer*) ou simplesmente “camada 3”.
- Todas as camadas que estiverem entre a entrada e a saída são chamadas de “escondidas”.
- Como o valor de x_0 é sempre igual a 1, não costumamos representá-lo no modelo.

2.8.5 Arquitetura

Chamamos de arquitetura o conjunto de camadas de uma rede neural. Por exemplo, a figura a seguir representa uma rede neural com uma camada adicional em relação à figura anterior.

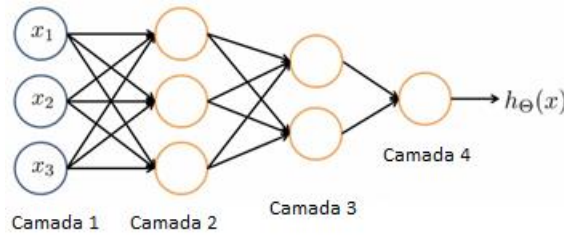


Figura 10 – Arquitetura de uma rede neural

Neste caso, o vetor \mathbf{x} corresponde novamente à camada 1, porém a saída corresponde à camada 4. Adicionalmente, temos duas camadas escondidas, correspondentes respectivamente aos vetores tri e bidimensionais (camadas 2 e 3). Recomenda-se que as camadas escondidas tenham um número de unidades da mesma ordem de grandeza que as unidades de entrada. Quanto mais camadas forem adicionadas, mais complexas ficam as possibilidades de aprendizado. Quanto à nomenclatura, denominamos redes neurais não-lineares auto-regressivas (NAR) as redes que utilizarem a série histórica de um certo dado para tentar prever o próximo dado desta série. Esta será o tipo de rede neural que utilizaremos neste trabalho.

2.8.6 Função custo

Definimos “função custo” como um indicador da qualidade do aprendizado. Quanto menor o valor da função, mais próximos estamos de um aprendizado perfeito. Portanto, para parametrizar o modelo, procederemos à minimização da função custo J , dada, no caso de uma rede neural, por:

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{k=1}^K y_k^{(i)} \log(h_{\theta}(x^{(i)}))_k + (1 - y_k^{(i)}) \log(1 - (h_{\theta}(x^{(i)}))_k) \right] + \frac{\lambda}{2m} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (\theta_{ji}^{(l)})^2$$

Onde:

- L é o número de camadas na rede.
- s_l é o número de unidades da camada l , descontando a unidade que é sempre 1.
- K é o número de classes, no caso de um classificador multi-classe ($s_L = K$).
- O termo $\frac{\lambda}{2m} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (\theta_{ji}^{(l)})^2$ corresponde à regularização do modelo, para evitar sobreajuste (overfitting).

3. ESTUDO COMPARATIVO DE MODELOS DE OTIMIZAÇÃO DE PORTFÓLIO SELECIONADOS

3.1. Metodologia

O estudo comparativo deste capítulo está organizado em diversas etapas. Inicialmente, apresentaremos seis fatores que, quando combinados, determinam o funcionamento dos modelos de otimização que estudaremos. Estes fatores possuem naturezas diversas; alguns correspondem ao tipo de dados que serão utilizados, outros se referem à incorporação de incertezas no modelo e, ainda, alguns se referem aos métodos de cálculos dos parâmetros.

Apresentados estes fatores, prosseguiremos à definição dos modelos em si. Naturalmente, cada modelo é fruto de uma combinação específica dos seis fatores supracitados. Em seguida, realizaremos uma análise destes modelos, do ponto de vista das métricas de rugosidade e desempenho definidas no Capítulo 2. Esta análise revelará qual o modelo com melhor desempenho, sob a nossa perspectiva.

Por fim, a última etapa consistirá da utilização de algoritmos evolutivos para otimizar dois dos fatores, no caso do modelo de melhor desempenho. Veremos que esta abordagem trará vantagens significativas em termos de retorno ao longo do tempo.

3.2. Fatores considerados nos modelos

Três dos seis fatores que definem os modelos são “externos” a eles, ao passo que estão relacionados apenas ao seu modo de utilização. Tais fatores serão doravante denominados de **extrínsecos**. Os outros três fatores são “internos” aos modelos, dado que estão diretamente relacionados às hipóteses e modelos matemáticos utilizados no seu funcionamento. Estes serão chamadas de **estruturais**.

3.2.1 Fatores extrínsecos

Enunciaremos primeiramente os fatores extrínsecos ao modelo, pois são de mais fácil compreensão.

A. Tipo de janela (móvel ou fixa).

Refere-se à maneira como atualizamos a base de dados quando da chegada de novas informações. Se excluirmos a informação mais antiga em detrimento da mais nova, denominamos o método como “janela móvel”, pois o intervalo de dados analisados está se alterando no tempo. Se adicionarmos a nova informação sem excluir a mais antiga, denominamos “janela fixa”.

B. Perfil do investidor (conservador ou arrojado).

O primeiro perfil (conservador) baseará as suas decisões em um amplo intervalo de tempo. No caso de janela fixa, o número mínimo de dias de negociação considerados será de 400, correspondendo a cerca de 1 ano e 7 meses. No caso de janela móvel, a quantidade de dias considerados será igualmente 400.

O segundo perfil (arrojado), por sua vez, fundamentará as decisões em um intervalo mais curto. O número mínimo de dias considerados, no caso de janela fixa, será de 30, que é também o valor constante de dias utilizado no caso de janela móvel.

Os perfis diferem também quanto ao retorno esperado. Como apresentado no Capítulo 2, os modelos de otimização de portfólio geram uma “fronteira eficiente”, na qual todos os pontos correspondem às melhores combinações de risco e retorno, isto é, cada ponto da fronteira corresponde ao portfólio de menor risco para um dado retorno ou, inversamente, ao portfólio de maior retorno para um dado risco. No entanto, quando geramos esta fronteira, devemos escolher apenas um destes portfólios.

Esta escolha diferirá de acordo com o perfil do investidor. Escolhemos, no caso do perfil conservador, o portfólio cujo retorno esperado anual seja de 12% e, para o perfil arrojado, o aquele com 25% de retorno anual.

Salientamos que estes perfis foram escolhidos com base em valores reais praticados por analistas de investimento.

C. Informação utilizada (low, close ou high).

Os modelos clássicos de otimização de portfólios costumam utilizar o preço de fechamento do ativo para todos os cálculos (retorno diário, desempenho do ativo, etc.). Neste artigo, será testado qual o efeito de se utilizar o preço mais baixo (*low*) ou o mais alto (*high*) que o ativo atingiu em um determinado dia, no lugar do preço de fechamento.

3.2.2 Fatores estruturais

Os fatores estruturais, muito mais importantes de um ponto de vista conceitual do que os extrínsecos, serão o grande foco das análises.

D. Método de cálculo dos retornos esperados (média histórica ou rede neural).

Estudaremos duas opções quanto ao cálculo do retorno esperado de cada ativo ao longo do tempo. Primeiramente, utilizaremos a média dos retornos históricos, da maneira que apresentamos para o modelo de Markowitz tradicional no Capítulo 2.

Alternativamente, utilizaremos redes neurais não-lineares auto-regressivas (NAR) para efetuar este cálculo (conferir Capítulo 2). Para cada ativo, uma sucessão de redes neurais foi criada: a cada 42 dias de negociação (2 meses em dias corridos), criamos uma rede nova levando em consideração os dados antigos conjuntamente com os novos dados que foram obtidos nestes últimos 42 dias. Esta nova rede substitui a anterior como modelo de previsão dos retornos futuros de cada ativo.

A arquitetura escolhida está descrita abaixo:

Camada de entrada: como a rede é auto-regressiva, a entrada é a mesma que a saída. Queremos prever o preço do ativo no dia seguinte, então a entrada consistirá dos preços nos dias anteriores. Escolhemos basear a predição nos preços ponderados dos 5 dias de negociação precedentes. O peso dos 5 preços é crescente, de acordo com quão recentes eles sejam.

Conjunto de treinamento: 599 exemplos na primeira rede, ocorrendo atualização a cada 42 dias de negociação, que equivalem a 1 bimestre do ano. Como realizamos 3400 previsões, criamos um conjunto de 80 redes neurais para cada ativo, cada uma das quais com complexidade e precisão crescentes. Portanto, a última rede criada possuía um conjunto de dados com 3360 exemplos. Separamos sempre 70% dos dados para treinamento, 15% para validação e 15% para os testes finais.

Camada escondida: uma única camada, com 10 unidades. Percebemos que a precisão da rede não crescia com o aumento do número de unidades além de 10, nem com o acréscimo do número de camadas.

Camada de saída: vetor unidimensional, correspondente à previsão de preço para o dia em questão.

A Figura abaixo mostra a representação no MATLAB da rede neural criada:

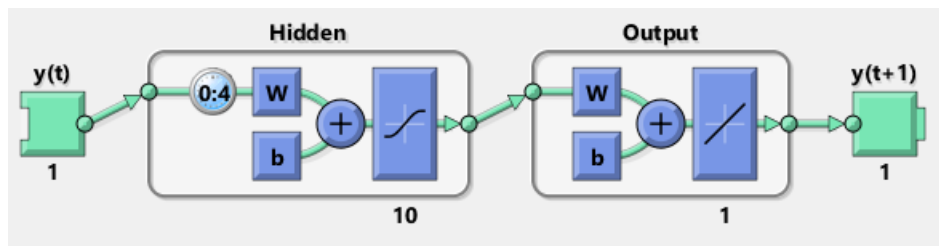


Figura 11 – Representação MATLAB da rede neural

Exemplificamos abaixo as previsões da rede neural para a ação KO (Coca-Cola). A linha azul corresponde aos preços reais; a linha vermelha, às previsões da rede.

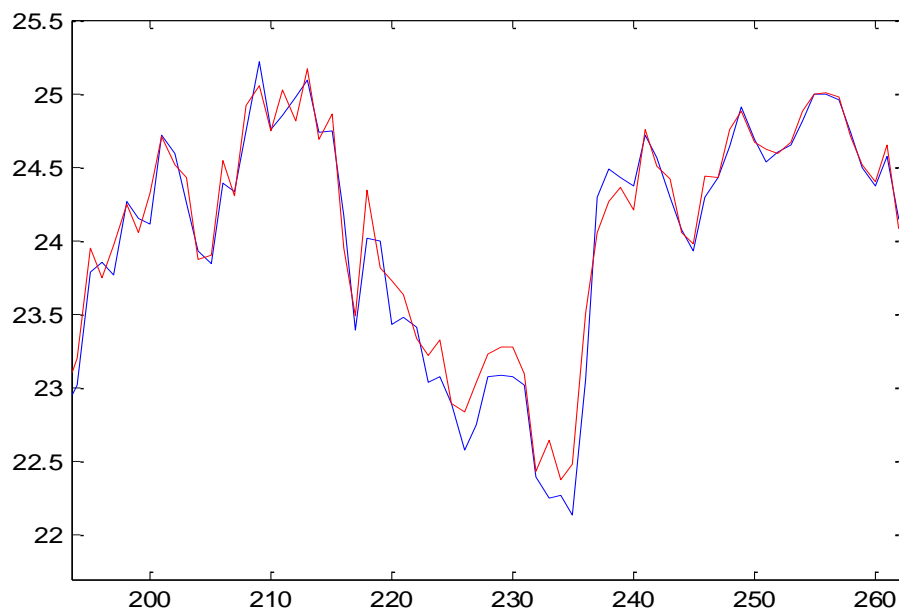


Figura 12 – Exemplo da capacidade de previsão das redes neurais

Como vemos, as redes são muito precisas. No entanto, constataremos posteriormente que outras características da rede são pouco interessantes do ponto de vista prático, tornando pouco eficiente a sua utilização.

E. Medida de risco utilizada (variância ou CVaR).

Conforme apresentamos no Capítulo 2, há diversas medidas de risco presentes na literatura acadêmica. Em especial, destacamos a variância, escolhida inicialmente por Markowitz (1952), e o CVaR, proposto por Rockafellar e Uryasev (2000).

A variância é, naturalmente, muito mais simples, em termos de cálculo e também de compreensão. O CVaR, por outro lado, requer a geração de cenários aleatórios, que rapidamente se tornam custosos computacionalmente. Neste trabalho, fizemos a escolha de gerar 10000 cenários aleatórios com 95% de nível de confiança, equivalentes a um programa de otimização com 10000 restrições.

F. Incorporação do risco (Markowitz ou Bayes).

O modelo de Markowitz, tal qual apresentado no Capítulo 2, considera que todos os parâmetros do modelo não possuam incertezas. Supõe-se, portanto, que o vetor de retornos $\vec{\mu}$ e a matriz de covariâncias Σ sejam ambos exatos.

No entanto, esta hipótese é considerada muito fraca por diversos autores (FABOZZI, F. J. *et al.*, 2007). Dentre as diversas abordagens propostas para contornar este problema, ganharam destaque o modelo de otimização robusta bayesiana e o modelo de Black-Litterman. Neste trabalho, contraporemos o modelo clássico de Markowitz à otimização bayesiana, de maneira a verificarmos o desempenho e a rugosidade de cada uma delas.

Utilizaremos o modelo proposto por Meucci (2005), com a variância como medida de risco, tal qual apresentamos no Capítulo 2.

3.3. Modelos estudados

Os seis fatores descritos acima geram um total de 72 combinações possíveis. Sendo impraticável estudar tal quantidade de combinações, restringimos o escopo deste trabalho a 48 delas.

Dois fatores – o tipo de janela e o perfil do investidor – serão tratados separadamente dos demais. Enquanto os outros fatores são categóricos em vez de numéricos, estes dois fatores se referem diretamente a variáveis numéricas: o retorno desejado e o tamanho do intervalo de aquisição de dados. Esta característica particular os torna passíveis de otimização, justificando esta separação. Desta maneira, temos quatro combinações possíveis para estes dois fatores:

1. Janela Fixa – Perfil Conservador
2. Janela Fixa – Perfil Arrojado
3. Janela Móvel – Perfil Conservador
4. Janela Móvel – Perfil Arrojado.

Quanto ao restante das escolhas, definiremos abaixo 12 combinações. Referir-nos-emos doravante a estas 12 combinações como “modelos”, conquanto a rigor elas não o sejam – afinal, os únicos “modelos” de fato estudados são Markowitz e Bayes. Este sutil abuso de linguagem simplificará a escrita e facilitará enormemente a compreensão.

Denominação do modelo	Informação utilizada	Cálculo do retorno	Medida de risco	Incorporação do risco no modelo
Markowitz (Close)	Close	Média	Variância	Não - Markowitz
Markowitz (Low)	Low	Média	Variância	Não - Markowitz
Markowitz (High)	High	Média	Variância	Não - Markowitz
Bayes (Close)	Close	Média	Variância	Sim - Bayes
Bayes (Low)	Low	Média	Variância	Sim - Bayes
Bayes (High)	High	Média	Variância	Sim - Bayes
CVaR (Close)	Close	Média	CVaR	Não - Markowitz
CVaR (Low)	Low	Média	CVaR	Não - Markowitz
CVaR (High)	High	Média	CVaR	Não - Markowitz
Neural (Close)	Close	Rede neural	Variância	Não - Markowitz
Neural (Low)	Low	Rede neural	Variância	Não - Markowitz
Neural (High)	High	Rede neural	Variância	Não - Markowitz

Tabela 1 – Os 12 modelos estudados

3.4. Detalhes Adicionais

O *software* utilizado é exclusivamente o MATLAB R2013b, em especial as caixas de ferramenta de Finanças e de Otimização.

O portfólio será composto pelas 10 ações da bolsa de Nova York (NYSE) abaixo:

- | | | |
|-------------------------|-------------------------|--------------------|
| - Coca-Cola (KO) | - Wal-Mart (WMT) | - ExxonMobil (XOM) |
| - Toyota Motors (TM) | - Procter & Gamble (PG) | - Pfizer (PFE) |
| - General Electric (GE) | - T&T (T) | - Citibank (C) |
| - J. P. Morgan (JPM) | | |

Foram escolhidos ativos “*blue chips*” (com alta capitalização de mercado). Além disso, diversas indústrias estão representadas, para diversificar o portfólio e, assim, diminuir as covariâncias entre os ativos.

Mostramos abaixo a interface de utilização desenvolvida:

The screenshot shows a MATLAB-style window titled 'parametros'. Inside, there's a main panel with a blue header 'MARKOWITZ'. Below this header, there are several sub-sections: 'Janela' with radio buttons for 'Móvel' (selected) and 'Fixa'; 'Perfil' with radio buttons for 'Conservador' (selected) and 'Arrojado'; 'Risco' with radio buttons for 'Variância' (selected) and 'CVaR' (with a value of 0.95); 'Critério de seleção' with radio buttons for 'Retorno' (selected, with a value of 0.12) and 'Sharpe' (with a value of 0.06); 'Data Inicial' and 'Data Final' with date pickers showing '01/07/2012' and '15/05/2014' respectively; 'Max Dias' and 'Potência' with input fields showing '70' and '1'; 'Tamanho da janela' with an input field showing '400'; 'Códigos das ações do portfólio' with a text box containing 'PG, TM, C, PFE, T, JPM, XOM, GE, WMT, KO' and a note 'Separe-os por vírgulas'; and 'Informação' with radio buttons for 'Close' (selected), 'High', and 'Low'. To the right of the 'MARKOWITZ' panel, there are two more panels: 'REDE NEURAL' with a checkbox for 'Rede Neural', a 'Período' input field showing '1000', and a 'Unidades' input field showing '10'; and 'BAYES' with a checkbox for 'Bayes'. At the bottom of the window, there are two buttons: a green one labeled 'Rodar otimizações' and a red one labeled 'Ir para o código'.

Figura 13 – Interface de utilização desenvolvida

Salientamos que, na interface, há diversas funcionalidades que foram desenvolvidas durante versões intermediárias do trabalho e abandonadas na versão final. Um total de 1300 linhas de código foram escritas.

3.5. Estudo de Rugosidade

3.5.1 Situação 1: 70 dias de negociação

A matriz das covariâncias, Σ , é um parâmetro importante para a determinação do portfólio ótimo. Verificaremos se os diferentes modelos se comportam diferentemente quando ocorrem pequenas alterações na matriz Σ . Em outras palavras, mediremos a robustez de Σ em relação aos seus dados de entrada, através da medida de rugosidade explanada anteriormente.

A Tabela 2 mostra o valor de rugosidade caso os 12 diferentes modelos fossem utilizados nos 70 dias de negociação que precedem o dia 15 de maio de 2014.

Janela e perfil					
Modelo	Móvel - Conservador	Móvel - Arrojado	Fixa - Conservador	Fixa - Arrojado	Média
Markowitz (Close)	2.55	9.62	1.25	14.91	7.08
Markowitz (Low)	2.51	6.76	1.14	14.97	6.34
Markowitz (High)	2.23	8.53	1.18	14.66	6.65
Bayes (Close)	3.42	9.63	1.43	12.08	6.64
Bayes (Low)	3.38	6.62	1.37	10.70	5.51
Bayes (High)	3.30	8.33	1.16	9.74	5.64
CVaR (Close)	11.80	10.44	10.35	18.05	12.66
CVaR (Low)	11.18	7.75	9.84	16.86	11.41
CVaR (High)	10.78	9.01	8.57	15.51	10.97
Neural (Close)	20.91	22.54	21.02	19.94	21.10
Neural (Low)	20.00	23.76	20.24	20.90	21.23
Neural (High)	20.85	22.84	20.74	21.50	21.48
Média	9.41	12.15	8.19	15.82	11.39

Tabela 2 – Estudo de rugosidade dos modelos (situação 1)

Conduziremos a seguir uma Análise de Variância (ANOVA, dois fatores, sem réplica, $\alpha = 5\%$) dos resultados acima, para verificar se podemos afirmar que as diferenças observadas são significativas:

Fonte da variação	Soma de Quadrados	Graus de Liberdade	Quadrado Médio	F	Nível descritivo	F crítico
Linhas	1805.5067	11	164.1370	21.6256	0.0000	2.0933
Colunas	412.2090	3	137.4030	18.1033	0.0000	2.8916
Residual	250.4680	33	7.5899			
Total	2468.1837	47				

Tabela 3 – Tabela ANOVA para o estudo de rugosidade dos modelos (situação 1)

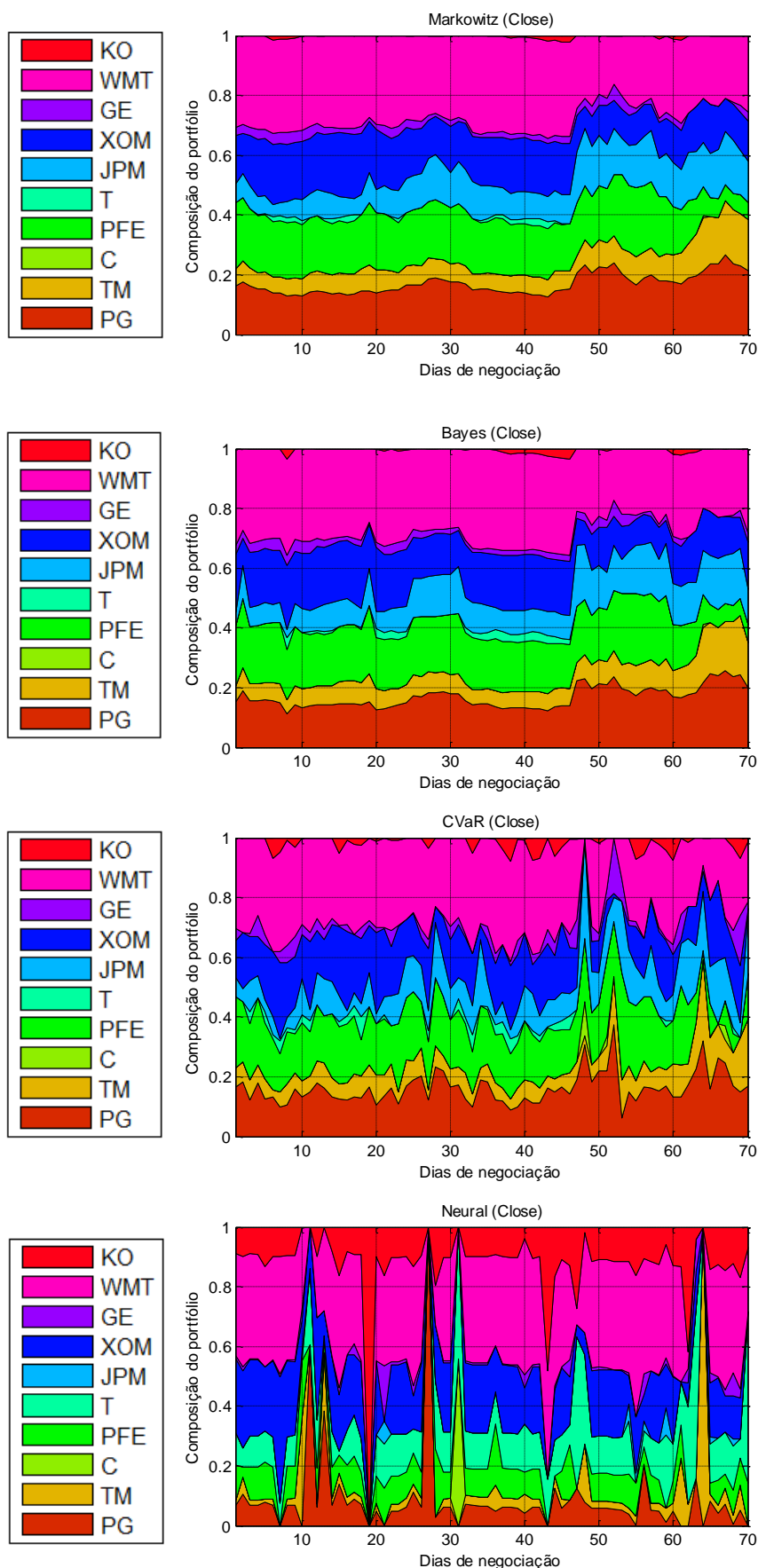
Os baixos valores de nível descritivo indicam que, com mais de 99,99% de significância, **existem diferenças de rugosidade** de acordo com tanto os fatores das linhas quanto com os das colunas. Repetimos o teste excluindo o CVaR e as redes neurais da análise e passamos a obter um nível descritivo de 58% para as linhas, mantendo o valor baixo para as colunas, indicando que eles eram os gerador da diferença entre as linhas.

Concluimos que:

- Os modelos com CVaR e redes neurais são significativamente mais rugosos do que Markowitz e Bayes.
- O perfil conservador, independentemente do tipo de janela usada, possui rugosidade menor do que o perfil arrojado.
- Otimização bayesiana de fato diminui a rugosidade dos portfólios, justificando a denominação “otimização robusta”. Nos modelos acima, há três pares que diferem apenas quanto à incorporação do risco. A saber:
 - Markowitz (Close) e Bayes (Close)
 - Markowitz (Low) e Bayes (Low)
 - Markowitz (High) e Bayes (High)

Em todos os pares, temos uma média de rugosidade menor para o modelo que usa otimização bayesiana, indicando que esta seja de fato mais robusta.

Ilustrativamente, providenciamos abaixo figuras que mostram as diferenças de rugosidade entre os quatro modelos básicos (Markowitz, Bayes, CVaR e Neural, usando close como informação, perfil conservador, janela móvel):



A composição segundo o modelo clássico de Markowitz é bem estável ao longo do tempo.

Neste intervalo específico, ela foi levemente mais estável do que com otimização bayesiana.

A otimização bayesiana, por ser robusta e incorporar as incertezas, é muito estável ao longo do tempo.

Como veremos a seguir, ela é, na média, um pouco mais estável do que o Markowitz clássico.

A otimização com o CVaR, além de demandar um tempo computacional extremamente longo, é muito pouco estável.

Podemos notar as “rugosidades” na figura ao lado.

A otimização através de redes neurais foi ainda mais rugosa do que o CVaR.

Veremos que esta é a desvantagem das redes, aos passos que o desempenho costuma ser muito bom.

Figura 14 – Variação da composição dos portfólios segundo o modelo

3.5.2 Situação 2: 3400 dias de negociação

Os resultados acima são relativos a um período de negociação de apenas 70 dias. Realizaremos a seguir um estudo para um período de 3400 dias de negociação, que se estende de 6 de novembro de 2000 a 15 de maio de 2014.

Nos nossos testes, os modelos que utilizam o CVaR tiveram, em média um tempo de cálculo 38 vezes maior do que o modelo clássico de Markowitz. Por comparação, os modelos com rede neural demoram 20% a mais do que Markowitz, ao passo que otimização bayesiana toma um tempo 80% maior. Dada a alta rugosidade e o altíssimo custo computacional, decidimos excluir o CVaR da análise de 3400 dias de negociação. Relembramos que o tempo de cálculo é decorrente da criação de 10000 cenários aleatórios, correspondentes em seguida à resolução de um problema de otimização com 10000 restrições.

A tabela abaixo mostra os resultados nesta segunda situação, para os outros nove modelos:

Janela e perfil					
Modelo	Móvel - Conservador	Móvel - Arrojado	Fixa - Conservador	Fixa - Arrojado	Média
Markowitz (Close)	208	618	100	73	250
Markowitz (Low)	191	559	82	71	226
Markowitz (High)	186	558	94	75	228
Bayes (Close)	191	581	83	58	228
Bayes (Low)	166	529	87	58	210
Bayes (High)	170	521	76	52	205
Neural (Close)	1082	1243	1041	1120	1121
Neural (Low)	1069	1275	1046	1117	1127
Neural (High)	1081	1300	1060	1128	1142
Média	483	798	408	417	526

Tabela 4 – Estudo de rugosidade dos modelos (situação 2)

Analogamente ao que fizemos anteriormente, procedemos com uma análise de variância (dois fatores, sem réplica). Obtemos níveis descritivos menores do que 10^{-10} para as linhas e para as colunas, indicando que ambas são influentes sobre a rugosidade dos modelos.

3.6. Estudo de Desempenho

3.6.1 Situação 1: 70 dias de negociação

Realizamos a seguir a análise de desempenho dos modelos. Numa primeira situação, faremos novamente observações correspondentes aos 70 dias de negociação que precedem 15 de maio de 2014. A Tabela 5 apresenta os resultados:

Janela e perfil					
Modelo	Móvel - Conservador	Móvel - Arrojado	Fixa - Conservador	Fixa - Arrojado	Média
Markowitz (Close)	2.48	3.39	2.79	-2.13	1.63
Markowitz (Low)	2.53	3.30	2.91	-0.84	1.97
Markowitz (High)	2.51	3.08	2.80	-1.85	1.64
Bayes (Close)	2.58	2.82	2.76	-0.66	1.87
Bayes (Low)	2.37	3.31	2.81	-0.57	1.98
Bayes (High)	2.58	2.82	2.76	-0.66	1.87
CVaR (Close)	3.15	3.51	2.62	-1.16	2.03
CVaR (Low)	2.54	3.06	2.54	-0.79	1.84
CVaR (High)	2.26	2.80	3.02	-1.99	1.52
Neural (Close)	4.05	4.46	4.01	4.55	4.27
Neural (Low)	4.09	4.06	4.09	3.85	4.02
Neural (High)	4.23	4.81	4.21	4.64	4.47
Média	2.95	3.45	3.11	0.20	2.43

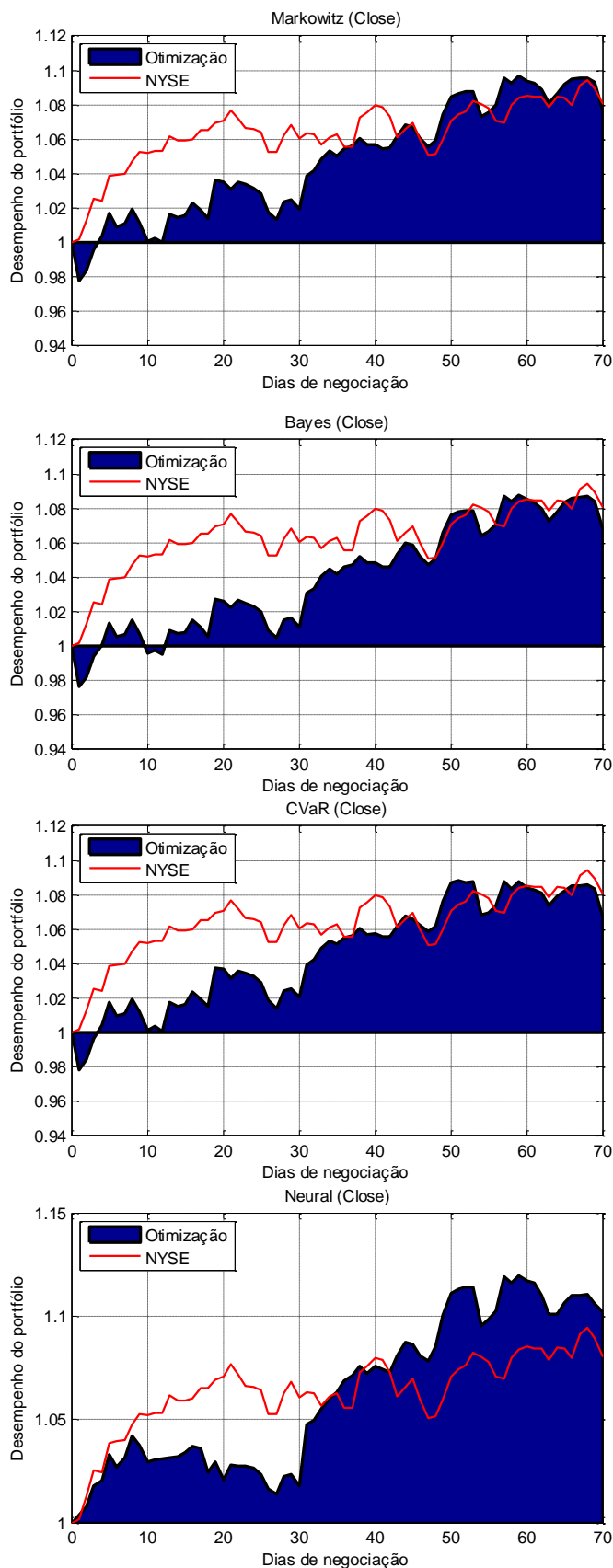
Tabela 5 – Estudo do desempenho dos modelos (situação 1)

Realizamos novamente uma análise de variância (dois fatores, sem réplica, $\alpha = 5\%$). Concluimos, novamente, que tanto as linhas quanto as colunas são influentes sobre o desempenho da otimização de portfólio.

Fonte da variação	Soma de Quadrados	Graus de Liberdade	Quadrado Médio	F	Nível descritivo	F crítico
Linhas	54.9374	11	4.9943	5.0745	0.0001	2.0933
Colunas	81.0408	3	27.0136	27.4474	0.0000	2.8916
Residual	32.4784	33	0.9842			
Total	168.4566	47				

Tabela 6 – Tabela ANOVA para o estudo do desempenho dos modelos (situação 1)

Ilustramos abaixo o desempenho dos métodos básicos (close como informação) ao longo do tempo, perfil arrojado e janela móvel.



Neste intervalo específico, Markowitz teve dificuldades de acompanhar o desempenho geral da bolsa.

No entanto, próximo do fim, notamos uma boa recuperação.

A otimização bayesiana teve desempenhos muito similares aos de Markowitz.

Percebemos algumas diferenças sutis, sobretudo do dia 45 em diante.

A otimização com o CVaR foi muito similar às duas acima.

Notamos sutis diferenças em relação ao Markowitz nos últimos 10 dias do período.

Similarmente ao Markowitz e ao Bayes, o modelo com Rede Neural teve alguma dificuldade no começo, porém ultrapassou o índice da bolsa no fim do período.

Figura 15 – Desempenho ao longo do tempo segundo os modelos

3.6.2 Situação 2: 3400 dias de negociação

Analogamente ao que fizemos com a análise de rugosidade, repetimos o estudo para o intervalo de 3400 dias de negociação (13 anos) que precede 15 de maio de 2014. Excluimos, novamente, os modelos que utilizam o CVaR como medida de risco, devido às razões já mencionadas. Obtemos os resultados a seguir:

Janela e perfil

Modelo	Móvel - Conservador	Móvel - Arrojado	Fixa - Conservador	Fixa - Arrojado	Média
Markowitz (Close)	982	-1680	-1934	-367	-750
Markowitz (Low)	356	-1627	-1408	-667	-836
Markowitz (High)	161	-1493	-1568	-704	-901
Bayes (Close)	516	-1512	-1122	-332	-613
Bayes (Low)	466	-1492	-1105	-197	-582
Bayes (High)	357	-1257	-1197	-318	-604
Neural (Close)	969	101	1116	743	732
Neural (Low)	547	-172	1022	682	520
Neural (High)	378	-111	682	412	340
Média	526	-1027	-613	-83	-299

Tabela 7 – Estudo do desempenho dos modelos (situação 2)

Ao repetir a análise de variância (dois fatores, sem réplica) chegamos exatamente às mesmas conclusões que na situação 1: tanto as linhas quanto as colunas são influentes sobre o desempenho da otimização de portfólio.

Notamos também que os desempenhos são quase todos negativos, exceto com a combinação Janela Móvel – Perfil Conservador e com a utilização de redes neurais.

3.7. Estudo da razão Desempenho/Rugosidade

Os resultados acima foram interessantes, porém inconclusivos. Analisar rugosidade e desempenho separadamente traz diversas limitações, já que as métricas devem ser consideradas conjuntamente quando de uma comparação holística entre modelos de otimização.

Tendo em vista este fato, propusemos uma nova métrica, que será a razão entre o desempenho e a rugosidade. Esta métrica, de fácil interpretação, traduz quanto de desempenho é trazido, em média, por cada variação na composição do portfólio.

A literatura acadêmica vem tratando deste balanço há bastante tempo. Skaf e Boyd (2009), entre outros autores, propuseram um modelo de otimização de portfólio considerando os custos de transação já na função objetivo. O próprio MATLAB possui em sua biblioteca um modelo desta natureza. A razão pela qual optamos por não considerar custos de transação é puramente computacional: considerar tais restrições torna o tempo necessário para as análises substancialmente maior, dificultando análises tão amplas quanto as que este trabalho se propôs a fazer. Alternativamente, optamos por estudar a razão desempenho/rugosidade. Esta solução nos permitiu realizar comparações entre um número considerável (48) de modelos e, simultaneamente, verificar o impacto das mudanças na composição do portfólio.

3.7.1 Situação 1: 70 dias de negociação

Os resultados da análise para a primeira situação se encontram resumidos na Tabela 8:

Janela e perfil					
Modelo	Móvel - Conservador	Móvel - Arrojado	Fixa - Conservador	Fixa - Arrojado	Média
Markowitz (Close)	0.97	0.35	2.24	-0.14	0.86
Markowitz (Low)	1.01	0.49	2.55	-0.06	1.00
Markowitz (High)	1.13	0.36	2.37	-0.13	0.93
Bayes (Close)	0.75	0.29	1.93	-0.05	0.73
Bayes (Low)	0.70	0.50	2.06	-0.05	0.80
Bayes (High)	0.78	0.34	2.37	-0.07	0.86
CVaR (Close)	0.27	0.34	0.25	-0.06	0.20
CVaR (Low)	0.23	0.39	0.26	-0.05	0.21
CVaR (High)	0.21	0.31	0.35	-0.13	0.19
Neural (Close)	0.19	0.20	0.19	0.23	0.20
Neural (Low)	0.20	0.17	0.20	0.18	0.19
Neural (High)	0.20	0.21	0.20	0.22	0.21
Média	0.55	0.33	1.25	-0.01	0.53

Tabela 8 – Razão desempenho/rugosidade (situação 1)

Notamos claramente que, quando as métricas são consideradas conjuntamente, Markowitz e Bayes têm resultados significativamente melhores do que CVaR e Neural.

3.7.2 Situação 2: 3400 dias de negociação

Este será o teste final, responsável por decidir qual modelo será utilizado para otimizar a matriz energética brasileira:

Janela e perfil

Modelo	Móvel - Conservador	Móvel - Arrojado	Fixa - Conservador	Fixa - Arrojado	Média
Markowitz (Close)	4.7	-2.7	-19.3	-5.1	-5.6
Markowitz (Low)	1.9	-2.9	-17.2	-9.4	-6.9
Markowitz (High)	0.9	-2.7	-16.6	-9.4	-7.0
Bayes (Close)	2.7	-2.6	-13.5	-5.7	-4.8
Bayes (Low)	2.8	-2.8	-12.8	-3.4	-4.0
Bayes (High)	2.1	-2.4	-15.7	-6.1	-5.5
Neural (Close)	0.9	0.1	1.1	0.7	0.7
Neural (Low)	0.5	-0.1	1.0	0.6	0.5
Neural (High)	0.3	-0.1	0.6	0.4	0.3
Média	1.9	-1.8	-10.3	-4.2	-3.6

Tabela 9 – Razão desempenho/rugosidade (situação 2)

Discutamos primeiramente as combinações de janela e perfil. Exceto para Janela Fixa e Perfil Conservador, todas as combinações têm médias negativas. Realizamos um teste ANOVA (dois fatores, sem réplica, $\alpha = 5\%$) para verificar se as diferenças são estatisticamente significativas, de acordo com a Tabela 10:

Fonte da variação	Soma de Quadrados	Graus de Liberdade	Quadrado Médio	F	Nível descritivo	F crítico
Linhas	327.5896	8	40.9487	2.3312	0.0520	2.3551
Colunas	700.7483	3	233.5828	13.2980	0.0000	3.0088
Residual	421.5664	24	17.5653			
Total	1449.9043	35				

Tabela 10 – Tabela ANOVA para desempenho/rugosidade (situação 1)

Concluímos que há diferenças entre as colunas. Portanto, na busca pelo melhor modelo de otimização, consideraremos exclusivamente a combinação Janela Móvel – Perfil Conservador, por ser a única de média positiva.

Em seguida, passemos à discussão dos modelos (linhas). Notamos que todos os que utilizam redes neurais, embora evitem desempenhos negativos para quase quaisquer combinações de janela e perfil, têm resultados bastante aquém do esperado. É notável que a média dos modelos com redes neurais é abaixo da dos modelos Markowitz e Bayes; consequentemente, desconsiderá-los-emos.

Resta-nos a difícil escolha entre o modelo clássico de Markowitz e o modelo de otimização robusta bayesiano. Se analisamos as médias, continuamos num impasse, pois elas são iguais a dois algarismos significativos.

Optaremos, nestas condições, pelo modelo com maior valor para a razão desempenho/rugosidade: Markowitz (Close). A razão é, no mínimo, 70% maior do que a de qualquer outro modelo. Interessantemente, este é o modelo tradicional desenvolvido em 1952 por Markowitz, sem nenhuma das melhorias que analisamos, que foram propostas pela literatura acadêmica ao longo de décadas. Segundo os resultados de nossos estudos, as propostas acadêmicas, embora capazes de melhorar pontos específicos do modelo, são incapazes de melhorar o balanço entre desempenho e rugosidade.

A nossa interpretação é que as redes neurais têm um poder de predição tão grande que as variações na composição no portfólio se tornam bruscas, causando instabilidades ao longo do tempo. Isto justifica o fato de os desempenhos serem muitas vezes excelentes, porém acompanhados de altíssima rugosidade.

No caso da otimização bayesiana, a interpretação é precisamente a oposta. Este modelo é extremamente robusto, com uma composição muito estável ao longo do tempo. Entretanto, esta robustez revela uma dificuldade de acompanhar as mudanças no mercado de capitais, de tal forma que, a longo prazo, o desempenho deixa fortemente a desejar.

Neste contexto, Markowitz mantém o seu papel como pilar unificador dos modelos, capaz de manter consistentemente um balanço quase perfeito entre qualidade de predição – traduzida em desempenho – e robustez do portfólio – traduzida em rugosidade.

Em conclusão, a seção seguinte deste trabalho focará em trabalhar para tornar ainda mais poderosas as características do modelo de otimização de portfólio com melhores resultados, de acordo com as nossas métricas.

3.8. Aprimoramento do modelo de Markowitz através de algoritmos evolutivos

O modelo de melhor desempenho do estudo anterior foi Markowitz (Close), combinado com Janela Móvel e Perfil Conservador. Nesta seção, aprimoraremos o desempenho desta combinação através de algoritmos evolutivos.

Como havíamos indicado no início deste capítulo, separamos os fatores Janela e Perfil porque eles eram numéricos e não categóricos. Este fato é importantíssimo, pois nos permitirá utilizar as poderosas técnicas conhecidas como algoritmos evolutivos para otimizá-los.

A Janela e o Perfil definem dois valores numéricos: o tamanho da janela e o retorno desejado (que, relembramos, é o critério de escolha do portfólio ótimo). Para perfil conservador, o tamanho da janela era de 400 dias e o retorno desejado era de 12% ao ano. Neste contexto, o nosso atual objetivo é encontrar os valores de tamanho de janela e de retorno desejado que maximizem o potencial de geração de valor do modelo. Em outras palavras, queremos encontrar a combinação de valores que irá maximizar a nossa medida de desempenho ao longo dos 13 anos de negociação que estamos estudando.

Para fazê-lo, separamos os dados em dois conjuntos, tal qual sugerido pela revisão bibliográfica (conferir Capítulo 2). Extraímos da bolsa de valores americanas dados relativos a um período total de 4000 dias de negociação e separamos cerca de 20% para rodar o algoritmo de otimização. Os dias restantes (3400) serão o conjunto onde verificaremos a qualidade da nossa otimização. Relembramos que essa separação do conjunto de dados é recomendada para evitar “*overfitting*”, isto é, que a nossa otimização seja excelente, porém pouco generalizável para conjuntos de dados diferentes daqueles nos quais realizamos a otimização.

Utilizamos o algoritmo evolutivo presente na Optimization Toolbox do MATLAB 2013b, configurando-o de acordo com os parâmetros abaixo:

Classificação	Parâmetro	Escolha
População	Tamanho	20
	Função de criação	Dependente das restrições
Seleção	Função de seleção	Uniforme estocástica
Reprodução	Elite count	1
	Fração de <i>crossover</i>	0.8
Mutação	Função de mutação	Dependente das restrições
Crossover	Função de <i>crossover</i>	Scattered
Migração	Direção	Forward
	Fração	0.2
	Intervalo	20
Restrições	Penalidade inicial	10
	Fator de penalidade	100
Fitness scaling	Scaling function	Rank
Critérios de parada	Gerações	60
	Limite de tempo	Infinito
	Limite de fitness	- Infinito
	Stall generations	50
	Stall time limit	Infinito
	Tolerância da função	1.00E-06
	Tolerância não-linear	1.00E-06

Tabela 11 – Parâmetros do algoritmo evolutivo (MATLAB).

Os parâmetros mais importantes são o tamanho da população – 20 indivíduos – e o número total de gerações – 60 –, que compõem o número total de indivíduos durante a otimização – 1200. Obtivemos, ao final destas 60 gerações, a combinação de 250 dias para a janela e 20% para o retorno. Com estes parâmetros, o desempenho da otimização passa a ser 4231.9150, contra 981.7624 obtidos com os parâmetros anteriores – uma melhora substancial de 331%. Curiosamente, 250 dias de negociação equivalem a, aproximadamente, 1 ano em dias corridos.

Para verificar a validade destes resultados, recalculamos o desempenho e a rugosidade de **todos** os modelos, de acordo com estes novos parâmetros. Sabemos que esta comparação não é justa, dado que os valores foram otimizados especificamente para Markowitz (Close), porém fá-la-emos não obstante este fato, para não fugir do escopo do trabalho. Obtemos a tabela abaixo:

Modelo	Desempenho	Rugosidade	Desempenho/Rugosidade
Markowitz (Close)	4232	291	14.6
Markowitz (Low)	2354	278	8.5
Markowitz (High)	2599	241	10.8
Bayes (Close)	2648	258	10.3
Bayes (Low)	1959	244	8.0
Bayes (High)	2048	233	8.8
Neural (Close)	989	1135	0.9
Neural (Low)	570	1122	0.5
Neural (High)	423	1128	0.4

Tabela 12 – Resultados da otimização para todos os modelos

Verificamos, apesar das melhoras enormes em todos os modelos, que o modelo Markowitz (Close) obteve, indiscutivelmente, os melhores resultados. Quanto à rugosidade, notamos novamente que os modelos bayesianos são levemente menos rugosos (mais robustos) do que o modelo clássico de Markowitz. Comparamos abaixo os desempenhos (d'_j), para Markowitz (Close), antes e após a otimização através de algoritmos evolutivos:

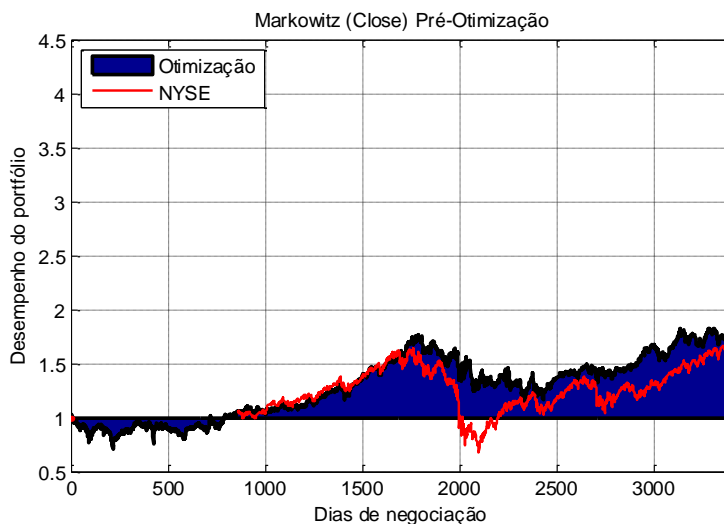


Figura 16 – Desempenho de Markowitz antes da otimização

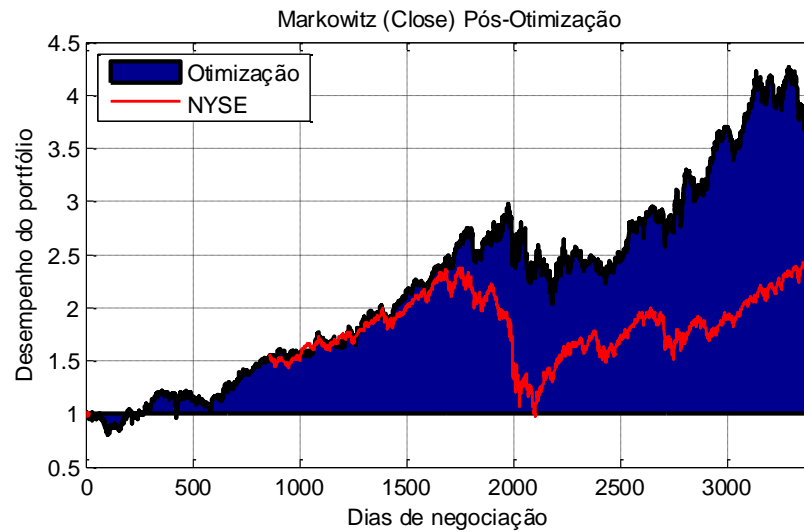


Figura 17 – Desempenho de Markowitz depois da otimização

Observação: o atual índice geral da NYSE (linha vermelha) foi criado em 2004, daí a sua ausência no começo do período de estudo.

Relembramos que a linha azul corresponde ao atual valor do portfólio, em múltiplos do valor inicial investido. Em particular, vemos que, ao término dos 3400 dias, o valor do portfólio foi multiplicado por 3.7: R\$10.000 investidos em 2000 ter-se-iam tornado R\$37.000 em 2014.

Consideramos esta melhora de desempenho de 331% como prova do poder de se utilizar algoritmos evolutivos no contexto de otimização de portfólio.

4. OTIMIZAÇÃO DA MATRIZ ENERGÉTICA BRASILEIRA

4.1. Mercado de energia no Brasil

Iniciamos este capítulo com uma breve explanação da dinâmica da geração de energia no Brasil. A matriz energética brasileira está fortemente concentrada em usinas hidrelétricas, com mais de 70% da energia proveniente desta fonte em 2013 (Balanço Energético Nacional, 2014):

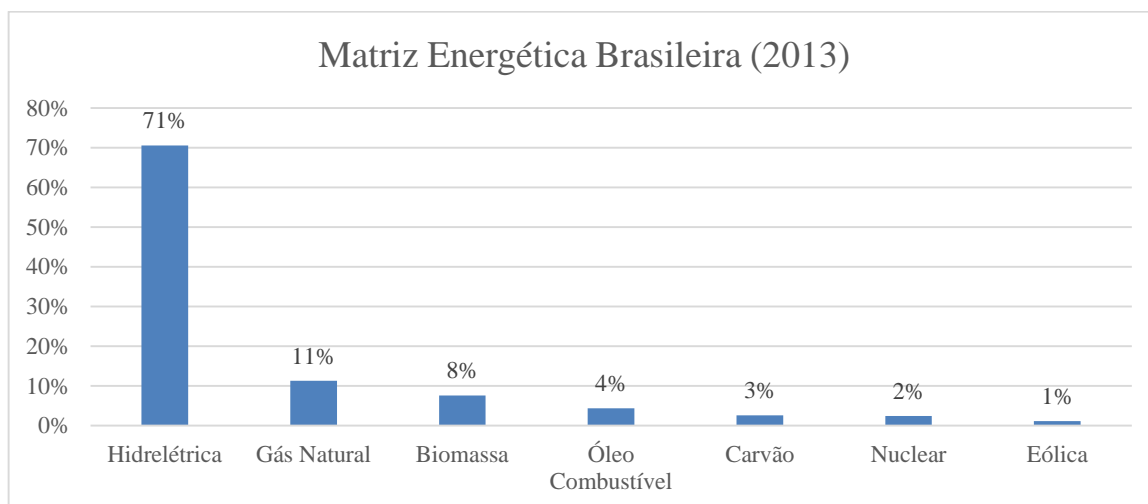


Figura 18 – Matriz energética brasileira (2013)

Notamos que a proporção de fontes renováveis (excetuando hidrelétrica) é baixíssima, com 8% advindo de usinas de incineração de biomassa e 1% de parques eólicos. Salientamos também que, diferentemente de países europeus, o Brasil nunca optou por fortalecer a produção por fontes nucleares, que ainda hoje é de apenas 2%. Quanto à energia térmica convencional fóssil, ela é considerada menos competitiva economicamente do que a hídrica ou as demais fontes renováveis (REGO, 2012), justificando a sua baixa contribuição.

O órgão governamental regulador deste mercado é a ANEEL (Agência Nacional de Energia Elétrica), que organiza leilões dos recursos naturais do país que possam ser usados como fontes de energia. Nestes leilões, a empresa que atender as especificações do edital e der o menor lance para o preço da energia comercializada, em R\$/MWh, receberá a concessão da geração de energia durante um período que costuma durar entre 30 e 40 anos.

Caso a empresa tenha mais capacidade de geração do que a que está sendo fornecida ao SIN (Sistema Elétrico Nacional), ela pode negociar a energia excedente no chamado mercado

“spot”. Este mercado é similar à bolsa de valores, no que tange à determinação de preços, que ocorre com base em oferta e procura. Empresas com alto consumo de energia elétrica (produtoras de alumínio ou aço, por exemplo) podem negociar a compra de energia elétrica no mercado spot, diretamente com as usinas geradoras, como alternativa a utilizar exclusivamente a energia elétrica proveniente do SIN.

4.2. Planejamento energético baseado em gestão de portfólios

Modelos de análise de portfólio são amplamente utilizados por investidores para alocar os recursos de maneira a garantir os retornos esperados, a despeito das condições econômicas (Roudier, 2007). Para fazê-lo, são contrabalanceados os retornos e os riscos de cada ativo, escolhendo-se a combinação que maximiza o retorno para um dado risco escolhido pelo investidor ou, alternativamente, a que minimiza o risco para um dado retorno. Neste capítulo, aplicaremos estes modelos à matriz energética do Brasil. Portanto, é necessário primeiramente interpretar de forma correta o que representam os parâmetros do modelo neste novo contexto.

Enquanto os modelos de teoria de portfólios maximizam o **retorno** – dado um nível de risco –, o interessante, do ponto de vista de geração de energia, seria minimizar o **custo**. Este fato é facilmente corrigido, pois há uma relação muito simples entre o retorno e o custo de uma fonte de energia. O retorno é a geração por unidade monetária investida (MWh/R\$), ao passo que o custo é o investimento por unidade de energia produzida (R\$/MWh). Portanto, para podermos utilizar o modelo clássico de maximização do retorno, basta interpretarmos o retorno como o inverso do custo e utilizarmos os dados relativos aos custos.

Conforme Awerbuch e Berger (2003), no caso de políticas energéticas, três pontos são fundamentais à matriz energética de um país: suprir a demanda, ter baixo custo e apresentar poucos riscos. Quanto à demanda, os atuais modelos de previsão são considerados suficientemente precisos, dado que a evolução do consumo é sempre gradual, sem saltos abruptos. Consequentemente, as políticas governamentais podem ser feitas com base nestas previsões e de acordo com um coeficiente de segurança a ser estabelecido. Além disto, países podem, em caso de necessidade, aumentar a sua capacidade energética temporariamente, através de importação ou de utilização de fontes intermitentes.

Quanto aos outros dois aspectos fundamentais – custo e risco –, a análise se torna substancialmente mais complexa. Como geração de energia envolve necessariamente a realização de escolhas entre tecnologias concorrentes, a comparação de custos deve ser precisa

e padronizada, de forma a mitigar distorções provenientes de cadeias de valor diferentes. Adicionalmente, como cada tecnologia têm custos de implementação e operação, eficiência, duração do ciclo de vida, pegada ecológica e fator de capacidade³ distintos, devem ser estabelecidas métricas comparativas abrangentes e multidimensionais.

Em relação aos riscos, estes se classificam essencialmente em duas classes: os riscos de disfunção do fornecimento e os de aumentos inesperados dos custos. Como vimos anteriormente, os riscos de disfunção do fornecimento podem ser reduzidos através do coeficiente de segurança em relação à demanda prevista e do uso eventual de fontes intermitentes ou mesmo de importação. No entanto, o segundo tipo de risco, calculado através da variância dos custos ao longo do tempo, requer uma análise profunda de diversos fatores, qualitativos e quantitativos, macro e microeconômicos.

Segundo Awerbuch e Spencer (2007), flutuações no preço do petróleo são exemplos de riscos relacionados a custos. Tais flutuações são capazes de reduzir a atividade econômica de todo um país, no caso de nações altamente dependentes de importação de petróleo. Até mesmo pequenos aumentos percentuais podem trazer perdas significativas, como desemprego e desaceleração da economia. Desta maneira, a redução deste tipo de risco é essencial para a escolha de uma boa matriz energética.

Ao aplicarmos a teoria de portfólios à produção de energia, determinamos portfólios que têm baixo custo e que, simultaneamente, minimizam a exposição de um país a flutuações de custos. Esta abordagem é completamente diferente da tradicional utilizada pelos governantes, que consiste em focar exclusivamente na minimização dos custos. A superioridade da abordagem de portfólios provém do fato de que é impossível prever qual será, daqui a 30 anos, a tecnologia mais barata. Consequentemente, focar exclusivamente em fontes de energia baratas

³ Fator de capacidade é a razão entre a energia de fato produzida durante um certo período e a energia que teria sido produzida se fosse possível operar uma usina produtora de energia em capacidade nominal total durante todo o período. Por exemplo, parques eólicos possuem fator de capacidade da ordem de 30%, dado que durante aproximadamente 70% do ano os ventos são insuficientes para produzir quantidades significativas de energia. Usinas nucleares têm fatores de capacidade que chegam a 90% (Electric Power Annual, 2009).

é análogo a investir durante 30 anos na ação que teve o melhor desempenho no dia em que tomamos a decisão (AWERBUCH, 2000).

O mesmo ponto é levantado por Corey (1981). A otimização da matriz energética de um país pressupõe a escolha de tecnologias robustas, de forma a diminuir a susceptibilidade à influência de fatores externos. Esta é uma característica inerente aos modelos de otimização de portfólio presentes na literatura acadêmica e analisados nos capítulos anteriores deste trabalho.

Em suma, propomos a escolha de um “portfólio eficiente” – em termos do balanço entre custos e riscos – no lugar dos típicos e simplistas “portfólios de baixo custo”. Para fazê-lo, utilizaremos o modelo tradicional de Markowitz, adaptado a este contexto:

- Fatores extrínsecos: Janela e perfil não têm sentido neste contexto. Utilizaremos todos os dados da base, porque ela contém apenas 19 datas. A escolha de um único portfólio da fronteira será realizada de uma maneira discutida posteriormente, e não simplisticamente selecionando um retorno alvo. A base BNEF possui três cenários (baixo custo, alto custo e intermediário). Utilizamos o intermediário, já que, nas análises do Capítulo 3 foi a utilização do preço Close que trouxe mais benefícios.
- Fatores estruturais: utilizaremos Markowitz tradicional, ou seja, calculado com a média dos retornos e a variância como risco e sem incorporar as incertezas.

Exemplificamos abaixo a fronteira eficiente (conferir Capítulo 2) calculada no contexto de matrizes energéticas. No eixo das ordenadas, temos o retorno em KWh/R\$, e, no eixo das abscissas, o risco, também em KWh/R\$.

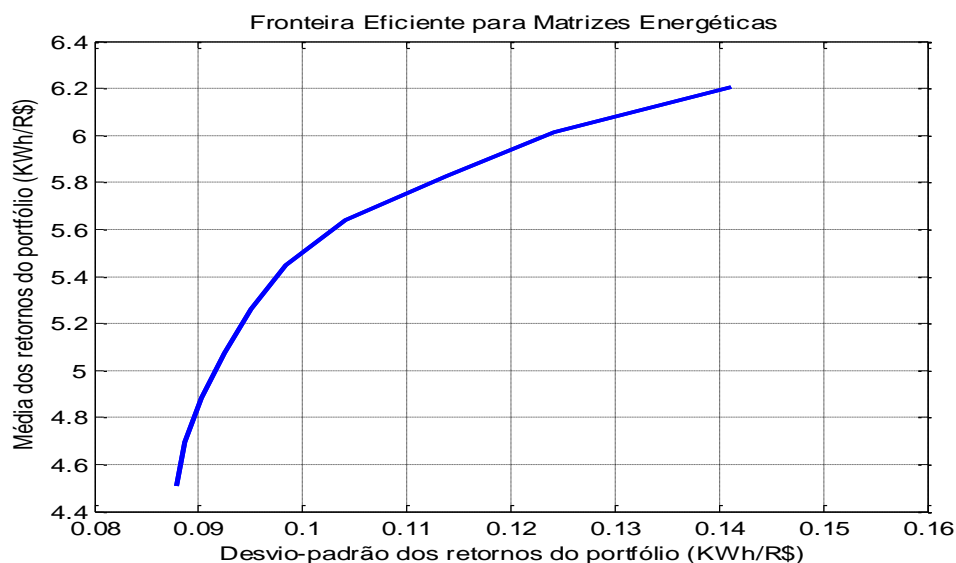


Figura 19 – Fronteira eficiente no contexto de geração de energia

4.3. Modelo de custos utilizado

Para obter os dados numéricos necessários aos algoritmos de otimização de portfólio, faremos uso de um modelo denominado LCOE (“custo nivelado da energia”, do inglês *levelized cost of energy*). Segundo Branker *et al.* (2011), o LCOE é uma métrica muito utilizada quando da comparação entre diferentes formas de geração de energia, correspondendo conceitualmente à razão entre o custo total de um empreendimento (descontado no tempo) e o total de energia elétrica que ele gerará (também descontado no tempo):

$$LCOE = \frac{\text{Custo total do empreendimento ao longo do ciclo de vida}}{\text{Produção total de energia ao longo do ciclo de vida}}$$

A seguinte fórmula é proposta para o cálculo exato do LCOE:

$$LCOE = \frac{\sum_{t=1}^n \frac{I_t + M_t + F_t}{(1+r)^t}}{\sum_{t=1}^n \frac{E_t}{(1+r)^t}}$$

Onde:

- I_t são os gastos com investimentos no ano t .
- M_t são os gastos com manutenção e operações no ano t .
- F_t são os gastos com combustíveis no ano t .
- E_t é a energia gerada no ano t .
- r é a taxa de desconto escolhida.
- t é tempo de vida do sistema.

O resultado do cálculo é geralmente expresso em dólares por megawatt-hora (\$/MWh). Conceitualmente, este custo é equivalente ao preço médio por MWh que teria de ser pago pelos consumidores para que o realizador do projeto obtivesse uma taxa de retorno igual à taxa de desconto escolhida.

De acordo com Branker *et al.* (2011), LCOE é frequentemente usado para contrapor o custo de geração de energia através de diferentes tecnologias. Em particular, ele pode ser utilizado para determinar qual a fonte de energia mais eficiente, do ponto de vista de recursos investidos. O modelo também pode ser empregado para comparar o custo da energia gerada por fontes

novas com o custo de fontes já existentes. Neste contexto, ele é extremamente útil em termos de decisões sobre a matriz energética de um país, para decidir como as demandas futuras serão atendidas e quais tecnologias devem ser apoiadas.

Ao somar todos os custos, o método contempla também as diferenças estruturais entre as tecnologias. Energias renováveis tipicamente têm um investimento inicial altíssimo, seguido de custos de combustíveis próximos de zero; plantas de gás natural, por outro lado, têm custos de combustíveis elevados, apesar do investimento inicial ser baixo. Estas e outras diferenças são todas levadas em consideração no cálculo.

Além de permitir a comparação entre diferentes tecnologias, o LCOE permite também a comparação de variações da mesma tecnologia. Empresas podem, por exemplo, estudar qual é a especificação de células fotovoltaicas que trará o maior retorno, considerando o ciclo de vida total do projeto, ou analisar quais as áreas onde pesquisas de redução de custo trarão mais retornos. O método permite até mesmo a comparação entre os custos de geração de energia com o valor efetivamente pago pelos usuários na conta de energia.

Naturalmente, há limitações para este método. Branker *et al.* (2011) ressalta que, a despeito das vantagens deste método, outras métricas também devem ser utilizadas para avaliar o projeto, como o *triple bottom line*. Segundo esta metodologia, além do retorno financeiro, são ponderados os impactos sociais e ambientais.

Além dos custos enunciados na fórmula acima, é possível incluir custos de emissão de gases de efeito estufa. Segundo Meneguín (2012), os custos de emissão de gases de efeito estufa podem ser particularmente significativos no caso dos países desenvolvidos. O Protocolo de Quioto estabeleceu três mecanismos de redução da emissão de gases estufa, conhecidos como Comércio de Emissões, Implementação Conjunta e o Mecanismo de Desenvolvimento Limpo (MDL). No caso de países em desenvolvimento, como o Brasil, apenas o MDL é atuante. Portanto, diferentemente dos países desenvolvidos, o Brasil não possui a meta de reduzir a emissão de gases estufa em pelo menos 5,2% em relação aos níveis de 1990 no período entre 2008 e 2012.

A despeito da ausência desta meta e da tributação que a acompanharia, o MDL ainda incentiva as instituições brasileiras a reduzir a emissão de gases estufa: o mecanismo concede às instituições o direito de receber créditos de carbono para comercializar. Além do gás carbônico, outros gases de efeito estufa podem ser “convertidos” numa quantidade de carbono

equivalente, de tal forma que a iniciativa é ampla quanto à redução da pegada ecológica das empresas.

Pessoas jurídicas podem propor à Comissão Executiva do MDL um projeto de redução de gases, independentemente da natureza jurídica da instituição (governos, ONGs, cooperativas, associações e empresas). Depois que um projeto filiado ao MDL entra em vigor, o Conselho Executivo do MDL emite, de tempos em tempos, a Redução Certificada de Emissões (RCE), documento eletrônico que especifica os créditos de carbono alcançados por cada projeto. Estas RCEs equivalem à redução de uma tonelada de gás carbônico. A bolsa de valores brasileira, BM&FBovespa, possui um ambiente eletrônico de negociação desenvolvido para viabilizar o fechamento de negócios envolvendo RCEs gerados por projetos de MDL.

Ainda de acordo com Meneguín (2012), o sucesso desta medida tem sido considerado relevante no Brasil: apenas a usina termelétrica a biogás instalada no Aterro Bandeirantes, localizado em Perus, permitiu à Prefeitura de São Paulo a comercialização de 1.262.793 RCEs até 2012. Desta maneira, a comercialização dos RCEs se tornou uma fonte de arrecadação importante para instituições como a Prefeitura de São Paulo. Notavelmente, em setembro de 2007, o banco holandês Fortis Bank NV/SA desembolsou num único lance R\$ 34 milhões pelo lote de 808.450 créditos de carbono colocados em negociação pela Prefeitura de São Paulo na Bolsa de Mercadorias e Futuros.

Em conclusão, no caso dos países desenvolvidos, os custos de emissão de gases de efeito estufa são diretos, dado que as empresas são taxadas caso a produção esteja acima do limite estabelecido pelas políticas governamentais locais. Diferentemente, as empresas localizadas em países em desenvolvimento não são taxadas. Em vez disto, elas podem usufruir da receita adicional gerada pela comercialização de RCEs e, naturalmente, interpretar esta receita como redução dos custos de geração. As empresas de países desenvolvidos que produzem menos do que o limite legal estabelecido também podem comercializar créditos de carbono similares às RCEs brasileiras. Desta maneira, é necessário considerar no modelo LCOE a influência financeira – positiva ou negativa – das políticas internacionais de redução de gases de efeito estufa.

4.4. Metodologia

Utilizaremos dados provenientes da BNEF (Bloomberg New Energy Finance) e da BM&FBovespa (a bolsa de valores brasileira).

A BNEF tem calculado e registrado, desde 2009, os custos nivelados de energia (LCOE) para 19 fontes, incluindo subclasses de uma mesma tecnologia. Como o modelo utilizado é rigorosamente o mesmo para todas as tecnologias, a padronização dos dados permite comparações consistentes entre tecnologias e entre datas. As outras bases de dados que encontramos são menos abrangentes ou menos pragmáticas.

Salientamos, no entanto, que os dados são uma média global. Infelizmente, não existem empresas brasileiras que compilem sistematicamente dados nacionais sobre geração de energia para fins de criação de uma bases de dados pública de LCOE como a Bloomberg. Tivemos, conseqüentemente, que recorrer às médias internacionais, uma abordagem que julgamos suficiente precisa, dado o escopo deste trabalho.

Esta metodologia difere da adotada por Losekann *et al.* (2013) em um dos artigos mais recentes focados em otimização da matriz energética brasileira do ponto de vista de teoria de portfólios. Para contornar o grave problema de escassez de dados sobre o Brasil, os autores recorreram a simulações de Monte Carlo. Para cada tecnologia, eles simularam os custos de acordo com um conjunto de parâmetros (eficiência, fator de capacidade, custos de investimento e operação, preço de emissão de gás carbônico, etc.). Esta abordagem, nas palavras dos autores, se trata de uma aproximação empírica: *“Since no one knows the exact value of these parameters, we postulate a probability distribution for each of them based on the information available (IEA, NEA, questions to experts in Brazil)”* (LOSEKANN *et al.*, 2013).

Comparamos os dados obtidos pelos autores deste artigo com os preços pagos pelo governo nos leilões brasileiros. Os preços simulados de geração a partir de usinas hidrelétricas grandes e pequenas são, respectivamente, 50,24 e 69,09 dólares por MWh. No entanto, entre 16 de dezembro de 2005 e 14 de dezembro de 2012, o preço médio pago pelo MWh, segundo os dados públicos da ANEEL foi de, respectivamente, 83 e 134 reais. A despeito da taxa de câmbio utilizada, vemos que a diferença entre as tecnologias segundo Losekann *et al.* é de 37,5%, enquanto que segundo a média das concessões é de 61,4%. Julgamos esta disparidade muito grande para um país onde 71% da geração de energia provém de usinas hidroelétricas (Balanço Energético Nacional, 2014).

Os dados da BNEF nos permitiram evitar outro problema encontrado por Losekann *et al.* Dada a escassez de dados, eles tiveram que limitar o estudo a apenas 8 fontes (gás, carvão, nuclear, óleo combustível, biomassa, eólica e hidrelétrica grandes e pequenas). A Bloomberg compila dados para um total de 19 fontes, incluindo fontes mais modernas e suas subclasses, como células fotovoltaicas à base de silício e de filmes finos, três tipos diferentes de biomassas e ondas do mar e marés⁴.

Finalmente, evitamos a abordagem de Losekann *et al.* porque os autores misturaram dados das simulações de Monte Carlo que eles próprios fizeram com dados do artigo de Awerbuch e Spencer de 2007. Conquanto eles tenham sido capazes de estimar os custos e os riscos de cada uma das 8 tecnologias, eles tiveram que recorrer aos dados publicados por Awerbuch e Spencer para obter valores para as covariâncias entre as variáveis. Julgamos que esta metodologia seja pouco precisa, dada que os dados e contextos de cada artigo são diferentes. Ao utilizarmos exclusivamente a base Bloomberg, asseguramos completa consistência entre os valores estimados para custos, riscos e covariâncias.

Como discutido anteriormente, o modelo LCOE calcula os custos de produção de acordo com a fonte. No entanto, como otimizaremos a matriz energética de todo um país, não são os custos incorridos pelas empresas que importam, mas sim o preço pago pelo governo por unidade de energia. Em outras palavras, o custo do modelo de otimização não é o mesmo custo do modelo LCOE, mas sim o preço de venda da energia.

A diferença entre o custo total dos produtos vendidos e receita total de vendas é a margem bruta das empresas. Das diferentes margens (bruta, operacional, líquida, EBIT/LAJIR, EBITDA/LAJIDA, etc.) de uma empresa, julgamos que a mais conveniente para a nossa aproximação seja a margem bruta, pois esta corresponde exatamente às receitas totais da empresa, descontado o custo dos produtos vendidos (CPV). Desta maneira, podemos utilizar a

⁴ Os mecanismos de geração de energia por ondas do mar e por marés são distintos, dado que as ondas do mar são originárias essencialmente do vento, enquanto que as marés advêm das forças gravitacionais entre a Terra, a Lua e o Sol.

margem bruta como aproximação da diferença entre o custo de produção da energia e o preço de venda. Como o custo estimado pelo modelo LCOE é por MWh, esta aproximação é razoável.

Para realizar esta adaptação, recorreremos a dados públicos da bolsa de valores BM&FBovespa e estimamos a margem média das empresas do setor elétrico. Apesar de as empresas atuarem em diferentes estágios da cadeia de valor (a saber, a cadeia é composta por geração, transmissão e distribuição) e nem todas serem públicas, consideramos que a margem média das empresas negociadas na bolsa seja suficientemente representativa da média das empresas do setor. As dez empresas mais líquidas da bolsa de valores do setor elétrico entre 2003 e 2013 foram AES Tietê (controladora da Eletropaulo), Cemig, Cesp, Copel, Light S/A, Tractebel, AES Elpa, Ampla, Celpe, Coelce, Elektro e Coelba. Consideramos que estas empresas sejam representativas do conjunto de empresas do setor elétrico, dado que as outras tiveram índices de liquidez extremamente baixos.

As empresas do setor elétrico costumam ter margens altas e, graças a este fato, terem um desempenho excelente na bolsa de valores. Notavelmente, a AES Tietê, a Coelce, a Cemig e a Tractebel foram, respectivamente, as 2^a, 3^a, 5^a e 7^a empresas que mais distribuíram dividendos entre 2003 e 2013, dentre todas as negociadas na bolsa (cálculo do autor com dados da BM&FBovespa).

As margens brutas em 2013 para as dez empresas supracitadas foram:

AES Tietê	Cemig	Cesp	Copel	Light S/A	Tractebel
66,7%	36,1%	64,7%	19,2%	26,1%	47,7%
AES Elpa	Ampla	Celpe	Coelce	Elektro	Coelba
4,8%	26,4%	12,9%	14,5%	17,2%	27,2%

Tabela 13 – Margens das empresas do setor elétrico mais líquidas

As empresas têm, portanto, uma média de margem de lucro bruta de 30,3%. Por simplicidade, assumiremos a hipótese de que a margem das empresas é em média constante, isto é, que as variações nos seus custos são sempre compensadas por variações proporcionais no preço de venda ao governo.

Em resumo, a nossa abordagem consistirá em converter os dados do modelo LCOE da BNEF para o mercado brasileiro através da informação sobre a margem bruta das empresas

obtida na BM&FBovespa e da taxa de câmbio vigente no período analisado. Em seguida, calcularemos os custos, riscos e covariâncias entre as fontes, que são as informações necessárias para a aplicação efetiva dos modelos de otimização de portfólio. Relembramos que o risco corresponde à variância dos custos.

4.5. Análise preliminar dos dados

Fontes de Energia

Apresentamos abaixo as 19 fontes catalogadas pela BNEF, em ordem decrescente de custo médio de geração de energia, com explicações para as mais modernas.

1. **Gás Natural e Petróleo.**
2. **Geotérmica:** utiliza-se como fonte o calor proveniente do interior da Terra (até 64 quilômetros de profundidade), como gêiseres e vulcões.
3. **Cogeração:** é um processo onde eletricidade e calor são gerados conjuntamente, através do aproveitamento de mais de 70% da energia térmica proveniente da queima dos combustíveis. Geralmente, o combustível utilizado é gás natural.
4. **Hidrelétrica.**
5. **Carvão.**
6. **Eólica em terra (*onshore*):** gerada através da instalação de parques eólicos que utilizam a energia do vento.
7. **Geotérmica binária:** similar à geotérmica normal, porém permite a utilização de fontes de temperatura moderada.
8. **Resíduos sólidos:** gerada a partir de dejetos urbanos, através da biomassa contida.
9. **Nuclear.**
10. **Biomassa por incineração:** gerada a partir de biomassa, como lenha, bagaços de cana, papel e palha de arroz.
11. **Biomassa anaeróbica:** o processo de decomposição da biomassa é feito por bactérias, que, ao decompor o material, produzem biogás (metano e dióxido de carbono).
12. **Biomassa por gaseificação:** a conversão do combustível sólido em gás ocorre por meio de reações termoquímicas.
13. **Fotovoltaica filme fino:** transformam a energia solar em energia elétrica através do efeito fotovoltaico. A primeira geração de células fotovoltaicas era constituída por

silício, porém as novas gerações utilizam um filme fino que reduz os custos de manufatura e manutenção.

14. **Fotovoltaica móvel com silício:** é a primeira geração de células fotovoltaicas, com um rendimento maior graças à mobilidade, permitindo direcionar às células em direção aos raios de sol.
15. **Eólica no mar (*offshore*).**
16. **Fotovoltaica fixa com silício:** a primeira geração de células fotovoltaicas, sem o mecanismo de direcionamento das células.
17. **Refletor linear:** um conjunto de espelhos que focam a luz do sol em um ponto, de tal forma a deixá-la até 30 vezes mais intensa. Neste ponto, um líquido térmico (capaz de manter estado líquido mesmo a altas temperaturas) é aquecido e, em seguida, utilizado para a vaporização de água.
18. **Marés:** é obtida por meio do aproveitamento da energia proveniente do desnível das marés. Para que essa energia seja revertida em eletricidade é necessária a construção de barragens, eclusas e unidades geradoras de energia. O sistema utilizado é semelhante ao de uma usina hidrelétrica.
19. **Ondas do mar:** transformação da energia das ondas em energia elétrica.

Custos

O gráfico a seguir mostra os retornos por fonte, em R\$/KWh, segundo a base de dados da BNEF. Por consistência, denominamos os retornos de LROE (retornos nivelados de energia, ou *levelized return of energy*). Relembramos que os retornos são os inversos dos custos.

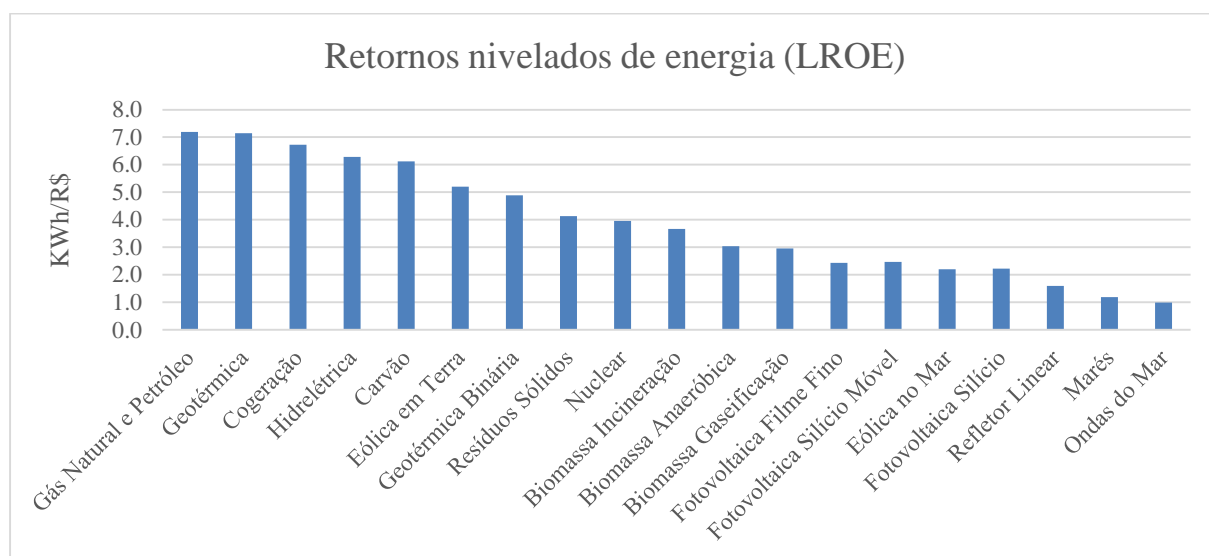


Figura 20 – Retornos nivelados de energia (LROE) de acordo com a BNEF

O retorno da geração por hidrelétricas é de 6,28KWh/R\$, correspondendo a um custo de geração de R\$159/MWh, sem distinção entre hidrelétricas grandes e pequenas. Apesar de este valor estar acima da média paga pelo governo pela energia advinda de hidrelétricas, no leilão de 30 de julho de 2010, o preço pago pela ANEEL pela energia das usinas hidrelétricas de Pirapora, Canaã, Jamari e Santa Cruz do Montenegro (todas na região Sudeste) foi de, em média, R\$154. Após esta validação da fonte mais importante da matriz brasileira, consideramos que a faixa de valores que obtivemos é plausível.

Notamos que as tecnologias mais tradicionais são consideravelmente mais baratas do que as mais modernas. Com exceção de eólica em terra e geotérmica, que têm um ótimo retorno, as outras tecnologias inovadoras ainda possuem em geral custos bastante elevados.

Riscos

O gráfico abaixo compara os riscos de produção, por fonte, segundo a base de dados da BNEF. O risco é medido nos modelos de otimização através da variância dos retornos, porém mostraremos na figura abaixo o desvio-padrão, pois este é mais fácil de interpretar, por ser expresso na mesma unidade que o retorno (KWh/R\$).

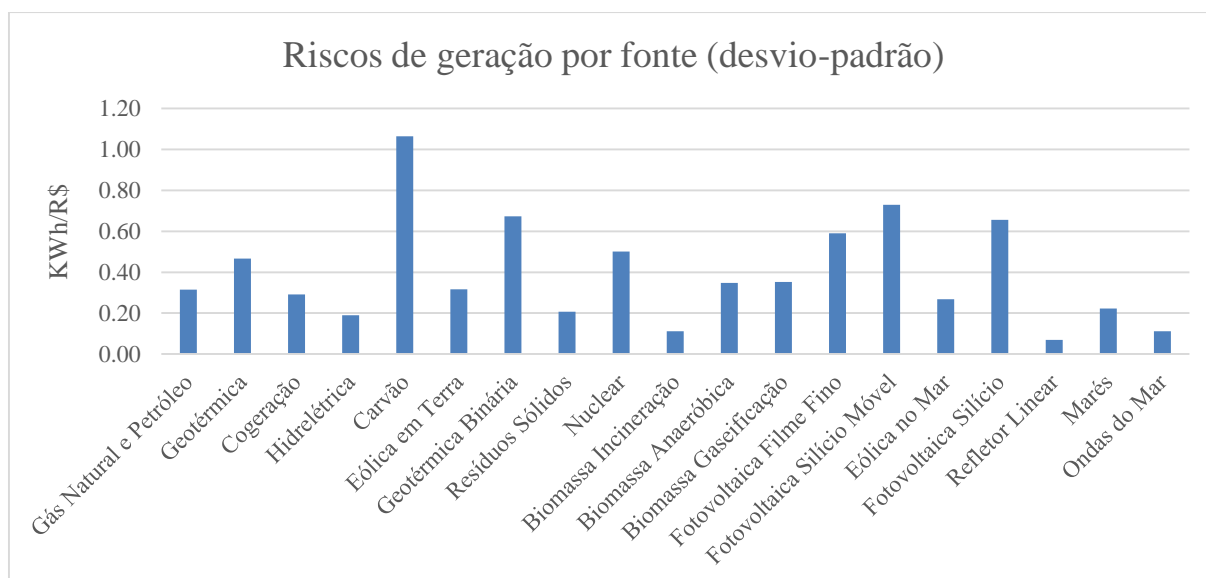


Figura 21 – Riscos das fontes de energia de acordo com a BNEF

Quanto aos custos, não há nenhum padrão claro que diferencie as tecnologias. Notamos, no entanto, que hidrelétricas possuem, simultaneamente, baixo custo e baixo risco. Esta combinação de características, extremamente benéfica para o Brasil, fará com que esta

tecnologia seja considerada um ativo muito interessante, do ponto de vista de teoria de portfólios.

Covariâncias

A matriz das covariâncias se encontra no Apêndice, devido ao seu tamanho (19x19).

4.6. Resultados

A figura abaixo contrapõe a fronteira eficiente da matriz energética brasileira, calculada de acordo com o modelo clássico de Markowitz, e quatro portfólios importantes para o Brasil: o atual portfólio de geração e os portfólios A, B e C, que serão detalhados a seguir.

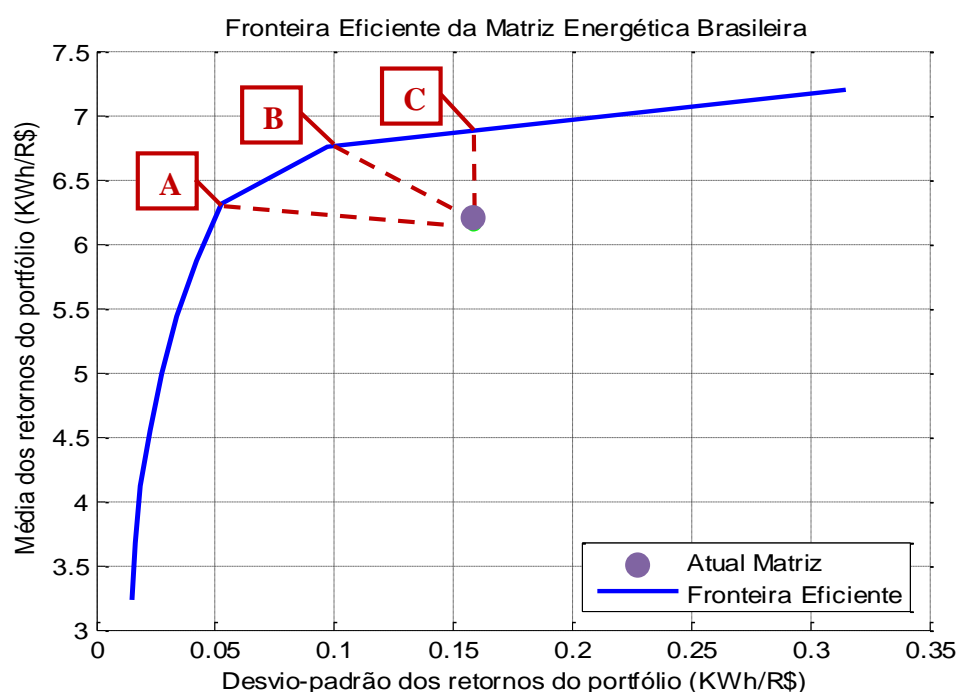


Figura 22 – Fronteira eficiente da matriz energética brasileira

Os três portfólios destacados na figura, chamados de A, B e C, são essenciais para a otimização do portfólio brasileiro:

- A. A derivada da fronteira eficiente visivelmente sofre grandes alterações nos pontos A e B. Em particular, o ponto A corresponde a um retorno de 6,314 KWh/R\$ e a um risco (desvio-padrão) de 0,053 KWh/R\$. A partir de A, o ganho de retorno passa a ser pequeno em comparação com o aumento de risco. Como o retorno máximo do portfólio é de 7,193 KWh/R\$ e o risco máximo é 0,315 KWh/R\$, o ponto A tem

a interessante propriedade de equivaler a 88% do retorno máximo com apenas 17% do risco máximo.

Além disso, o portfólio A está praticamente na mesma abscissa que a atual matriz energética brasileira, sendo portanto próximo do portfólio de menor risco mantendo o atual retorno.

- B.** O portfólio B, por sua vez, equivale a um segundo ponto onde há uma grande alteração na inclinação da fronteira eficiente. Ele seria, portanto, o “último” portfólio (na direção de risco crescente) que ainda possui um balanço interessante entre risco e retorno interessante. A curva que interliga os portfólios A e B é, segundo a nossa análise, a região mais interessante da fronteira.
- C.** O portfólio C se encontra na mesma ordenada que a atual matriz brasileira e, conseqüentemente, equivale ao portfólio com maior retorno se o risco for mantido constante.

A tabela abaixo explicita a composição dos portfólios A, B e C:

Tecnologia	% em A	% em B	% em C
Gás e Petróleo	3,3%	24,1%	35,0%
Geotérmica	10,1%	13,4%	19,3%
Cogeração	32,8%	30,1%	27,0%
Hidrelétrica	48,1%	32,4%	18,7%
Fotovoltaica fixa à base de silício	5,8%		

Tabela 14 – Composição dos portfólios A, B e C

Notamos claramente que as tecnologias renováveis modernas, exceto uma, estão ausentes nos portfólios. A razão é simples e já foi discutida anteriormente: elas possuem simultaneamente altos custos e altos riscos. A exceção é a célula fotovoltaica fixa à base de silício, que tem a interessante propriedade de possuir covariância de praticamente 0 com cogeração e com hidrelétrica e covariância bastante negativa com geotérmica. Como cogeração, hidrelétrica e geotérmica compõem 91% do portfólio A, a tecnologia fotovoltaica fixa com silício atua como um “contraponto” para o risco do portfólio.

Como era de se esperar, dados os baixos custos e baixos riscos, as tecnologias mais presentes são gás e petróleo, geotérmica, cogeração e hidrelétrica. Como explanado anteriormente, a energia hidrelétrica é muito atraente do ponto de vista de teoria de portfólios e, conseqüentemente, é a componente majoritária dos portfólios A e B.

Apesar de interessantes, os resultados acima são pouco realistas: não é plausível propor mudanças radicais no portfólio energético de um país, devido ao volume de investimentos necessários para alterar a matriz energética. Levando este fato em consideração, recalculamos a fronteira eficiente, adicionando restrições em relação à composição do portfólio, de maneira a tornar mais factíveis as mudanças propostas. As restrições adicionadas foram:

- Gás Natural e Petróleo: mínimo de 3% e máximo de 15%. O mínimo decorre da plausibilidade do modelo e o máximo para evitar que o país seja excessivamente dependente de petróleo. Conforme Awerbuch e Spencer (2007), variações no preço do petróleo estão entre as maiores fontes de risco para portfólios de energia.
- Geotérmica e cogeração: máximo de 3%, por plausibilidade.
- Hidrelétrica: mínimo de 65% e máximo de 85%, por questões de factibilidade, de dificuldade de alteração da matriz energética e para garantir o uso de outras fontes.
- Carvão: máximo de 20%.
- Biomassa: entre 3% e 20%, para manter o uso da energia renovável.
- Nuclear: entre 1% e 20%, por questão de compatibilidade com a atual matriz.
- Geotérmica binária: máximo de 5%.
- Eólica em terra: mínimo de 3% e máximo de 10%, devido às boas propriedades e para incentivar o uso de energias renováveis, porém mantendo a coerência com a atual matriz.
- Fotovoltaica fixa com silício: mínimo de 3%, pela boa combinação risco-retorno e para aumentar o uso de fontes modernas e renováveis.
- Demais energias: máximo de 10%.

A figura a seguir contrapõe a nova fronteira eficiente e a atual matriz:

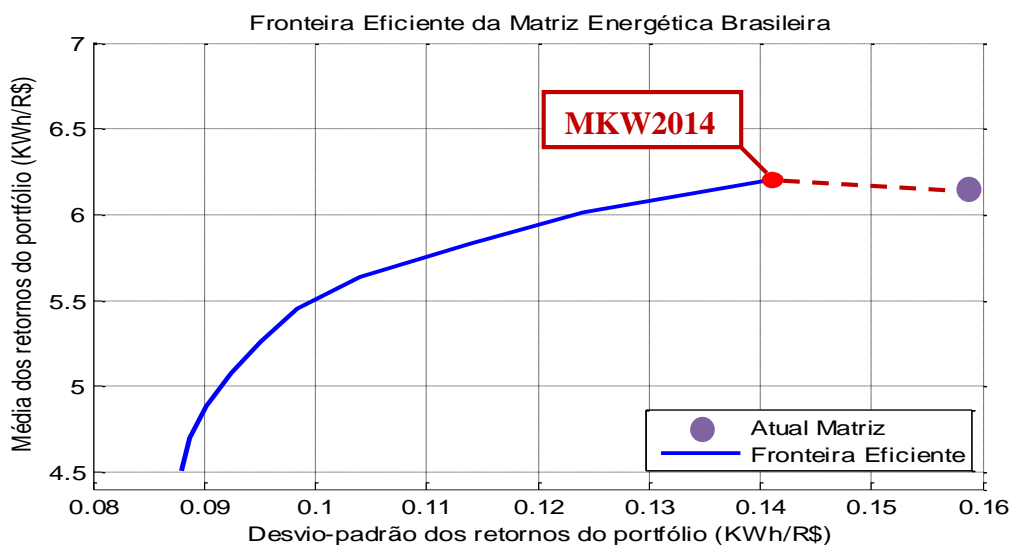


Figura 23 – Fronteira eficiente com restrições de factibilidade

Notamos que o comprimento da fronteira eficiente foi reduzido após a introdução das restrições. Este fenômeno é esperado, conjuntamente com o deslocamento da fronteira, que foi levada na direção sudeste, com retornos menores e riscos maiores. Tal qual aconteceu com a fronteira sem restrições, também focaremos a atenção sobre um portfólio específico da fronteira. Este portfólio, denominado de MKW2014 por ter sido obtido através do modelo de Markowitz com dados até 2014, é o mais próximo da atual matriz brasileira e possui ótimas propriedades de risco e retorno. Para que possamos comparar o portfólio MKW2014 com a fronteira eficiente irrestrita, apresentamos o gráfico abaixo.

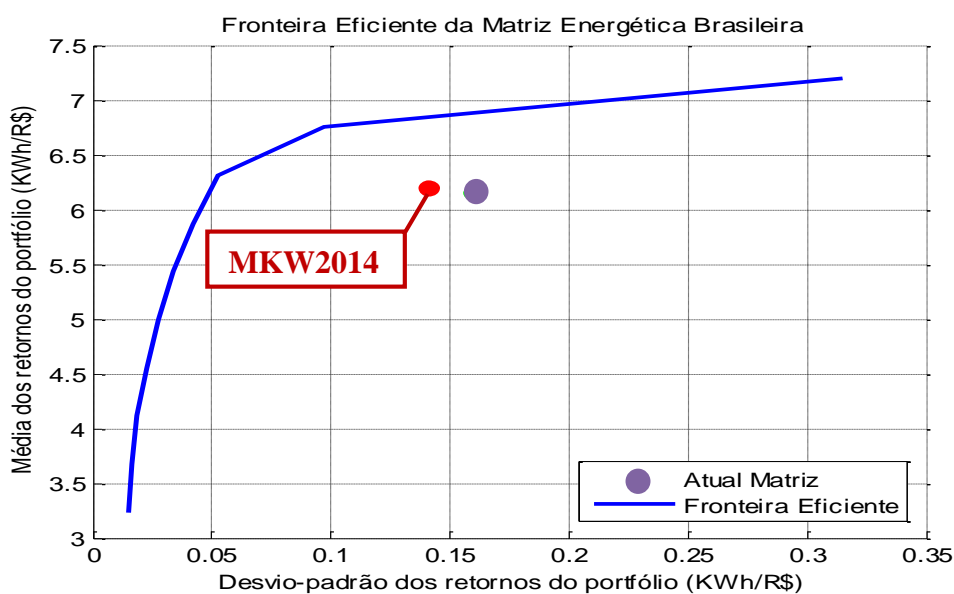


Figura 24 – Fronteira eficiente irrestrita e o portfólio MKW2014

Notamos que, apesar da inclusão das restrições, o portfólio MKW2014 não se afastou tanto da fronteira. Em particular, mover a matriz em direção ao portfólio MKW2014 equivale a movê-la na direção da curva A-B da fronteira irrestrita, que, como vimos anteriormente, tem ótimas combinações risco-retorno. Portanto, a recomendação final deste trabalho é que o Brasil tente aproximar a matriz energética atual tanto quanto possível deste portfólio, explicitado abaixo:

Tecnologia	% em MKW2014
Gás e Petróleo	15,0%
Geotérmica	3,0%
Cogeração	3,0%
Hidrelétrica	69,0%
Eólica em terra	3,0%
Nuclear	1,0%
Biomassa incineração	3,0%
Fotovoltaica móvel com silício	3,0%

Tabela 15 – Composição do portfólio MKW2014

Este portfólio possui um retorno de 6,204 KWh/R\$ e um risco (desvio-padrão) de 0,141 KWh/R\$. Em comparação com a atual matriz brasileira, com retorno de 6,13 KWh/R\$ e risco de 0,165 KWh/R\$, a melhora é notável, especialmente quanto ao risco, que caiu 14,5%.

4.7. Análise de sensibilidade à energia hidroelétrica

Realizaremos a seguir um estudo de quão sensível a fronteira eficiente é em relação à restrição de proporção mínima de energia hidrelétrica no portfólio. Como sabemos, a matriz brasileira é altamente dependente de recursos hídricos, com 71% da energia brasileira provindo de centrais hidrelétricas. Esta é, naturalmente, uma enorme fonte de riscos, devido simultaneamente à concentração do portfólio e à alocação majoritária em uma componente que não é a que tem a melhor relação risco-retorno.

A perspectiva de portfólios não é a única que evidencia os perigos de uma proporção tão alta de energia hidroelétrica. Com a escassez de chuvas no país, a chamada “crise hídrica” atualmente assola o país e em especial a região Sudeste. Desde 12 de junho de 2014, é o volume morto do sistema Cantareira que abastece as casas de cerca de 9 milhões de pessoas na capital

e região metropolitana de São Paulo. Com mais de 70% da energia do país dependente de água, os benefícios da diversificação da matriz energética do país são evidentes também sob as perspectivas geográfica e geológica (Balanço Energético Nacional, 2014).

Para testar qual a variação do risco e do retorno em função da principal fonte de energia do Brasil, testamos o modelo com cinco restrições diferentes. Para manter a plausibilidade do modelo, começamos com restrição de mínimo de 55% de energia hidrelétrica no Brasil e avançamos a passos de 5% até 75%.

Abaixo encontram-se as cinco fronteiras eficientes, em ordem crescente (da esquerda para a direita) de proporção mínima de energia hidrelétrica.

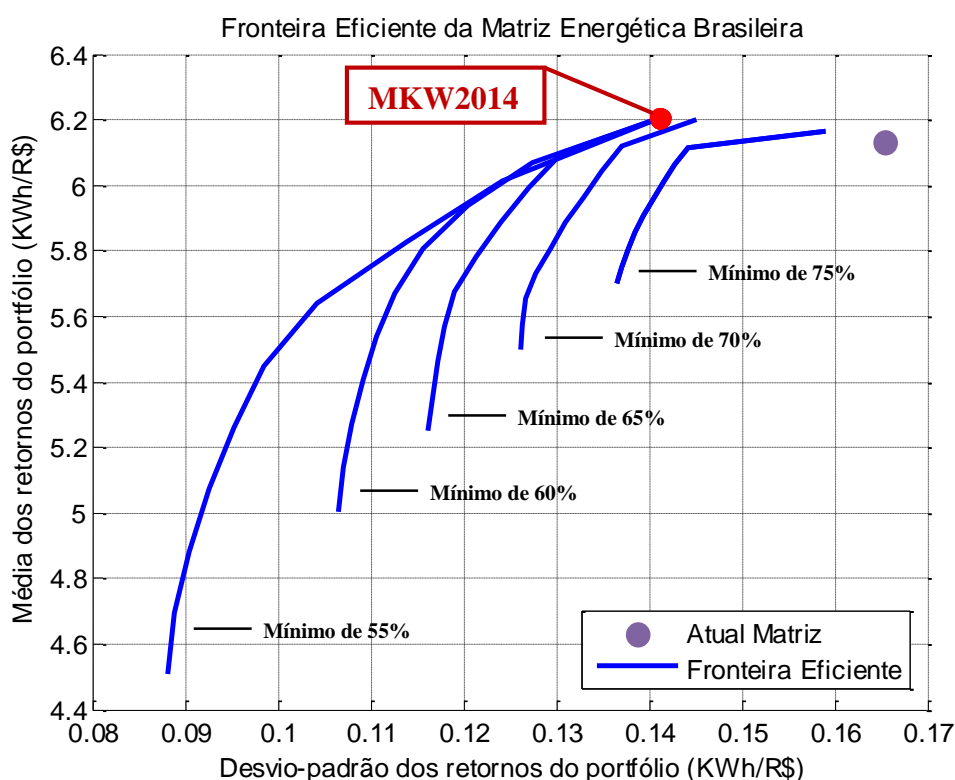


Figura 25 – Sensibilidade da fronteira eficiente à energia hidrelétrica

Torna-se evidente, a partir da figura, que a fronteira eficiente do Brasil é extremamente sensível à proporção mínima de energia hidrelétrica. As fronteiras se deslocam no sentido de menor risco ao passo que a restrição é relaxada, indicando que grande parte do risco da atual matriz provém da excessiva proporção de energia hidrelétrica.

Alternativamente, podemos interpretar os deslocamentos das matrizes do ponto de vista dos retornos. Não obstante a alteração da proporção mínima, o retorno máximo da fronteira (ponto mais à direita de cada fronteira) se mantém praticamente inalterado. Este fato indica que, concomitantemente ao alto risco decorrente da concentração do portfólio, o Brasil não recebe nenhum acréscimo minimamente significativo do ponto de vista do retorno.

4.8. Comparação com outros resultados

Por fim, realizamos um breve estudo comparativo entre os resultados deste trabalho e os de estudos similares, dos quais destacamos três:

- PDE2020 (Ministério de Minas e Energia, 2012): A Secretaria de Planejamento e Desenvolvimento Energético do Ministério de Minas e Energia publica anualmente um estudo chamado de Plano Decenal de Expansão de Energia (PDE2020). O relatório leva em consideração projeções de diversos indicadores sociais, técnicos e macroeconômicos, como PIB, população, e oferta e demanda de energia, e estima qual será a matriz energética brasileira em 2020.
- PSW2020 (WWF-Brazil, 2012): O WWF (World Wildlife Fund) realizou um estudo denominado de Power Switch 2020 (PSW2020), focado majoritariamente em uma matriz energética verde e sustentável proposta para o ano de 2020.
- BAU2020: No mesmo relatório em que o WWF propõe um cenário sustentável, um cenário menos consciente ecologicamente é proposto, chamado de *business-as-usual* (BAU2020). Este cenário é baseado em projeções acerca da situação macroeconômica e energética do país em 2020, caso poucas políticas verdes sejam adotadas.

A figura a seguir localiza no plano retorno-risco os três portfólios acima, conjuntamente com a fronteira eficiente irrestrita e o portfólio MKW2014 que propusemos.

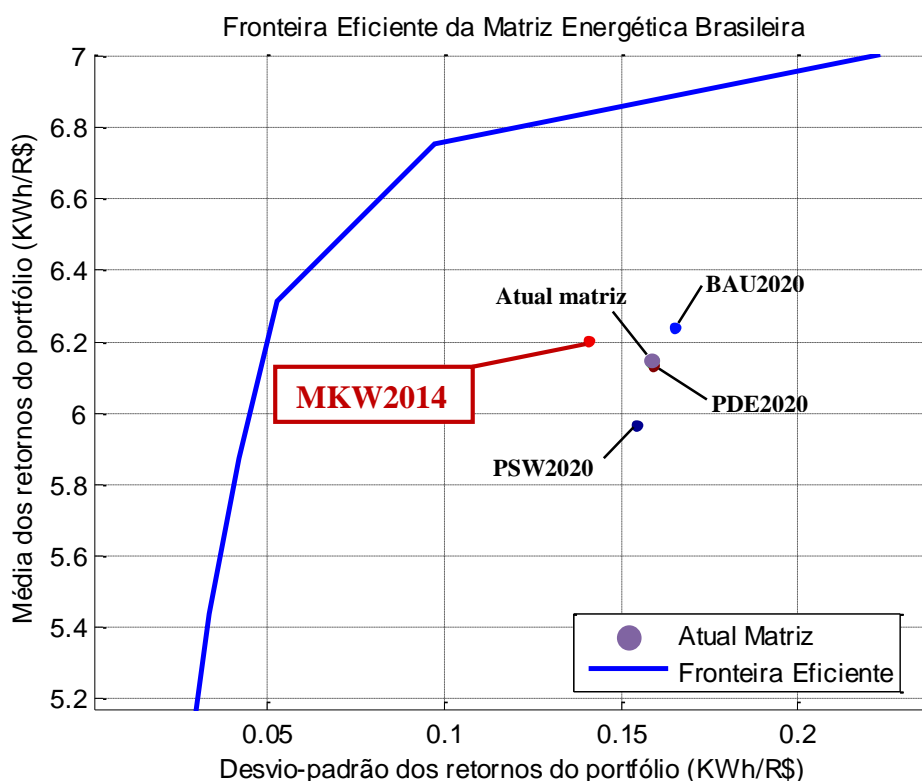


Figura 26 – Comparação com projeções de outros estudos

O primeiro ponto importante desta figura é que, de acordo com o modelo de Markowitz e os dados da BNEF, o portfólio MKW2014 é o único que melhora tanto o retorno quanto o risco da matriz energética brasileira. O portfólio PSW2020 diminui levemente os riscos, ao passo que reduz drasticamente os retornos. Por sua vez, o portfólio BAU2020 apresenta um grande acréscimo de retorno, mas que vem acompanhado de um aumento também do risco. O portfólio PDE2020 é praticamente equivalente à atual matriz brasileira, com pioras leves tanto de risco quanto de retorno. Diferentemente das projeções, o portfólio que nós propomos, MKW2014, caminha na direção noroeste do plano risco-retorno: ele aumenta o retorno ao mesmo tempo em que diminui o risco, sendo portanto a única proposta que alia os benefícios tanto do aumento do retorno quando de redução de riscos.

O segundo ponto relevante a ser notado é que as três projeções caminham em direções distintas do plano. PSW2020, por focar em fontes energéticas sustentáveis, resulta num decréscimo do retorno. No entanto, o risco também diminuiu, indicativo de que as tecnologias foram bem selecionadas. BAU2020 e PDE2020, dada a menor preocupação

ecológica, resultam em portfólios com retornos no mínimo iguais aos atuais. O cenário *business-as-usual* sacrificou a redução do risco em função do grande aumento de retorno, enquanto que o Plano Decenal de Expansão da Energia foi extremamente conservador quanto às mudanças propostas, alterando pouco a matriz segundo o plano risco-retorno.

O último ponto a salientar é que a natureza do portfólio MKW2014 é diferente da dos demais portfólios acima. PSW, BAU e PDE são projeções atreladas a hipóteses altamente realísticas. O portfólio MKW2014, apesar de adicionarmos restrições de factibilidade, corresponde a um ótimo teórico, um modelo a ser usado como guia para políticas energéticas futuras. Desta maneira, o papel desempenhado por estes portfólios é distinto. Enquanto que as projeções representam cenários prováveis, o portfólio MKW2014 simboliza a matriz energética ideal para o Brasil, do ponto de vista de retorno e risco – um alvo a ser alcançado.

A recomendação final deste trabalho é que, do ponto de vista de retorno e risco, a matriz energética brasileira caminhe na direção ao portfólio MKW2014, aliando aumento de energia por unidade monetária investida (retorno) e redução de riscos.

5. CONCLUSÕES

Com o objetivo de realizar uma análise da matriz energética brasileira sob a perspectiva de teoria de portfólios, iniciamos o estudo selecionando um conjunto de modelos de otimização de portfólio da literatura acadêmica composto por 48 modelos. Estes modelos são gerados a partir das combinações entre seis fatores, classificados entre extrínsecos, caso alterem apenas a o modo de utilização do modelo, ou estruturais, caso tragam diferenças profundas no funcionamento dos modelos.

Analizamos estes modelos sob as métricas de desempenho e rugosidade, determinando que o modelo tradicional de Markowitz tem o melhor balanço global entre desempenho e rugosidade. Este modelo, não obstante o grande número de variações analisadas neste trabalho, indubitavelmente mantém a sua posição como modelo de melhor desempenho: no nosso teste final, os resultados do Markowitz tradicional foram pelo menos 70% superiores aos dos outros modelos.

Adicionalmente, confirmamos estudos anteriores e verificamos que utilizar otimização bayesiana diminui a rugosidade dos modelos, sendo portanto um modelo mais robusto de otimização. Infelizmente, obtivemos evidência de que, simultaneamente, este modelo piora o desempenho global da otimização.

Constatamos que o CVaR teve uma média de desempenho sempre abaixo do modelo clássico de Markowitz, bem como uma rugosidade sempre acima. Desta maneira, tais características não justificam a pesada tarefa computacional de calcular o CVaR em cada iteração. As redes neurais, por sua vez, demonstraram um ótimo desempenho; porém, como a rugosidade delas aumentou desproporcionalmente em relação ao modelo tradicional de Markowitz, a razão desempenho/rugosidade diminuiu.

Após este estudo preliminar, prosseguimos à otimização de dois parâmetros numéricos do modelo tradicional de Markowitz, através de algoritmos evolutivos. Encontramos uma combinação de parâmetros capaz de melhorar em 331% o desempenho ao longo do período dos 3400 dias de negociação que antecederam 15 de maio de 2014 (13 anos). Ao reparametrizar os outros modelos, todos tiveram melhoras significativas (entre 320% e 1518%). No entanto, Markowitz (Close) continuou sendo o melhor, com um desempenho 60% melhor do que o

segundo colocado, Bayes (Close). Novamente, a otimização bayesiana apresenta resultados mais robustos, conquanto a razão entre desempenho e rugosidade continue pior do que no modelo clássico de Markowitz.

Em seguida, já em posse do modelo de melhor desempenho, prosseguimos à aplicação do modelo à matriz energética nacional. Localizamos a atual matriz em relação ao plano de risco-retorno e demonstramos através da fronteira eficiente que ela é sub-ótima.

De acordo com diversas restrições de factibilidade que impusemos, tais que reduzir a proporção de energia hidrelétricas apenas até certo ponto que consideramos plausível, propusemos uma matriz energética que concilia uma redução considerável do risco do portfólio e um aumento significativo do retorno, que é definido como a energia gerada por cada unidade monetária paga pelo governo. Incentivamos, nesta matriz, o uso de fontes energéticas renováveis não-hídricas. A recomendação final deste trabalho, em relação à matriz brasileira, é se aproximar deste portfólio ótimo que propomos, que denominamos de MKW2014.

Posteriormente, analisamos a sensibilidade da otimização em relação à proporção mínima de energia hidrelétrica, concluindo que, quanto menos energia hidrelétrica for requerida, melhor é para o país, em termos de risco-retorno.

Finalmente, comparamos os nossos resultados com outras três projeções encontradas na literatura, sob a perspectiva do nosso modelo. De acordo com a posição da atual matriz brasileira, verificamos que as outras projeções trariam mais consequências negativas do que positivas – diferentemente da nossa proposta, que seria benéfica tanto em termos de risco quanto de retorno.

REFERÊNCIAS BIBLIOGRÁFICAS

AHMADI-JAVID, A. **Entropic Value-at-Risk: A New Coherent Risk Measure**. Journal of Optimization Theory and Applications, volume 155, pp 1105-1123, 2012.

AWERBUCH, S. **Investing in photovoltaics: risk, accounting and the value of new technology**, Energy Policy, volume 28, pp 1023-1035, 2000

AWERBUCH, S.; BERGER, M. **Applying Portfolio Theory To EU Electricity Planning And Policy-Making**. IEA/EET Working Paper, 2003

AWERBUCH, S.; SPENCER, Y. **Efficient electricity generating portfolios for Europe: Maximising energy security and climate change mitigation**. EIB Papers, 2007.

BRANKER, K. *et al.* **A review of solar photovoltaic levelized cost of electricity**. Renewable and Sustainable Energy Reviews, volume 15, 2011.

COREY, G. R. **An Economic Comparison of Nuclear, Coal, and Oil-Fired Electric Generation in the Chicago Area**. Annual Review of Energy, volume 6, pp 417-443, 1981.

DELARUE, E. *et al.* **Applying portfolio theory to the electricity sector: energy versus power**. Energy Economics, vol. 33, 2011.

FABOZZI, F. J *et al.* **Robust Portfolio Optimization and Management**. New Jersey (EUA): John Wiley & Sons, Inc., 2007.

GOLDBERG, D. **Genetic Algorithms in Search, Optimization and Machine Learning**. Addison-Wesley Professional, 1989.

GRUBER, M. J. *et al.* **Modern Portfolio Theory and Investment Analysis**. 9ª edição. J. Wiley and Sons, 2014.

KOLM, P. N. *et al.* **60 Years of Portfolio Optimization: Practical Challenges and Current Trends**. European Journal of Operational Research 234, 2013.

LOSEKANN, L. *et al.* **Efficient Power Generating Portfolio in Brazil: Conciliating Cost, Emissions and Risk**. Energy Policy ,v.62, p.301–314, 2013.

MARKOWITZ, H. M. **Portfolio Selection**. The Journal of Finance 7: 77–91, 1952.

WWF-Brazil. **Sustainable Power Sector Vision 2020**, 2012. Disponível em: <<http://goo.gl/rwbdCq>>. Acessado em: 10 de agosto de 2014.

MEUCCI, A. **Robust Bayesian Allocation**, 2011. Disponível em: <<http://ssrn.com/abstract=681553>>. Acesso em: 10 de julho de 2014.

Ministério de Minas e Energia. **Balanco Energético Nacional**, 2014. Disponível em: <<http://goo.gl/2RsdMU>>. Acesso em: 10 de Agosto de 2014.

Ministério de Minas e Energia. **Plano Decenal de Expansão de Energia**, 2012. Disponível em: <<http://goo.gl/u7bqmm>>. Acesso em: 10 de Agosto de 2014.

MITCHELL, M. **An Introduction to Genetic Algorithms**. MIT Press, 1996.

MOČKUS, J. **On Bayesian Methods for Seeking the Extremum**. Automatika i vychislitel'naja tekhnika, N 3, 1972.

REGO, E.E. **Proposta de aperfeiçoamento da metodologia dos leilões de comercialização de energia elétrica no ambiente regulado: aspectos conceituais, metodológicos e suas aplicações**. Tese (Doutorado em Ciências) – Programa de Pós-Graduação em Energia da Universidade de São Paulo, São Paulo, 2012.

ROCKAFELLAR, R. T.; URYASEV, S. **Optimization of Conditional Value-at-Risk**. The Journal of Risk, Vol. 2, No. 3, 2000,

ROUDIER, F. **Portfolio Optimization and Genetic Algorithms**. Tese de Mestrado apresentada ao Swiss Federal Institute of Technology (ETH, Zurich) e à École Nationale de Ponts et Chaussées (ENPC, Paris), 2007.

SCHMIDHUBER, J. *et al.* **Multi-Column Deep Neural Network for Traffic Sign Classification**. Neural Networks, 2012.

U.S. Department of Energy. **Electric Power Annual 2009**. Disponível em: <<http://pubupws.nrc.gov/docs/ML1104/ML110410547.pdf>>. Acesso em: 5 de julho de 2014.

VAN HOEVE, W.-J. **Operations Research Techniques in Constraint Programming**, Tese de Doutorado apresentado à Universiteit van Amsterdam, 2005.

APÊNDICE

Matriz das Covariâncias

De acordo com os dados da BNEF e na ordem de fontes utilizada no corpo do trabalho:

0.10	-0.03	0.00	0.00	-0.11	0.06	-0.05	0.05	-0.01	0.01	0.03	0.05	0.06	0.09	-0.05	0.08	0.00	-0.02	-0.02
-0.03	0.22	-0.02	0.01	0.42	-0.09	0.29	-0.04	0.01	-0.01	0.08	-0.11	-0.20	-0.28	0.10	-0.26	0.00	0.09	0.04
0.00	-0.02	0.08	-0.05	-0.06	-0.01	0.00	0.00	-0.12	0.00	0.00	0.01	-0.03	-0.03	0.00	-0.01	0.00	0.00	0.00
0.00	0.01	-0.05	0.04	0.05	0.01	0.00	0.00	0.09	0.00	0.00	-0.01	0.01	0.02	-0.01	-0.01	0.00	0.00	0.00
-0.11	0.42	-0.06	0.05	1.13	-0.21	0.58	-0.13	0.07	0.01	0.18	-0.25	-0.46	-0.64	0.23	-0.61	0.01	0.20	0.10
0.06	-0.09	-0.01	0.01	-0.21	0.10	-0.14	0.05	0.02	0.02	-0.02	0.05	0.10	0.17	-0.07	0.14	0.01	-0.05	-0.02
-0.05	0.29	0.00	0.00	0.58	-0.14	0.45	-0.08	-0.03	0.00	0.14	-0.11	-0.24	-0.39	0.15	-0.33	0.00	0.12	0.06
0.05	-0.04	0.00	0.00	-0.13	0.05	-0.08	0.04	0.01	0.00	-0.01	0.04	0.06	0.10	-0.04	0.08	0.00	-0.03	-0.02
-0.01	0.01	-0.12	0.09	0.07	0.02	-0.03	0.01	0.25	0.00	-0.01	0.01	0.07	0.10	-0.03	0.03	0.00	-0.01	0.00
0.01	-0.01	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00
0.03	0.08	0.00	0.00	0.18	-0.02	0.14	-0.01	-0.01	0.01	0.12	0.00	-0.04	-0.09	0.03	-0.08	0.00	0.03	0.01
0.05	-0.11	0.01	-0.01	-0.25	0.05	-0.11	0.04	0.01	0.00	0.00	0.12	0.19	0.21	-0.06	0.20	-0.01	-0.06	-0.03
0.06	-0.20	-0.03	0.01	-0.46	0.10	-0.24	0.06	0.07	0.00	-0.04	0.19	0.35	0.40	-0.11	0.37	-0.02	-0.12	-0.06
0.09	-0.28	-0.03	0.02	-0.64	0.17	-0.39	0.10	0.10	0.00	-0.09	0.21	0.40	0.53	-0.17	0.46	-0.01	-0.16	-0.08
-0.05	0.10	0.00	-0.01	0.23	-0.07	0.15	-0.04	-0.03	-0.01	0.03	-0.06	-0.11	-0.17	0.07	-0.14	0.00	0.05	0.03
0.08	-0.26	-0.01	-0.01	-0.61	0.14	-0.33	0.08	0.03	0.00	-0.08	0.20	0.37	0.46	-0.14	0.43	-0.01	-0.14	-0.07
0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	-0.01	-0.02	-0.01	0.00	-0.01	0.00	0.00	0.00
-0.02	0.09	0.00	0.00	0.20	-0.05	0.12	-0.03	-0.01	0.00	0.03	-0.06	-0.12	-0.16	0.05	-0.14	0.00	0.05	0.02
-0.02	0.04	0.00	0.00	0.10	-0.02	0.06	-0.02	0.00	0.00	0.01	-0.03	-0.06	-0.08	0.03	-0.07	0.00	0.02	0.01