
UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

**DESENVOLVIMENTO DE MODELOS PARA PREVISÃO
DE *SPREAD* DO PREÇO DE ENERGIA ELÉTRICA NO
MERCADO DE CURTO PRAZO**

Autor:
Jefferson Rampazo Amadeu

Nº USP:
5910623

Orientador:
Eng. Rodrigo Sacchi, D. Sc.

Supervisor:
Prof. Dr. Adriano Alber de França Mendes Carneiro

São Carlos, SP
2011

JEFFERSON RAMPAZO AMADEU

DESENVOLVIMENTO DE MODELOS PARA PREVISÃO DE *SPREAD* DO PREÇO DE ENERGIA ELÉTRICA NO MERCADO DE CURTO PRAZO

Trabalho de Conclusão de Curso
apresentado à Escola de Engenharia de
São Carlos da Universidade de São Paulo

Curso de Engenharia Elétrica com Ênfase
em Sistemas de Energia e Automação

Orientador: Eng. Rodrigo Sacchi, D.Sc.

Supervisor: Prof. Dr. Adriano Alber de
França Mendes Carneiro

São Carlos
2011

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

A481d Amadeu, Jefferson Rampazo
Desenvolvimento de modelos para previsão de spread do
preço de energia elétrica no mercado de curto prazo. /
Jefferson Rampazo Amadeu ; orientador Rodrigo Sacchi --
São Carlos, 2011.

Monografia (Graduação em Engenharia Elétrica com
ênfase em Sistemas de Energia e Automação) -- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2011.

1. Comercialização de energia. 2. Mercado de curto
prazo. 3. Spread. 4. Modelos de previsão. 5. Modelos
autorregressivos. 6. Redes neurais. 7. Séries temporais
I. Título.

FOLHA DE APROVAÇÃO

Nome: Jefferson Rampazo Amadeu

Título: "Desenvolvimento de Modelos para Previsão de Spread do Preço de Energia Elétrica no Mercado de Curto Prazo"

Trabalho de Conclusão de Curso defendido e aprovado
em 29/11/2011,

com NOTA por 9,0, pela comissão julgadora:



Prof. Associado Ivan Nunes da Silva - EESC/USP



Prof. Dr. Adriano Alber de França Mendes Carneiro - EESC/USP



Prof. Associado Homero Schiabel
Coordenador da CoC-Engenharia Elétrica
EESC/USP

DEDICATÓRIA

Dedico este trabalho a meus pais, Edson e Márcia, pelas grandes oportunidades a que me proporcionaram e pelos conselhos que levarei comigo por toda a vida.

AGRADECIMENTOS

À todos os grandes colegas que fiz durante estes cinco anos de graduação, cuja companhia foi fundamental durante esta etapa de minha vida.

Á todos os amigos que me apoiaram e acreditaram no meu sucesso, especialmente à Yumi, Flávia, Mônica e Diogo, pela grande presença em minha vida nestes últimos anos.

Á minha família, cuja importância é imensurável.

Ao Rodrigo Sacchi, por sua orientação e oportunidade de desenvolvimento profissional.

Ao professor Adriano Alber de França Mendes Carneiro, pelo ensino durante a graduação e supervisão deste trabalho.

Á todos colegas da CPFL Energia pela oportunidade e pelo ótimo ambiente de trabalho.

“Let us think the unthinkable, let us do the undoable, let us prepare to grapple with the ineffable itself, and see if we may not eff it after all.”
Douglas Adams

SUMÁRIO

DEDICATÓRIA	5
AGRADECIMENTOS	7
LISTA DE ILUSTRAÇÕES	11
LISTA DE TABELAS	13
LISTA DE SIGLAS, SÍMBOLOS E ABREVIATURAS	15
RESUMO	17
ABSTRACT	19
CAPÍTULO 1	21
INTRODUÇÃO	21
1.1) FORMAÇÃO DE PREÇOS DE ENERGIA ELÉTRICA NO MERCADO BRASILEIRO	22
1.2) OBJETIVOS E ORGANIZAÇÃO DO TRABALHO	24
CAPÍTULO 2	25
SÉRIES TEMPORAIS	25
2.1) ANÁLISE DE SAZONALIDADE	30
2.2) AUTOCORRELAÇÃO PARCIAL	31
2.3) ANÁLISE DO COMPONENTE PRINCIPAL	33
CAPÍTULO 3	37
MODELOS LINEARES	37
3.1) MODELO NAIVE (INGÊNUO)	37
3.2) MODELOS AUTORREGRESSIVOS	39
3.2.1) <i>Modelo Autorregressivo de Primeira Ordem</i>	40
3.2.2) <i>Modelo Autorregressivo de Segunda Ordem</i>	42
3.3) MODELO AUTORREGRESSIVO EXÓGENO	43
3.3.1) <i>Modelo ARX de primeira ordem</i>	44
3.3.2) <i>Modelo ARX utilizando PCA</i>	45
CAPÍTULO 4	47
REDES NEURAIS ARTIFICIAIS	47
4.1) REDE PERCEPTRON MULTICAMADAS	49
4.1.1) <i>Spread como entrada da PMC</i>	51
4.1.2) <i>Variáveis defasadas de um mês como entradas da PMC</i>	53
4.1.3) <i>PCA como entradas da PMC</i>	53
4.1.4) <i>Parada Prematura</i>	57
CAPÍTULO 5	61
ANÁLISE DAS PREVISÕES	61
CAPÍTULO 6	67
CONCLUSÃO E DISCUSSÕES	67
BIBLIOGRAFIA	69
APÊNDICE A	71
ESTIMADORES DE ERRO	71
<i>Erro Quadrático Médio</i>	71
<i>Erro Médio Absoluto</i>	72
<i>Erro Percentual Absoluto Médio</i>	72

LISTA DE ILUSTRAÇÕES

FIGURA 1 - PREÇO DA ENERGIA X <i>SPREAD</i>	21
FIGURA 2 – RELAÇÃO ENTRE <i>SPREAD</i> E VARIÁVEIS DEFASADAS.....	22
FIGURA 3 – PROCESSO DE COMERCIALIZAÇÃO DE ENERGIA NA CCEE	23
FIGURA 4 – SÉRIE TEMPORAL: <i>SPREAD</i>	25
FIGURA 5 – SÉRIE TEMPORAL: PLD.....	26
FIGURA 6 – SÉRIE TEMPORAL: DEMANDA ENERGÉTICA.....	27
FIGURA 7 – SÉRIE TEMPORAL: GERAÇÃO HIDROELÉTRICA	27
FIGURA 8 – SÉRIE TEMPORAL: GERAÇÃO TÉRMICA.....	28
FIGURA 9 – SÉRIE TEMPORAL: ENA	29
FIGURA 10 – SÉRIE TEMPORAL: ENERGIA ARMAZENADA	29
FIGURA 11 – ANÁLISE DE SAZONALIDADE DA SÉRIE TEMPORAL <i>SPREAD</i>	31
FIGURA 12 – COEFICIENTES DE AUTOCORRELAÇÃO PARCIAL DA SÉRIE <i>SPREAD</i>	32
FIGURA 13 – APLICAÇÃO DE PCA EM DUAS VARIÁVEIS CORRELACIONADAS (SACCHI, 2009).....	33
FIGURA 14 – VARIÂNCIA EXPLICADA PELOS COMPONENTES PRINCIPAIS DA SÉRIE <i>SPREAD</i>	34
FIGURA 15 – COEFICIENTES DOS PRIMEIROS COMPONENTES PRINCIPAIS	35
FIGURA 16 – PARTICIPAÇÃO DAS VARIÁVEIS ORIGINAIS NOS COMPONENTES PRINCIPAIS	36
FIGURA 17 – <i>SPREAD</i> E PREVISÃO DO MODELO NAIVE	38
FIGURA 18 – <i>SPREAD</i> E PREVISÃO DO MODELO $AR(1)_{YULE-WALKER}$	41
FIGURA 19 - <i>SPREAD</i> E PREVISÃO DO MODELO $AR(1)_{MÍNIMOS\ QUADRADOS}$	42
FIGURA 20 – <i>SPREAD</i> E PREVISÃO DO MODELO $AR(2)$	43
FIGURA 21 – <i>SPREAD</i> E PREVISÃO DO MODELOS ARX DE PRIMEIRA ORDEM	44
FIGURA 22 – <i>SPREAD</i> E PREVISÃO DOS MODELOS ARX UTILIZANDO RESULTADOS DA PCA	45
FIGURA 23 – NEURÔNIO ARTIFICIAL DE UMA RNA.....	47
FIGURA 24 – EXEMPLOS DE FUNÇÕES DE ATIVAÇÃO	48
FIGURA 25 – ARQUITETURA BÁSICA DE UMA REDE PERCEPTRON MULTICAMADAS	50
FIGURA 26 – RESULTADOS DAS PMC COM ENTRADA DE $SPREAD_{T-1}$	51
FIGURA 27 – MELHORES RESULTADOS DAS PMC COM ENTRADAS HISTÓRICAS DO <i>SPREAD</i>	52
FIGURA 28 – RESULTADOS DA PMC COM VARIÁVEIS DO SISTEMA.....	53
FIGURA 29 – RESULTADOS DA PMC COM COMPONENTES PRINCIPAIS DE 85% DA VARIÂNCIA	54
FIGURA 30 – RESULTADOS DA PMC COM COMPONENTES PRINCIPAIS DE 90% DA VARIÂNCIA	55
FIGURA 31 – RESULTADOS DA PMC COM COMPONENTES PRINCIPAIS DE 95% DA VARIÂNCIA	55
FIGURA 32 – RESULTADOS DA PMC COM COMPONENTES PRINCIPAIS DE 100% DA VARIÂNCIA	56
FIGURA 33 – TREINAMENTO DE PMC COM CRITÉRIO DE PARADA	57
FIGURA 34 – AJUSTE DE PMC COM PCA 85% E CRITÉRIO DE PARADA	58
FIGURA 35 – AJUSTE DE PMC COM PCA 90% E CRITÉRIO DE PARADA	58
FIGURA 36 – AJUSTE DE PMC COM PCA 95% E CRITÉRIO DE PARADA	59
FIGURA 37 – AJUSTE DE PMC COM PCA 100% E CRITÉRIO DE PARADA	59
FIGURA 38 – PREVISÕES DO BLOCO (A)	62
FIGURA 39 – PREVISÕES DO BLOCO (B).....	63
FIGURA 40 – PREVISÕES DO BLOCO (C)	63
FIGURA 41 – PREVISÕES DO BLOCO (D)	64
FIGURA 42 – PREVISÕES DO BLOCO (E).....	65

LISTA DE TABELAS

TABELA 1 – ERROS DE AJUSTE DO MODELO <i>NAIVE</i>	38
TABELA 2 – ERROS DE AJUSTE DO MODELO $AR(1)_{YULE-WALKER}$	41
TABELA 3 – ERROS DE AJUSTE DO MODELO $AR(1)_{MÍNIMOS\ QUADRADOS}$	42
TABELA 4 – ERROS DE AJUSTE DO MODELO $AR(2)$	43
TABELA 5 – COEFICIENTES DO MODELO ARX DE PRIMEIRA ORDEM.....	44
TABELA 6 – ERROS DE AJUSTE DO MODELO ARX DE PRIMEIRA ORDEM	45
TABELA 7 – COEFICIENTES DOS MODELOS ARX UTILIZANDO RESULTADOS DA PCA	46
TABELA 8 - ERROS DE AJUSTE DA PMC COM ENTRADA DE $SPREAD_{t-1}$	51
TABELA 9 - ERROS DE AJUSTE DA PMC COM ENTRADAS HISTÓRICAS DO $SPREAD$	52
TABELA 10 - ERROS DE AJUSTE DA PMC COM VARIÁVEIS DO SISTEMA	53
TABELA 11 - ERROS DE AJUSTE DAS PMC COM PCA COMO ENTRADAS	56
TABELA 12 - ERROS DE AJUSTE DAS PMC COM PCA E CRITÉRIO DE PARADA	60
TABELA 13 – ERROS DE PREVISÃO DOS MODELOS APRESENTADOS.....	61

LISTA DE SIGLAS, SÍMBOLOS E ABREVIATURAS

<i>ACL</i>	Ambiente de Contratação Livre
<i>ACR</i>	Ambiente de Contratação Regulado
<i>ANEEL</i>	Agência Nacional de Energia Elétrica
<i>ANN</i>	Artificial Neural Networks
<i>AR</i>	Modelo Autorregressivo
<i>AR(p)</i>	Modelo Autorregressivo de ordem p
<i>ART</i>	Adaptive Resonance Theory
<i>ARX</i>	Modelo Autorregressivo Exógeno
<i>CCEE</i>	Câmara Comercializadora de Energia Elétrica
<i>CMO</i>	Custo Marginal de Operação
<i>EArm</i>	Energia Armazenada do Subsistema
<i>ENA</i>	Energia Natural Afluyente
<i>FACP</i>	Função de Autocorrelação Parcial
<i>GerH</i>	Geração Hidráulica do Subsistema
<i>GerT</i>	Geração Térmica do Subsistema
<i>GWmed</i>	Gigawatt médio
<i>GWh</i>	Gigawatt-hora
<i>MAE</i>	Erro Absoluto Médio - Mean Absolute Error
<i>MAPE</i>	Erro Absoluto Percentual Médio - Mean Absolute Percentual Error
<i>MLP</i>	Multilayer Perceptron
<i>MLT</i>	Média de Longo Termo
<i>MMQ</i>	Método dos Mínimos Quadrados
<i>MSE</i>	Erro Quadrático Médio - Mean Squared Error
<i>MWh</i>	Megawatt-hora
<i>N</i>	Submercado Norte
<i>Naive</i>	Modelo Naive
<i>NE</i>	Submercado Nordeste
<i>ONS</i>	Operador Nacional do Sistema
<i>PAR(p)</i>	Modelo Autorregressivo Periódico de ordem p
<i>PCA</i>	Análise do Componente Principal - Principal Componente Analysis
<i>PCA 100%</i>	Componentes Principais que correspondem a 100% da variância
<i>PCA 85%</i>	Componentes Principais que correspondem a 85% da variância
<i>PCA 90%</i>	Componentes Principais que correspondem a 90% da variância

<i>PCA 95%</i>	Componentes Principais que correspondem a 95% da variância
<i>PLD</i>	Preço de Liquidação das Diferenças
<i>PMC</i>	Perceptron Multicamadas
<i>R\$/MWh</i>	Reais por Megawatt-hora
<i>RNA</i>	Rede Neural Artificial
<i>S</i>	Submercado Sul
<i>SE-CO</i>	Submercado Sudeste/Centro-Oeste
<i>UTE</i>	Usina Termoelétrica

RESUMO

AMADEU, J. R. (2011). **Desenvolvimento de Modelos para a Previsão de *Spread* do Preço de Energia Elétrica no Mercado de Curto Prazo**. Monografia (Graduação) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2011.

Este trabalho consiste no desenvolvimento de modelos de previsão aplicados na predição do *Spread* (Ágio) do preço de energia elétrica praticado pelos agentes vendedores no mercado de curto prazo. Foram avaliados os desempenhos das técnicas de regressão linear (Modelo Autorregressivo - AR e Modelo Autorregressivo Exógeno - ARX) e dos modelos não lineares (Redes Neurais Artificiais - RNA). Como variáveis de entrada dos modelos de previsão foram utilizados históricos de variáveis influentes na operação energética brasileira: Geração Termelétrica, Geração Hidrelétrica, Preço de Liquidação das Diferenças (PLD), Demanda Energética, Energia Natural Afluenta (ENA), Energia Armazenada e o histórico do próprio *Spread*. Tais modelos de previsão podem ser utilizados como uma ferramenta de suporte no processo de planejamento de comercialização de energia elétrica.

Palavras-chave: comercialização de energia, mercado de curto prazo, *spread*, modelos de previsão, modelos autorregressivos, redes neurais, séries temporais.

ABSTRACT

AMADEU, J. R. (2011). **Forecast Model Design for Prediction of Energy Price Spread**. Monograph (Graduation) –Engineering School of São Carlos, University of São Paulo, São Carlos, 2011.

This research work aims to design of a forecast model applied on the electricity price spread used by the seller agents in Brazilian spot market. The performance of the linear regression technique (Auto-Regressive Model – AR and Auto-Regressive Exogenous Model – ARX) and the performance of the non-linear models (Artificial Neural Networks – ANN) were evaluated for the prediction. The historical of influent variables in the Brazilian power operation: Thermoelectric Power Generation, Hydroelectric Power Generation, Spot Price (PLD), Electricity Demand, Natural Inflow Energy (ENA), Stored Energy and the Spread itself were used as input variables on the prediction models. Such prediction models can be used as a support tool on the planning electricity commercialization.

Keywords: electricity commercialization, spot market, spread, forecast models, autoregressive models, neural networks, time series.

Capítulo 1

Introdução

A energia elétrica é hoje um insumo de suma importância na indústria mundial e, portanto seu preço tem grande influência no valor da produção de um país, impactando a todos.

No caso brasileiro, como a geração de energia elétrica é predominantemente hidráulica e esta depende de fatores naturais para a disponibilidade de recursos, os preços da energia elétrica variam em grandes proporções durante o ano, o que faz com que o planejamento de todos os agentes do setor energético e do setor industrial baseie-se fortemente em previsões de preço, visando ações lucrativas para os seus empreendimentos.

Com esta motivação, este trabalho visa obter modelos de previsão para o ágio praticado pelos agentes comercializadores na venda de energia elétrica. O ágio é normalmente referido no setor pelo seu termo em inglês, *Spread*, e varia tanto quanto o próprio preço da energia elétrica, como ilustra a Figura 1, tornando sua previsão bastante complicada, já que sua relação com outras variáveis da operação energética não é facilmente observada.

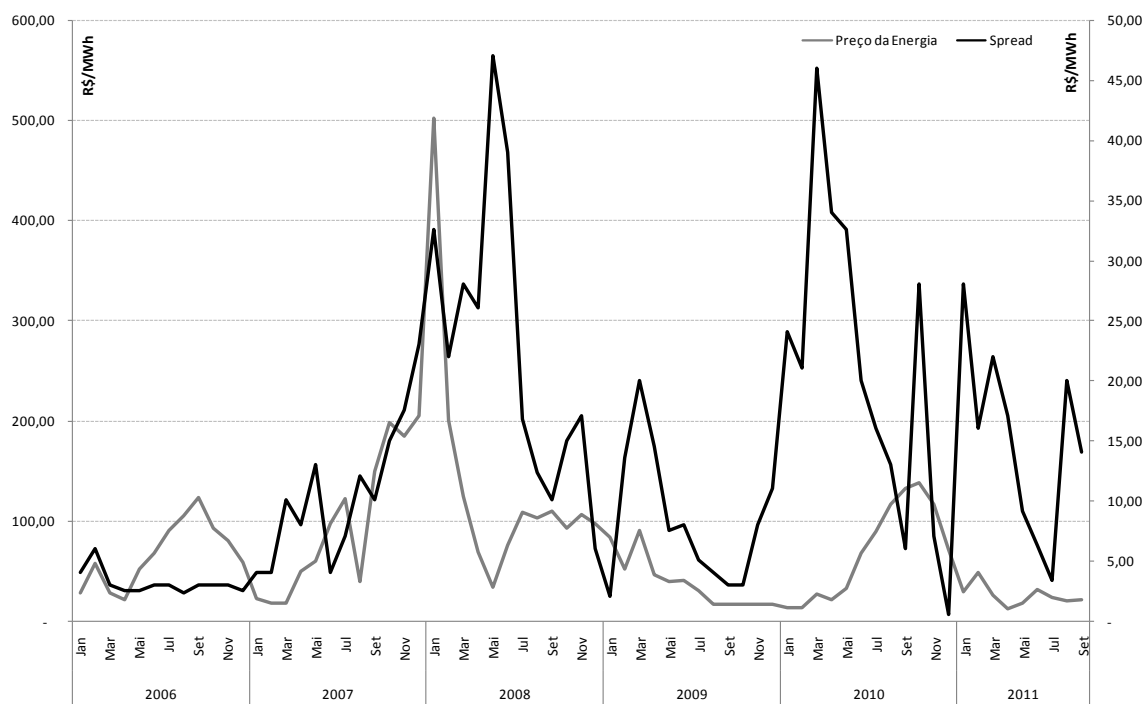


Figura 1 - Preço da Energia x *Spread*

A Figura 2 exibe a relação entre os *Spreads* praticados por uma comercializadora e os valores de algumas variáveis do sistema elétrico brasileiro defasadas em um, três, seis e doze

meses do *Spread*, segundo indicação dos resultados obtidos por (Sacchi, 2009). Percebe-se que os pontos são bastante esparsos e que uma relação entre o *Spread* e estas variáveis é difícil de ser visualizada graficamente.

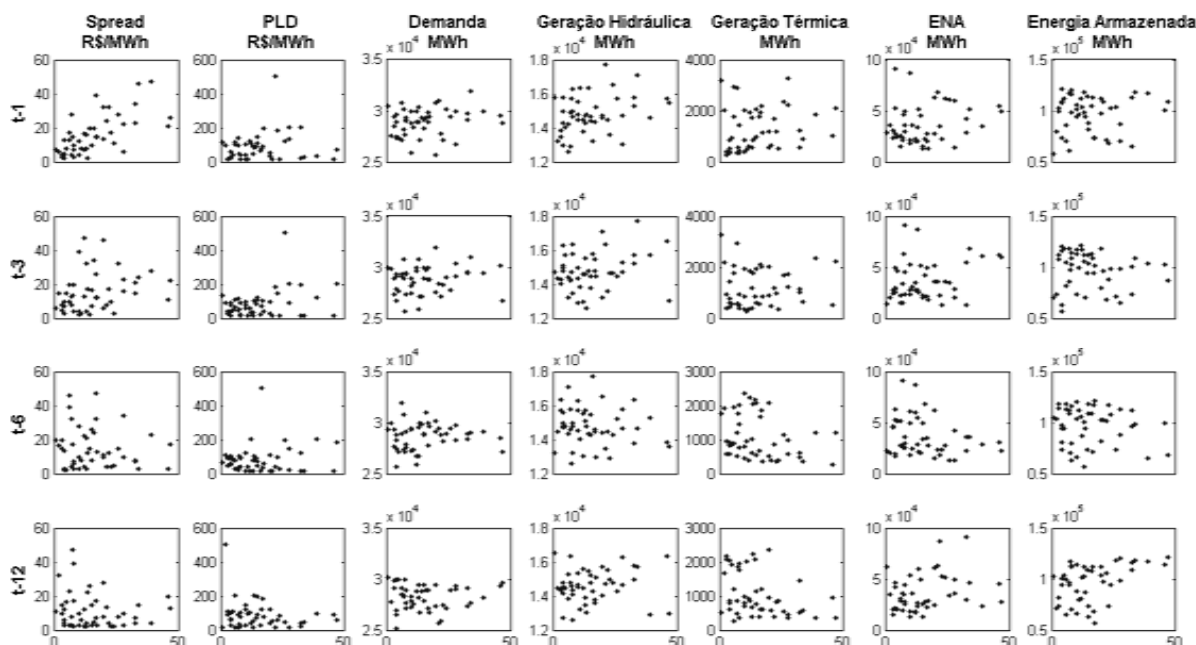


Figura 2 – Relação entre *Spread* e variáveis defasadas

1.1) Formação de preços de energia elétrica no mercado brasileiro

Atualmente o mercado de energia elétrica brasileiro é dividido em dois ambientes distintos de contratação, denominados Ambiente de Contratação Regulada (ACR) e Ambiente de Contratação Livre (ACL). Os preços praticados no ACR são definidos em leilões por menor valor e a energia é disponibilizada para as distribuidoras de energia, para que estas a repassem para os consumidores cativos a tarifas estipuladas pela Agência Nacional de Energia Elétrica (ANEEL). Já os preços praticados no ACL são livremente determinados em contratos bilaterais entre os agentes vendedores de energia (geradores e comercializadoras de energia) e os agentes compradores de energia (consumidores livres e comercializadoras de energia). Estes contratos são registrados pela Câmara Comercializadora de Energia Elétrica (CCEE).

A CCEE verifica a quantidade de energia produzida ou consumida por um agente e a quantidade de energia contratada do agente, conforme exibe a Figura 3. O balanço entre estas energias é também verificado em uma janela móvel de doze meses, sendo este montante

denominado lastro. No caso de um lastro negativo, o agente fica sujeito a penalidades junto à CCEE.

Durante os meses o balanço é liquidado ao Preço de Liquidação das Diferenças (PLD), que é definido semanalmente para cada um dos subsistemas brasileiros, denominados Sudeste / Centro-Oeste (SE-CO), Sul (S), Nordeste (NE) e Norte (N) e também para cada patamar de carga, estes dependentes do dia e horário em que a energia é demandada, sendo denominados patamares de carga leve, médio e pesado. O PLD tem como base o Custo Marginal de Operação (CMO) do sistema, limitado por um preço mínimo e um preço máximo. O mercado de curto prazo (mercado *Spot*) pode ser denominado então como o mercado que envolve as diferenças de montantes contratados e montantes consumidos ou produzidos pelos agentes (CCEE, 2011).



Figura 3 – Processo de comercialização de energia na CCEE

O CMO é o custo em reais para que o sistema gere 1MWh a mais do que o sistema demanda e é definido semanalmente para cada subsistema pela cadeia de modelos computacionais NEWAVE e DECOMP. O NEWAVE é um modelo de base mensal que define a política de operação ótima de médio prazo (até cinco anos à frente) do sistema elétrico, baseando-se no despacho (geração) de usinas cujo custo econômico e energético de operação seja o mais barato. Já o DECOMP é um modelo que utiliza saídas do NEWAVE, além de outras variáveis do sistema, para uma otimização do horizonte de curto prazo (até dois meses à frente), com base semanal para o primeiro mês.

A matriz energética brasileira é predominantemente hidráulica e a sua utilização máxima é a opção mais econômica em uma primeira análise, devido ao baixo custo de combustível em comparação com as Usinas Termoeletricas (UTE), mas a sua utilização excessiva pode causar déficits futuros para a geração, portanto os modelos ponderam as gerações, planejando uma operação ótima visando à confiabilidade futura de suprimento do sistema.

Ambos os modelos estão disponíveis para todos os agentes e, portanto são usados frequentemente, já que suas saídas estimam os valores dos CMOs e conseqüentemente de PLDs futuros, impactando diretamente na expectativa de preço do mercado de curto prazo e, deste modo, na operação de compra e venda de energia.

A energia no mercado de curto prazo pode ser comercializada livremente entre os agentes durante o seu mês de competência e até o nono dia útil do mês seguinte, trazendo outras possibilidades de preço além da liquidação ao PLD paga à (ou pela) CCEE, abrindo oportunidades de negociação de contratos de curto prazo.

Observando esta especulação de preços no mercado de curto prazo, os agentes vendedores de energia aplicam ágios ao PLD atual, conhecidos como *Spread*, aumentando a variação dos preços e dificultando ainda mais a previsão de preços, mesmo para futuros não muito distantes, como já foi mostrado na Figura 1.

1.2) *Objetivos e Organização do trabalho*

Este trabalho visa à predição do *Spread* que será aplicado pelos agentes vendedores no mercado de curto prazo no futuro, servindo de ferramenta na tomada de decisões contratuais dos consumidores livres e geradores de energia, utilizando variáveis do sistema elétrico como base de variáveis de entrada para modelos de previsão lineares e modelos de redes neurais artificiais.

O trabalho organiza-se da seguinte forma:

- **Capítulo 2:** Discorre sobre as séries temporais utilizadas neste trabalho e suas propriedades, visando à identificação das variáveis mais significativas para os modelos de previsão do *Spread* mensal.
- **Capítulo 3:** Apresenta os modelos lineares de previsão utilizados neste trabalho e seus resultados na aplicação da previsão do *Spread*.
- **Capítulo 4:** Apresenta as Redes Neurais Artificiais utilizadas neste trabalho e seus resultados na aplicação da previsão do *Spread*.
- **Capítulo 5:** Valida as previsões dos modelos dos dois capítulos anteriores com os dados históricos de 2011.
- **Capítulo 6:** Expõe a síntese dos resultados do trabalho e das pesquisas e discute a sua aplicação, além de sugerir possíveis pesquisas futuras

Capítulo 2

Séries Temporais

Uma série temporal é uma série de observações de uma variável ao longo do tempo, geralmente em períodos igualmente espaçados. No caso deste trabalho, estão sendo observadas sete séries temporais de variáveis do sistema elétrico brasileiro na região Sudeste / Centro-Oeste, com dados verificados mensalmente desde janeiro de 2006 até setembro de 2011.

A primeira série foi o objeto de estudo deste trabalho, o *Spread* médio aplicado por uma comercializadora brasileira sobre o PLD em R\$/MWh. Ela pode ser observada graficamente na Figura 4 e não aparenta seguir um padrão, em uma primeira análise, necessitando de análises mais elaboradas para uma melhor conclusão.

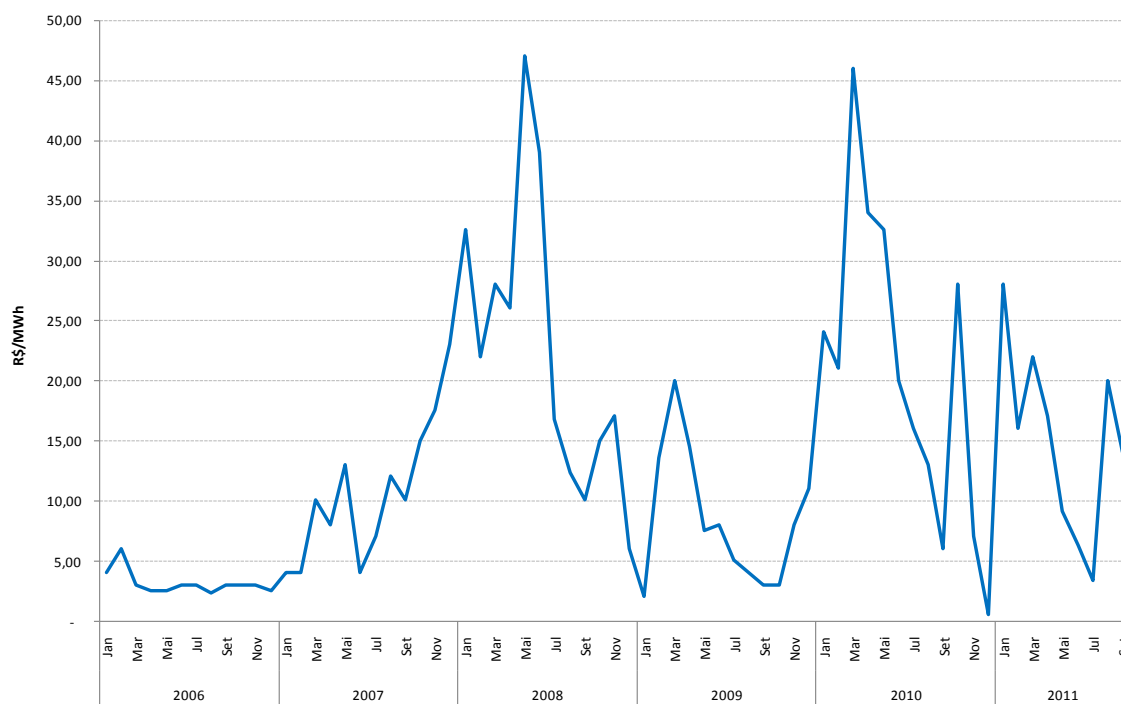


Figura 4 – Série temporal: *Spread*

Também foi observada a série histórica do PLD (CCEE, 2011), expresso em R\$/MWh para o mercado Sudeste/Centro-Oeste. A série pode ser visualizada na Figura 5 e nesta verifica-se um grande salto entre os anos de 2007 e 2008, causado pelo atraso do início do período chuvoso, normalmente esperado para o mês de Novembro.

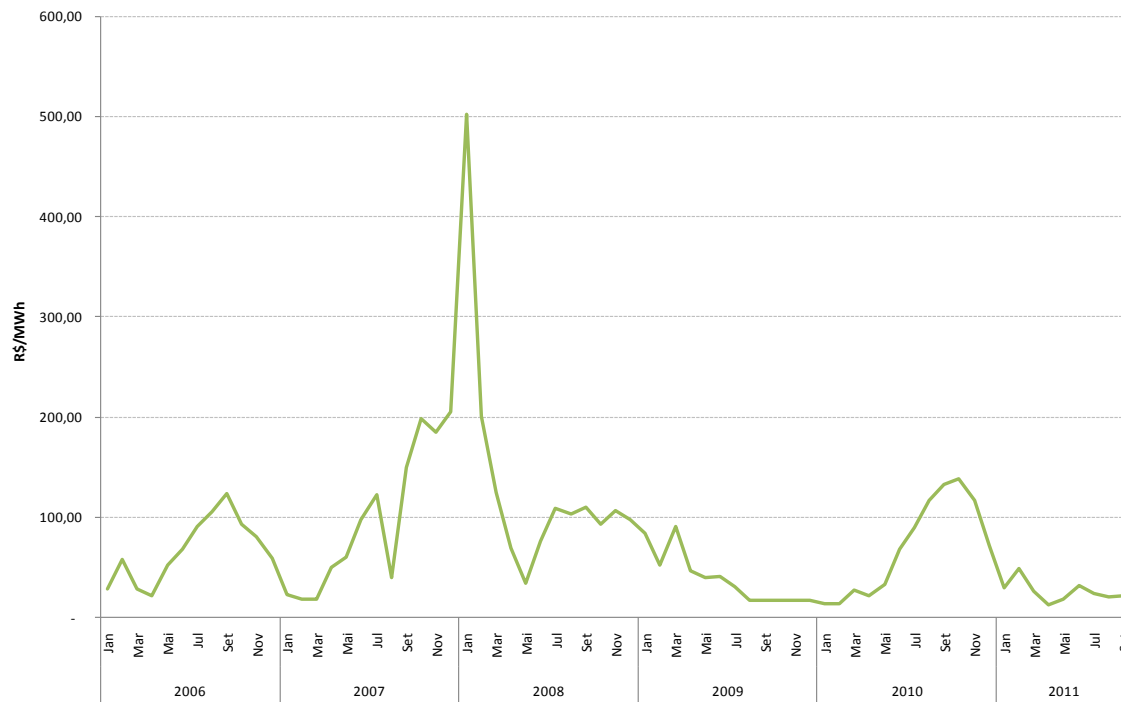
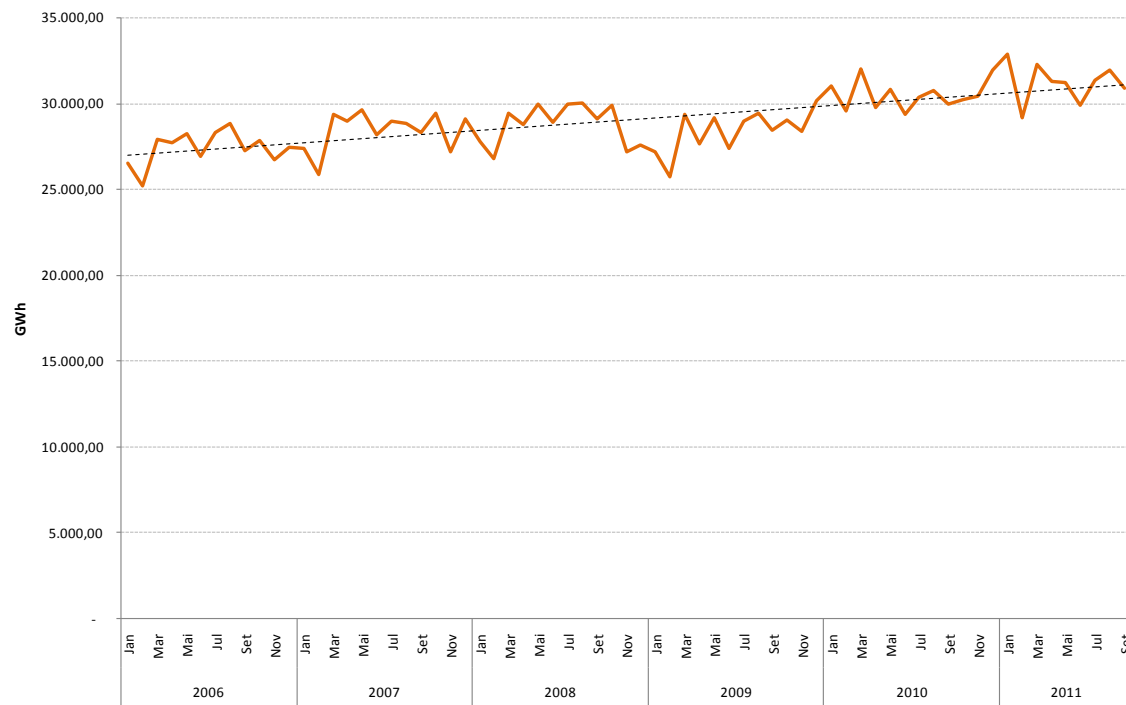
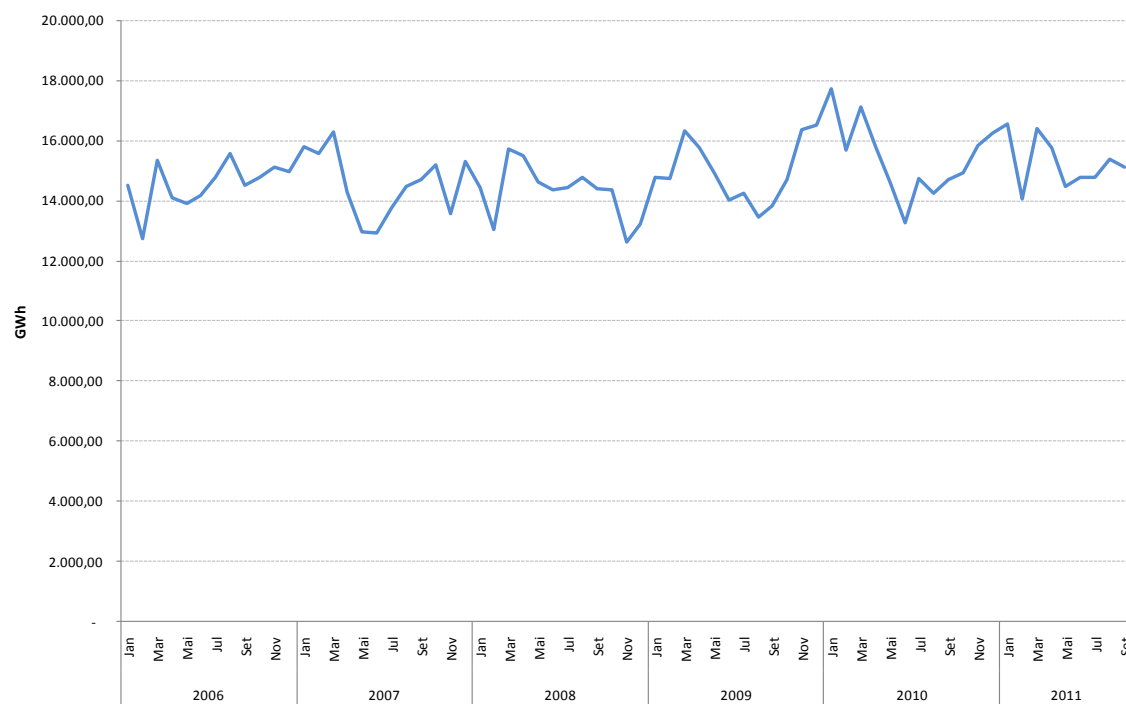


Figura 5 – Série temporal: PLD

A demanda energética do subsistema também foi verificada (ONS, 2011), em GWh. A Figura 6 exibe o histórico e é possível perceber através da linha de tendência tracejada que a demanda se mantém em ascensão durante os últimos cinco anos, crescendo por volta de cinco mil megawatts-hora, apesar de um afundamento entre 2008 e 2009 devido à crise financeira internacional.

**Figura 6 – Série temporal: Demanda energética**

A quarta série utilizada no trabalho foi a geração das usinas hidroelétricas (ONS, 2011), expressas em GWh e mostrada no gráfico da Figura 7.

**Figura 7 – Série Temporal: Geração Hidroelétrica**

A geração térmica em GWh do subsistema (ONS, 2011) também foi avaliada, em vista do seu forte impacto sobre o preço de energia, já que o seu custo é elevado. A Figura 8 exibe o seu histórico, onde se percebe uma relação entre seus picos e os picos de PLD.

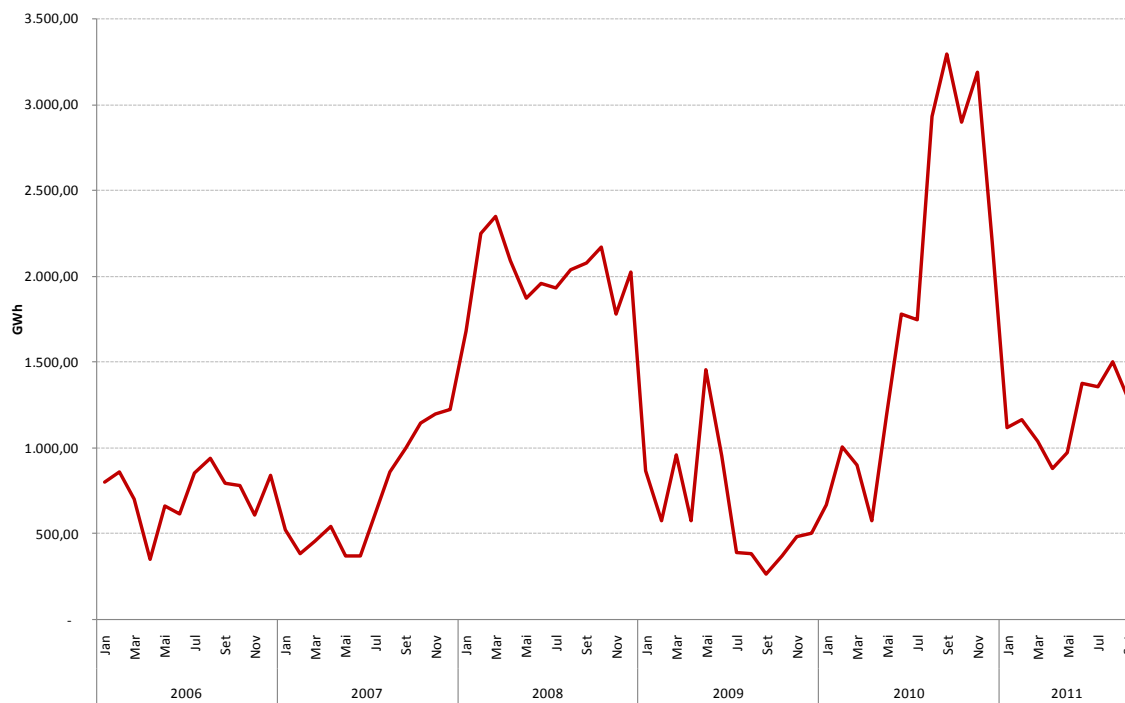


Figura 8 – Série temporal: Geração Térmica

Outra variável histórica do sistema observada foi a Energia Natural Afluentes (ENA) do subsistema Sudeste/Centro-Oeste (ONS, 2011). A ENA representa a energia que pode ser produzida a partir das vazões naturais afluentes aos reservatórios. Ela pode ser representada em Média de Longo Termo (MLT), percentual de variação da ENA atual em relação à média de 81 anos levantados pelo ONS ou pela sua correspondente em GW médios que pode ser observada para os últimos seis anos no gráfico da Figura 9, onde se percebe uma relação inversa com o PLD, havendo vales nos picos do PLD.

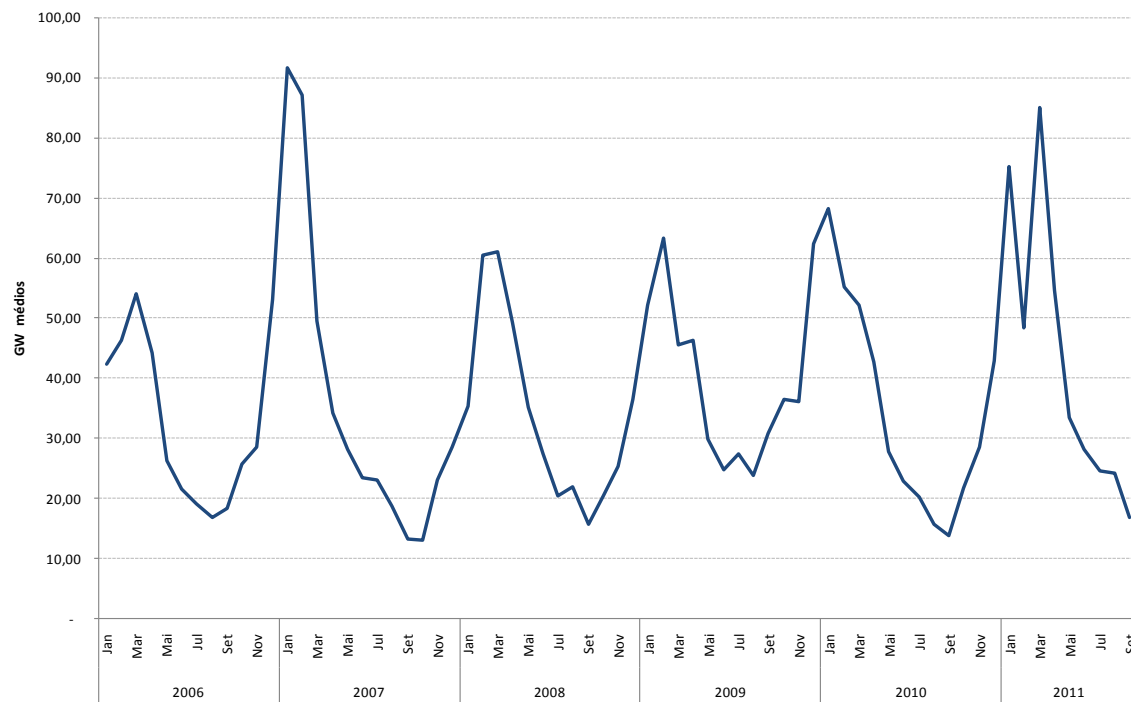


Figura 9 – Série temporal: ENA

A última série temporal observada foi a Energia Armazenada do sistema, exibida na Figura 10 em GWh, que corresponde em potencial de geração ao armazenamento dos reservatórios das hidrelétricas do sistema.

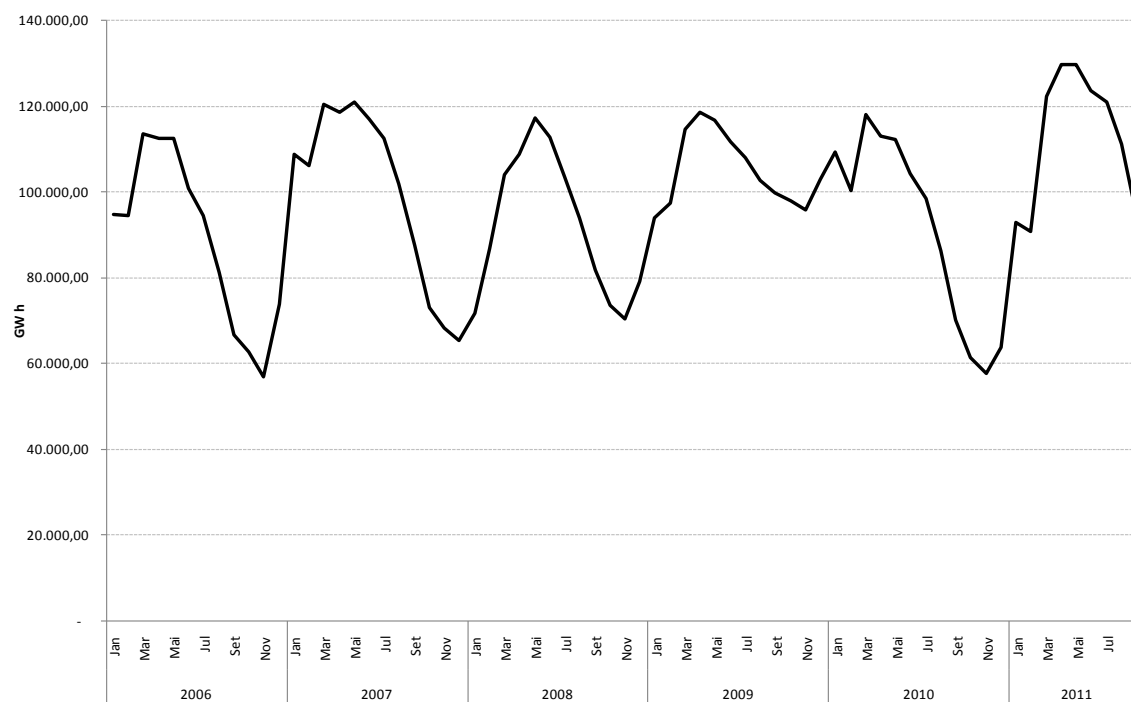


Figura 10 – Série Temporal: Energia Armazenada

2.1) *Análise de Sazonalidade*

Como a série temporal do *Spread* compreende cinco anos, é possível fazer a análise de sua sazonalidade anual com o objetivo de observar tendências mensais nos valores do *Spread*.

A análise é feita ao calcular a média dos valores de cada mês e a seu desvio padrão, cujas equações são, respectivamente:

$$\bar{x}_m = \frac{1}{N_m} \sum_{n=1}^{N_m} x_{m,n} \quad (1)$$

$$s_m = \sqrt{\frac{1}{N_m - 1} \sum_{n=1}^{N_m} (x_{m,n} - \bar{x}_m)^2} \quad (2)$$

sendo, para cada mês m :

$\bar{x}_m$	Média mensal;
N_m	Quantidade total de medições do mês m ;
$x_{m,n}$	Valor da série em sua n -ésima medição do mês m ;
s_m	Desvio padrão do mês.

Os resultados da análise podem ser visualizados na Figura 11, onde os pontos representam as médias, $\bar{x}_m$, e as barras de erro representam os desvios padrões, s_m . É possível perceber que a série é bastante estocástica, não existindo uma sazonalidade bem definida nos dados. As exceções são os meses de Agosto e Setembro, principalmente neste último cujos valores tem a menor variação. O que indica que os modelos periódicos, como o caso dos modelos Autorregressivos Periódicos (*Periodic Autoregressive Model* – Par(p)), podem não produzir melhorias significativas no desempenho de previsão.

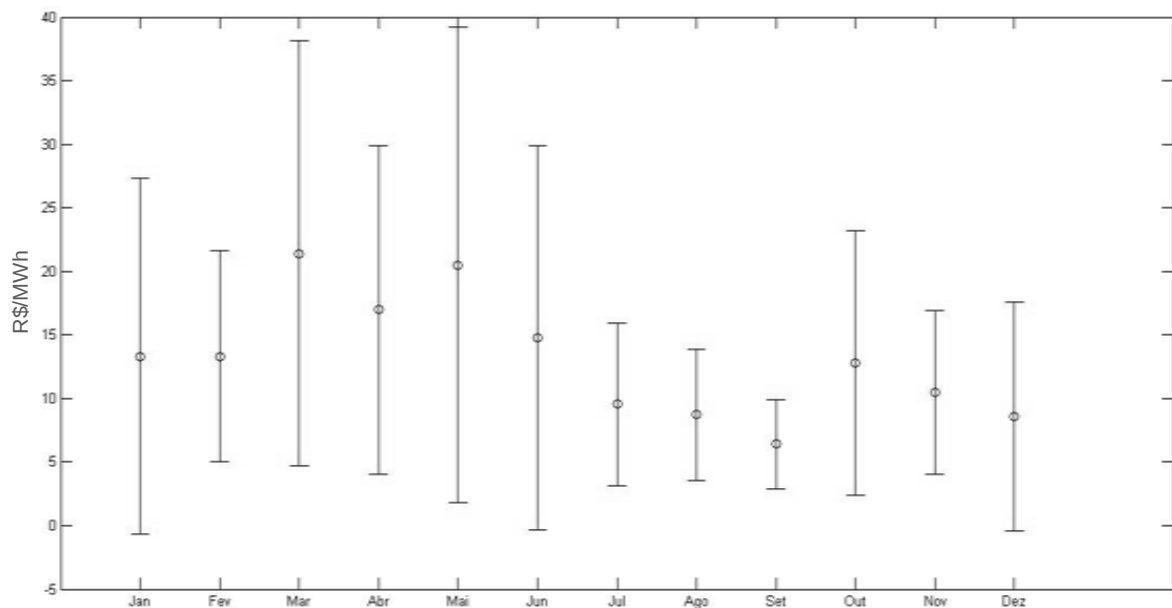


Figura 11 – Análise de Sazonalidade da série temporal *Spread*

2.2) Autocorrelação Parcial

A Função de Autocorrelação Parcial (FACP) tem um papel importante na determinação da ordem de modelos de previsão Autorregressivos, já que indica a integração linear entre duas amostras de uma série temporal defasadas no tempo, sem que os tempos intermediários influem em seu valor.

Existem diversos algoritmos para a estimação dos valores de Autocorrelação Parcial, como os propostos por (Box, Jenkins, and Reinsel 1970) e também por (Brockwell, 1991).

Uma das formas de se estimar os coeficientes de autocorrelação parcial utiliza as seguintes funções de autocorrelação e de autocorrelação parcial:

$$r_m = \frac{\sum_{i=1}^{n-m} (x_i - \bar{x})(x_{i+m} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3)$$

$$\Phi_{m,m} = \frac{r_m - \sum_{j=1}^{m-1} \Phi_{m-1,j} r_{m-1}}{1 - \sum_{j=1}^{m-1} \Phi_{m-1,j} r_j} \quad (4)$$

sendo, para atrasos $m = 1, 2, 3, 4, \dots, n/4$:

$\bar{x}$	Média;
x_i	i-ésimo valor da série;
n	Número de observações;
r_m	Coeficiente de Autocorrelação;
$\Phi_{m,m}$	Coeficiente de Autocorrelação Parcial.

A análise do resultado da Função de Autocorrelação Parcial deve ser feita através da comparação entre os coeficientes obtidos para cada atraso na série temporal e dos limites inferiores e superiores da região crítica expressa por:

$$r = \pm \frac{1,96}{\sqrt{n}} \quad (5)$$

sendo:

r	Limites da região crítica;
n	Número de amostras da série temporal.

Como no caso da série histórica do *Spread* existem sessenta e quatro meses, se resulta em uma região cujos limites $r = 0,245$. A Figura 12 exhibe os resultados da Função de Autocorrelação Parcial, incluindo os limites críticos exibidos nas linhas horizontais.

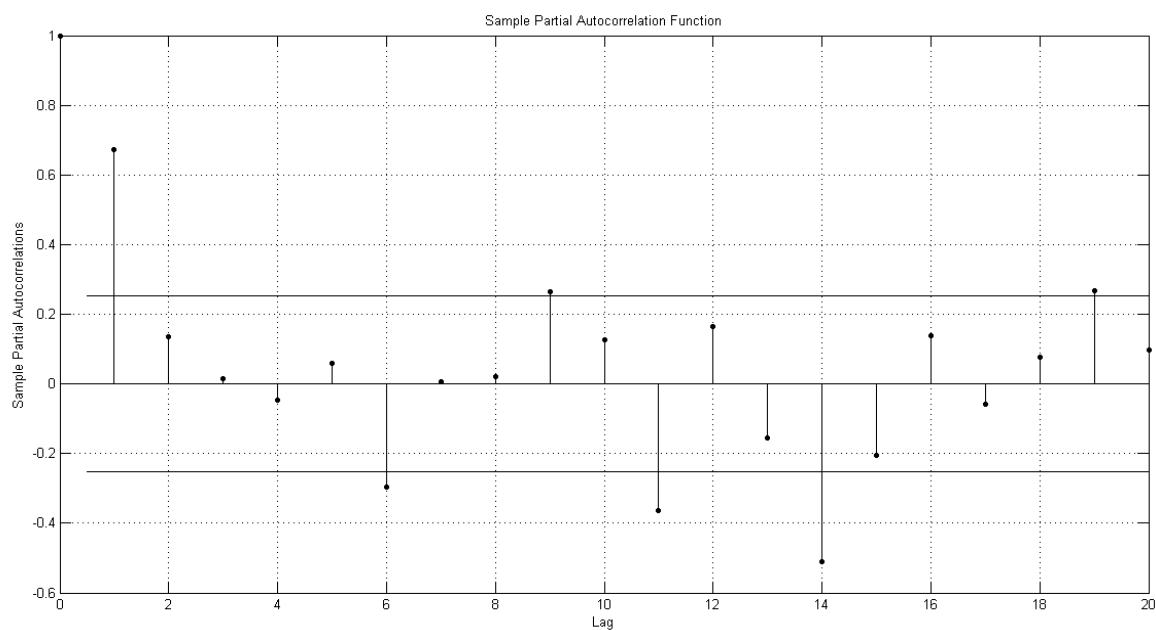


Figura 12 – Coeficientes de Autocorrelação Parcial da série *Spread*

É possível observar que existe uma forte interação entre o *Spread* do mês anterior e o *Spread* atual e que as outras ordens têm valores bem inferiores, indicando o uso de modelo Autorregressivo de ordem $p = 1$, o que será aplicado no Capítulo 4.

2.3) *Análise do Componente Principal*

O método de Análise do Componente Principal (Pearson, 1901) conhecido também por PCA (*Principal Component Analysis*, em inglês) consiste em utilizar transformações ortogonais para conversão de variáveis aparentemente relacionadas em um conjunto de variáveis não relacionadas entre si em um novo sistema de coordenadas, visando a redução do número de elementos a serem estudados, diminuindo assim a complexidade do caso.

A este novo conjunto de variáveis é dado o nome de componentes principais e apesar de seus valores não terem um significado a um primeiro olhar, eles explicam a variância das séries de variáveis utilizadas em sua modelagem.

A maior variância entre todos os componentes ocorre na direção da primeira coordenada do novo sistema, a segunda maior variância ocorre na direção da segunda coordenada e assim por diante. A quantidade de componentes após a transformação ortogonal é igual ao de variáveis inicial, mas devido aos primeiros componentes explicarem a maior parte da variância, os últimos componentes podem ser desconsiderados.

A Figura 13 exibe um gráfico relacionando duas variáveis, x_1 e x_2 , e o resultado da PCA aplicada ao caso, nas novas coordenadas z_1 e z_2 , demonstrando claramente que a maior variância reside ao longo da primeira coordenada, z_1 , do novo sistema.

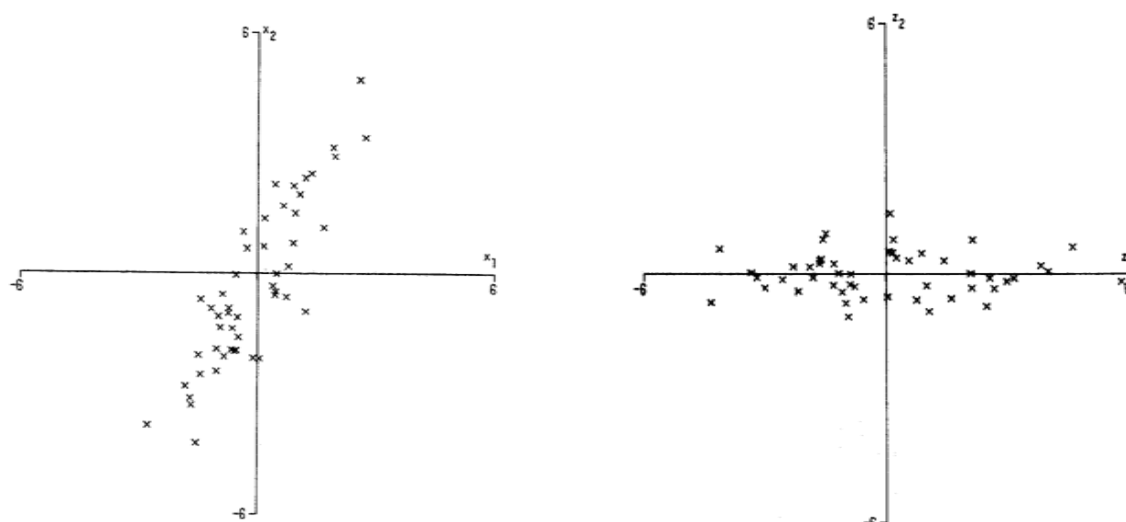


Figura 13 – Aplicação de PCA em duas variáveis correlacionadas (Sacchi, 2009)

A PCA foi aplicada no caso deste estudo utilizando como variáveis de entrada partes das seis séries temporais exibidas neste capítulo. A partir da série do *Spread* no mês t , foram utilizadas as sete séries defasadas, incluindo o *Spread*, em um, três, seis e doze meses, totalizando vinte e oito séries atrasadas (ou seja, vinte e oito variáveis) em relação à série cuja previsão será implementada. A discretização de três em três meses dos dados se deve à tentativa de se obter um *time delay* que garanta que os dados sejam independentes entre si, mas não ao ponto não terem nenhuma relação (Sacchi, 2009).

Ao separar as séries em partes distintas, o número de elementos em cada série passou a ser quarenta e oito, doze a menos que o conjunto de dados inicial já desconsiderados os meses de 2011, que serão utilizados para a validação das previsões, e os doze primeiros meses, devido às séries com defasagem de doze meses em relação ao *Spread*.

A Figura 14 exibe o resultado da PCA, indicando nas barras a percentagem da variância explicada por cada componente principal e a percentagem da variância acumulada até cada componente na linha. Os pontos destacados indicam a percentagem acumulada de no mínimo 85%, 90% e 95%, que correspondem às quantidades de nove, onze e quinze componentes, respectivamente, a partir das vinte e oito variáveis originais.

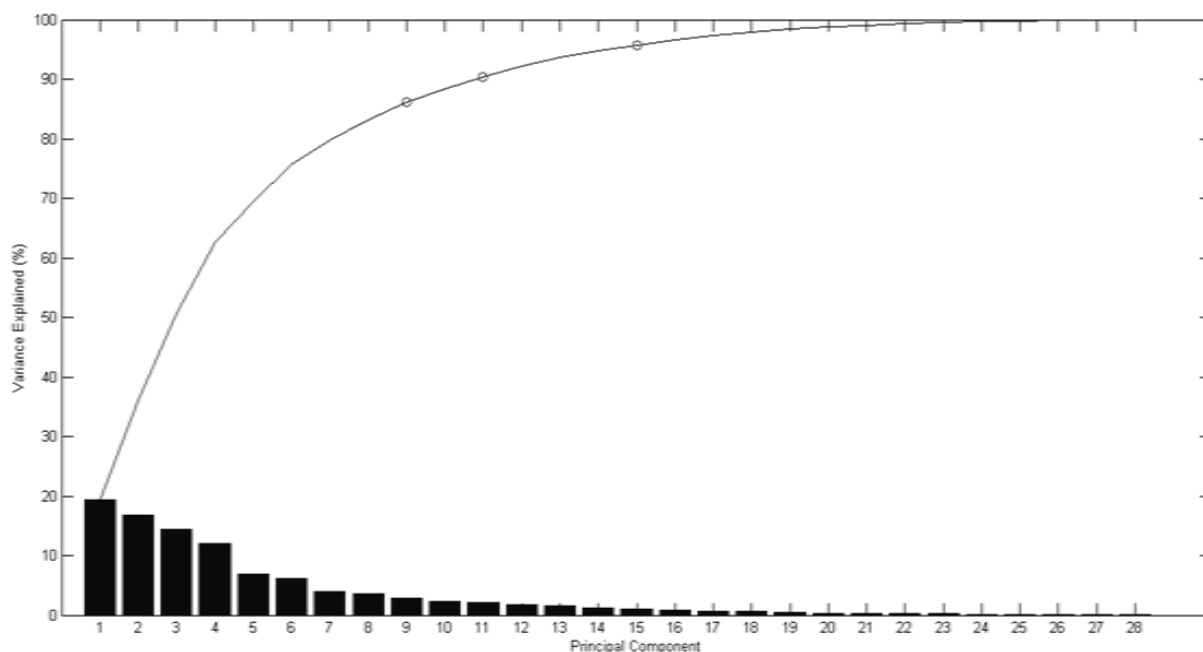


Figura 14 – Variância explicada pelos Componentes Principais da série *Spread*

O resultado da Análise dos Componentes Principais pode também ser analisado inversamente, a fim de se observar quais das variáveis originais mais influenciam nos componentes iniciais, ou seja, têm maior participação na variância geral do conjunto de dados.

Esta análise pode ser feita partindo do princípio que cada um dos componentes principais corresponde a uma dimensão e o seu valor corresponde à multiplicação das variáveis originais pelos coeficientes calculados. A Figura 15 exibe o valor destes coeficientes para os dois primeiros componentes e para os três primeiros componentes.

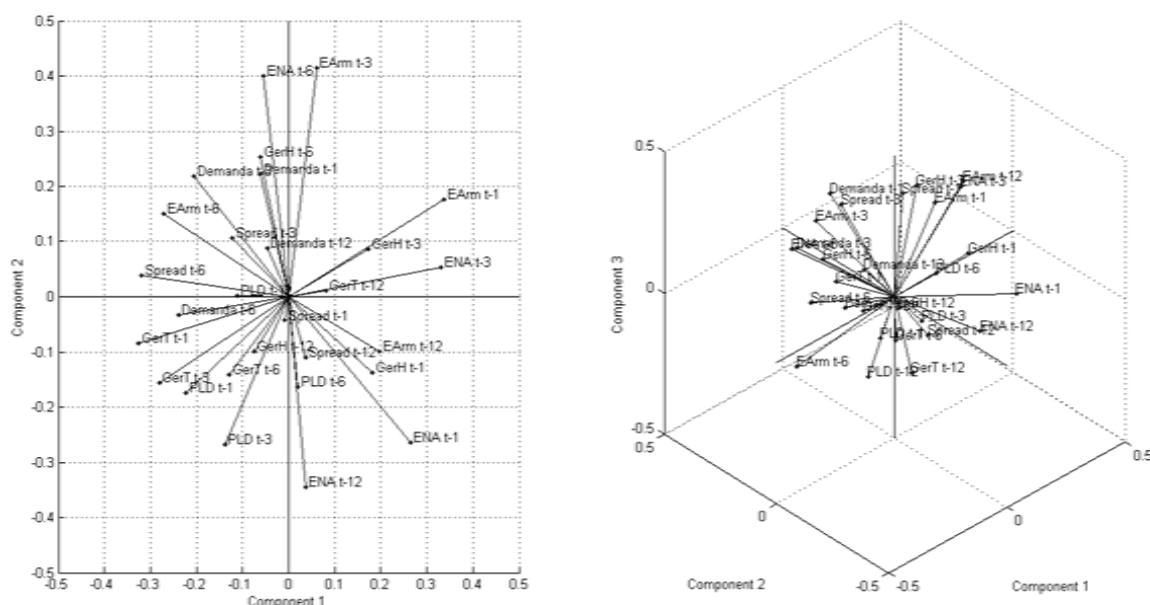


Figura 15 – Coeficientes dos primeiros componentes principais

Ao calcular a distância euclidiana da origem a cada um dos pontos que representam os coeficientes das variáveis originais, se obtém a contribuição desta variável às dimensões (componentes principais) consideradas. Caso todas as dimensões sejam consideradas, vinte e oito dimensões no caso desta análise, o resultado para todas as variáveis corresponderia à cem por cento de contribuição. A Figura 16 exibe a contribuição de todas as vinte e oito variáveis para as quantidades de componentes principais que correspondem a 85%, 90% e 95%, organizadas pela contribuição média de cada uma.

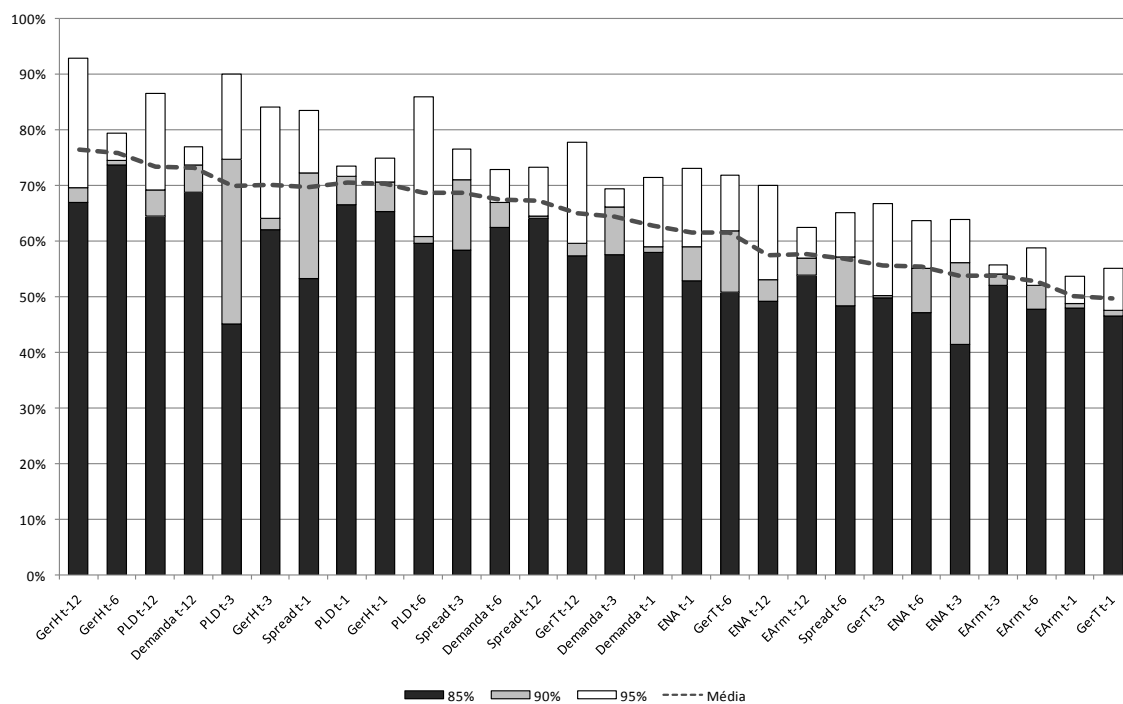


Figura 16 – Participação das variáveis originais nos componentes principais

Capítulo 3

Modelos Lineares

Os Modelos Lineares são de uma arquitetura simples, mas extremamente poderosos na predição de séries temporais. Neste trabalho foram explorados desde os modelos mais simples como o modelo *Naive* (Ingênuo) até modelos complexos como o modelo Autorregressivo Exógeno (Autoregressive Exogenous Model - ARX).

Tais modelos utilizam a premissa de que as séries são derivadas de relações lineares entre variáveis, sejam estas variáveis os próprios valores da própria série temporal, ou valores exógenos, como no caso do modelos Autorregressivo Exógeno (ARX).

Assim sendo, as saídas dos modelos são definidas como combinações de equações lineares que seguem a seguinte forma:

$$y_t = c + \sum_{p=1}^P (a_p y_{t-p}) + \sum_{n=1}^N (b_n x_n) + e \quad (6)$$

sendo, para cada previsão para o instante t:

y_t	Valor da série no instante t;
P	Ordem do modelo;
a_p	Coefficiente de ordem p;
N	Número de variáveis exógenas;
b_n	Coefficiente da n-ésima variável;
x_n	n-ésima Variável;
c	Termo constante;
e	Erro do modelo.

3.1) Modelo Naive (Ingênuo)

Os modelo *Naive*, ou modelo Ingênuo, em sua tradução ao português, é extremamente simples, sendo utilizado amplamente como comparação a modelos mais sofisticados ou quando se tem escassas informações sobre a série a ser prevista.

A previsão *Naive* define como sendo o próximo estado da variável o estado atual, sendo possível representá-la através da equação:

$$y_t = y_{t-1} \quad (7)$$

A utilização desse modelo é bastante direta e rápida, já que este tem ordem um e seu coeficiente de primeira ordem é sempre igual a um, ou seja, como será percebido adiante, é um caso simples do modelo Autorregressivo AR(1). A Figura 17 exibe em linha tracejada vermelha a sua aplicação na predição da série histórica do *Spread*, representada em linha cheia azul.

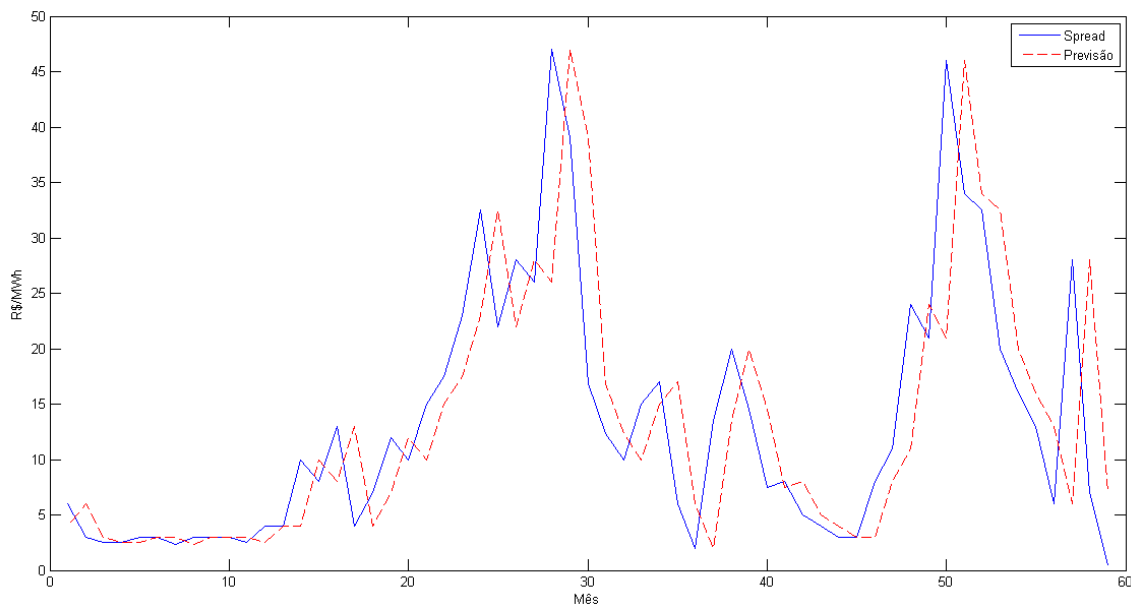


Figura 17 – *Spread* e previsão do Modelo *Naive*

Os modelos deste trabalho serão avaliados segundo três estimadores de erros, o Erro Quadrático Médio (Mean Squared Error - MSE), o Erro Absoluto Médio (Mean Absolute Error – MAE) e o Erro Absoluto Percentual Médio (Mean Absolute Percentual Error – MAPE), que são apresentados no Apêndice A. A Tabela 1 apresenta o resultado dos cálculos destes três estimadores de erro para a previsão resultante do modelo *Naive*.

Tabela 1 – Erros de ajuste do modelo *Naive*

MSE	70,14
MAE	5,71
MAPE	69,98%

3.2) Modelos Autorregressivos

Os Modelos Autorregressivos utilizam como variáveis de entrada apenas valores da própria série a ser prevista, excluindo assim o segundo somatório da equação (6), sendo portanto os modelos lineares mais simples.

Normalmente estes modelos são expressos como $AR(p)$, onde p indica a ordem do modelo, ou seja, qual o número de atrasos que o modelo leva em conta em sua equação.

A obtenção de seus coeficientes pode ser feita de diversas maneiras, seja utilizando as equações de Yule-Walker (Box, Jenkins, and Reinsel 1970) , que consiste na solução da seguinte equação matricial, considerando o termo constante como nulo:

$$\begin{pmatrix} 1 & r_1 & r_2 & r_3 & r_4 & \dots & r_{p-1} \\ r_1 & 1 & r_1 & r_2 & r_3 & \dots & r_{p-2} \\ r_2 & r_1 & 1 & r_1 & r_2 & \dots & r_{p-3} \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ r_{p-1} & r_{p-2} & r_{p-3} & r_{p-4} & r_{p-5} & \dots & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \cdot \\ \cdot \\ \cdot \\ a_p \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \cdot \\ \cdot \\ \cdot \\ r_p \end{pmatrix} \quad (8)$$

sendo:

r_n Coeficiente de correlação com atraso n , já definido pela equação (3);

a_p Coeficiente de ordem p .

Outra abordagem para a estimação dos coeficientes pode ser feita utilizando-se o método dos mínimos quadrados, que consiste na solução das seguintes equações matriciais:

$$X = \begin{pmatrix} 1 & y_{t-1} & y_{t-2} & \cdots & y_{t-P} \\ 1 & y_{t-2} & y_{t-3} & \cdots & y_{t-(P-1)} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & y_{t-n} & y_{t-(n+1)} & \cdots & y_{t-n-(P-1)} \end{pmatrix} \quad (9)$$

$$\begin{pmatrix} c \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_P \end{pmatrix} = (X'X)^{-1} X'y \quad (10)$$

sendo:

- y_t Valor da série y no instante t ;
- P Ordem do modelo;
- a_p Coeficiente de ordem p ;
- n Número de observações da série y .

3.2.1) Modelo Autorregressivo de Primeira Ordem

Também conhecido como modelo Markoviano, o Modelo Autorregressivo de Primeira Ordem, AR(1), é um modelo bastante simples, cuja equação final pode ser representada por:

$$y_t = a_1 y_{t-1} \quad (3.1)$$

A Função de Autocorrelação Parcial calculada para a série histórica do *Spread* no Capítulo 2 resultou em um modelo de primeira ordem, portanto teoricamente este é o modelo ideal para a sua previsão.

Ao aplicar o AR(1) na série histórica do *Spread*, utilizando o método Yule-Walker, o coeficiente α_1 resultante foi de 0,8838. A Figura 18 exibe o resultado da previsão pelo modelo em linha vermelha tracejada e a série temporal do *Spread* em linha azul contínua.

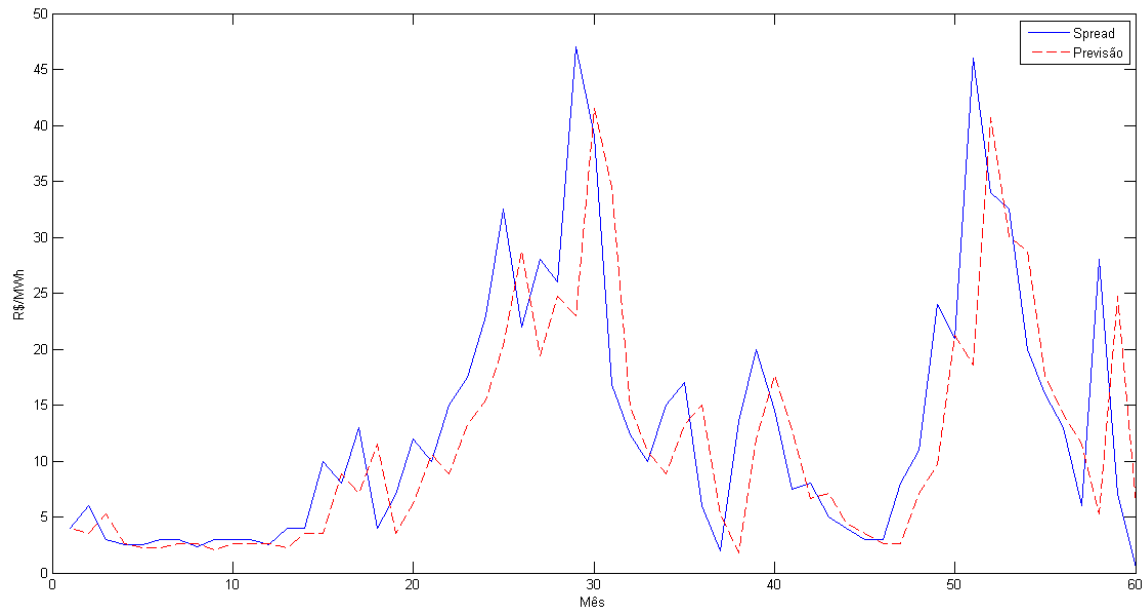


Figura 18 – *Spread* e previsão do modelo AR(1)_{Yule-Walker}

A Tabela 2 exibe o resultado para os três estimadores de erro.

Tabela 2 – Erros de ajuste do modelo AR(1)_{Yule-Walker}

MSE	66,05
MAE	5,33
MAPE	62,47%

Ao utilizar o método dos mínimos quadrados para a estimação da série, foram obtidos os valores de -0,1421 e 1,0062 para c e a_1 , respectivamente. O resultado pode ser observado na Figura 19.

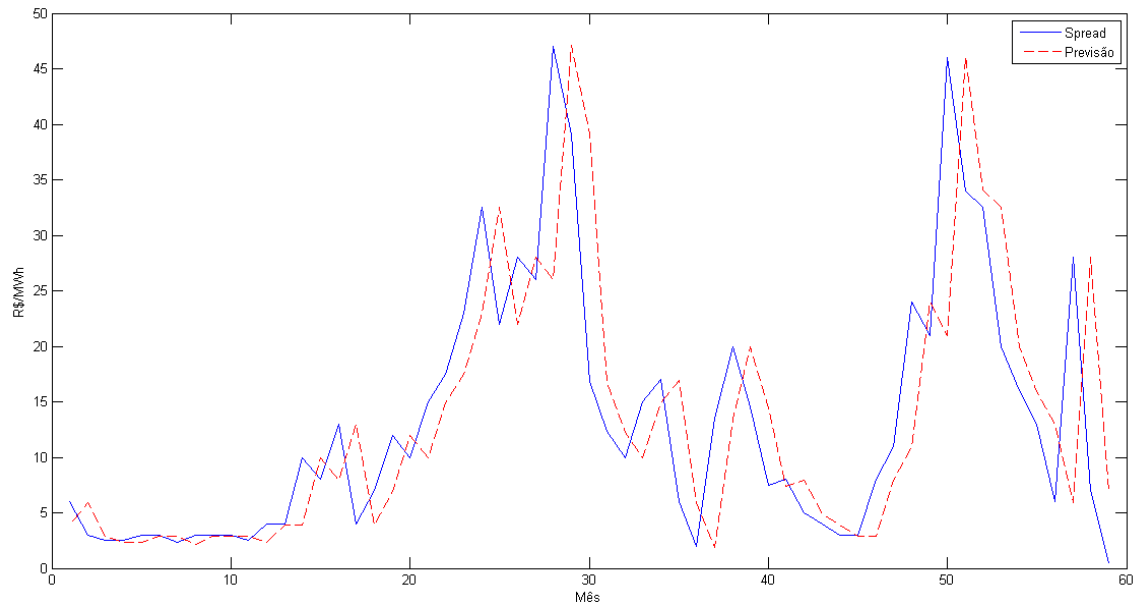


Figura 19 - Spread e previsão do modelo AR(1) Mínimos Quadrados

A Tabela 3 exibe o resultado do estudo de erros, onde é possível perceber a superioridade do método Yule-Walker.

Tabela 3 – Erros de ajuste do modelo AR(1) Mínimos Quadrados

MSE	70,57
MAE	5,74
MAPE	69,87%

3.2.2) Modelo Autorregressivo de Segunda Ordem

Apesar de a Função de Autocorrelação Parcial indicar o uso de um modelo de primeira ordem, devido a simplicidade dos modelos AR(p), decidiu-se aplicar a previsão utilizando o modelo AR(2), já que algumas ordens superiores apresentaram índices fora da região de confiança. A equação do modelo AR(2) segue como:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} \quad (11)$$

Os coeficientes a_1 e a_2 encontrados pelo método Yule-Walker foram respectivamente 0,7588 e 0,1422. Claramente o coeficiente de segunda ordem tem um peso menor na previsão, o que era esperado devido à sua baixa autocorrelação parcial.

A Figura 20 exibe o resultado da previsão utilizando o modelo AR(2) em tracejado vermelho, e a compara com o *Spread* real, em linha azul.

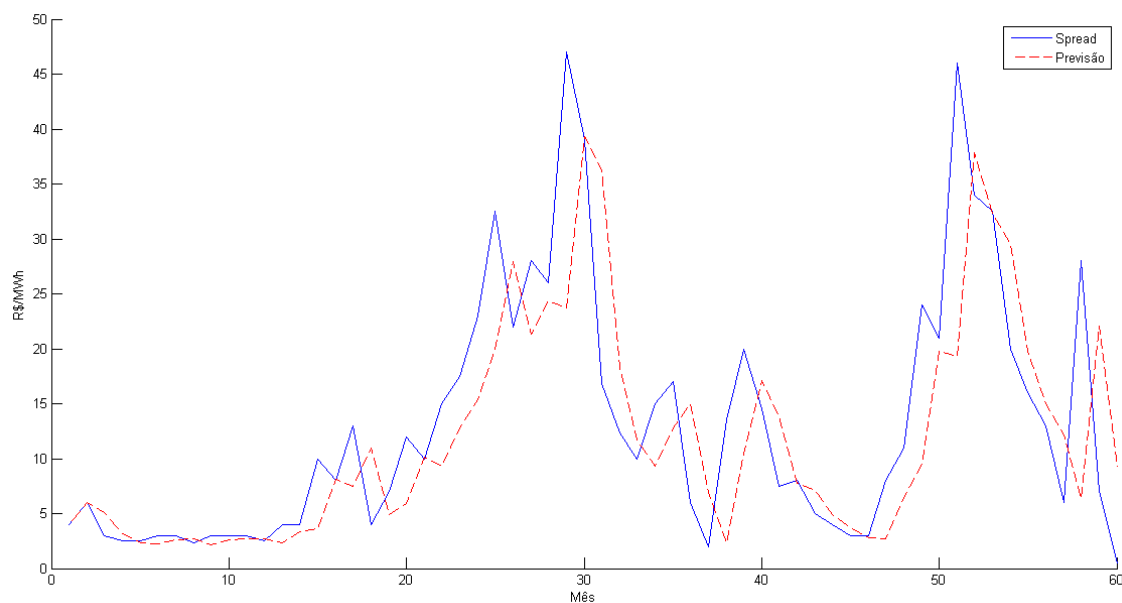


Figura 20 – *Spread* e previsão do modelo AR(2)

A Tabela 4 exibe o resultado da análise de erros da previsão do modelo AR(2).

Tabela 4 – Erros de ajuste do modelo AR(2)

MSE	65,73
MAE	5,38
MAPE	74,82%

3.3) *Modelo Autorregressivo Exógeno*

O Modelo Autorregressivo Exógeno, ou simplesmente ARX, utiliza premissas semelhantes aos Modelos AR(p), adicionando a estes modelos variáveis exógenas, independentes da série temporal estudada.

Para o caso estudado, foram definidas seis variáveis exógenas, já apresentadas no Capítulo 2 deste trabalho, além do histórico do próprio *Spread*, defasadas em quatro tempos distintos, resultando em vinte e oito séries de variáveis.

O modelo ARX de primeira ordem foi então implementado utilizando o método dos mínimos quadrados para a obtenção dos parâmetros e tomando-se a constante c como nula, para a simplificação do modelo. De uma forma geral, os resultados serão aplicados à seguinte equação, também derivada da Equação (6):

$$y_t = a_1 y_{t-1} + \sum_{n=1}^{24} (b_n x_n) \quad (12)$$

3.3.1) Modelo ARX de primeira ordem

Para uma primeira análise, o modelo utilizou como variáveis exógenas apenas os estados defasados em um mês. A Tabela 5 exibe os coeficientes de cada variável para o modelo, onde percebe-se uma grande dependência do *Spread* do mês anterior, e a Figura 21 exibe o resultado gráfico desta análise, com a previsão em vermelho tracejado e o *Spread* real em linha contínua azul.

Tabela 5 – Coeficientes do modelo ARX de primeira ordem

COEFICIENTE	VARIÁVEL	VALOR
a_1	$Spread_{t-1}$	0,5544
b_1	PLD _{t-1}	0,2929 x 10 ⁻³
b_2	Demanda _{t-1}	-0,4608 x 10 ⁻³
b_3	Geração Hidrelétrica _{t-1}	1,4652 x 10 ⁻³
b_4	Geração Térmica _{t-1}	0,0968 x 10 ⁻³
b_5	ENA _{t-1}	0,0804 x 10 ⁻³
b_6	Energia Armazenada	-0,0586 x 10 ⁻³

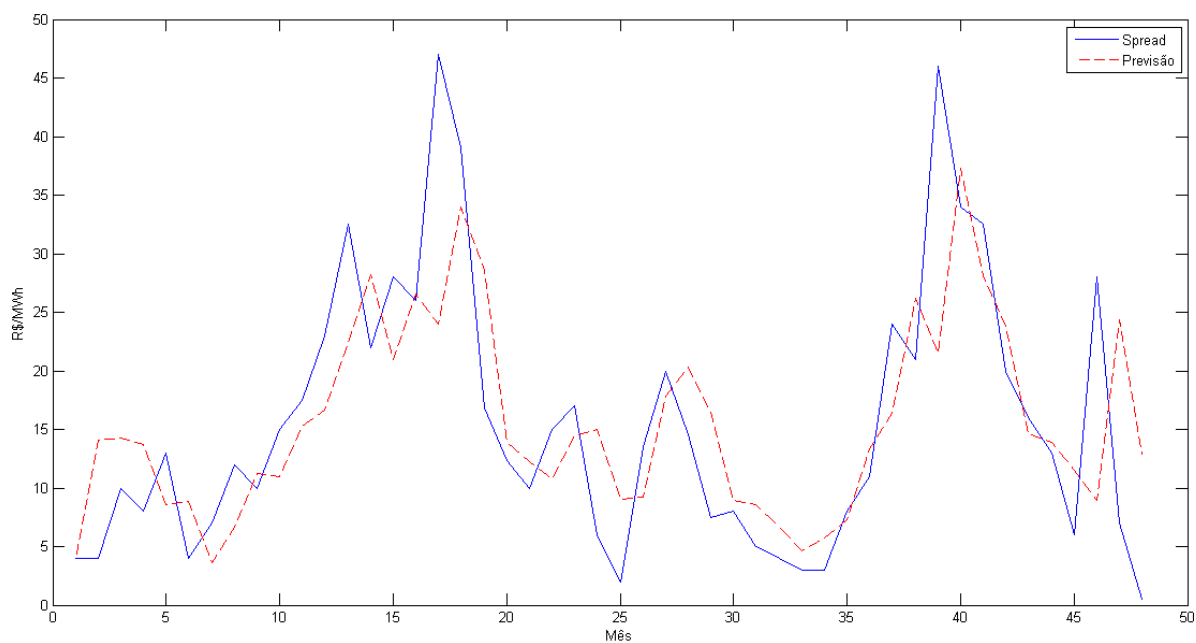


Figura 21 – *Spread* e previsão do modelos ARX de primeira ordem

A Tabela 6 exibe o resultado da análise de erros da previsão do modelo AR(2).

Tabela 6 – Erros de ajuste do modelo ARX de primeira ordem

MSE	66,05
MAE	6,04
MAPE	108,02%

3.3.2) Modelo ARX utilizando PCA

A próxima análise utilizou a Análise dos Componentes Principais, também apresentada no Capítulo 2, visando a redução do número de elementos e também a redução de uma possível redundância entre as variáveis utilizadas.

A análise foi feita quatro vezes, utilizando as quantidades de componentes principais que explicam um mínimo de 85%, 90%, 95% e 100% da variância das vinte e oito séries. A Figura 22 exibe o resultado gráfico das quatro análises.

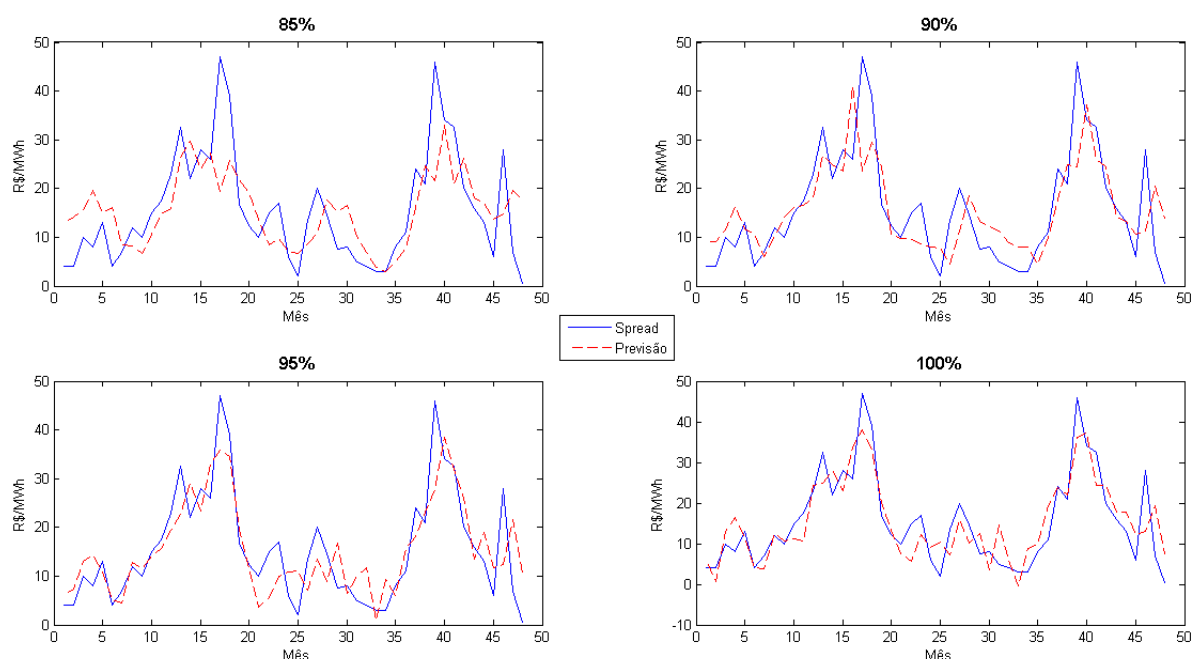


Figura 22 – Spread e previsão dos modelos ARX utilizando resultados da PCA

A Tabela 7 exibe uma compilação dos resultados das quatro análises sobre o modelo Autorregressivo Exógeno. É possível perceber que os valores dos coeficientes independem da análise, sendo os mesmos para qualquer uma das quatro análises e apenas sendo nulo caso o componente a que se refere não faça parte da análise.

Tabela 7 – Coeficientes dos modelos ARX utilizando resultados da PCA

<i>Análise</i>	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>
Componentes Principais	9	11	15	28
% da Variância Explicada	85%	90%	95%	100%
MSE	77,455	61,4365	45,0903	35,2356
MAE	6,8427	5,9839	5,4585	4,9374
MAPE	132,83%	111,87%	98,11%	81,24%
b_1	1,5889	1,5889	1,5889	-0,4742
b_2	0,9866	0,9866	0,9866	0,9866
b_3	-3,0621	-3,0621	-3,0621	-3,0621
b_4	1,0972	1,0972	1,0972	1,0972
b_5	-0,2928	-0,2928	-0,2928	-0,2928
b_6	-0,3956	-0,3956	-0,3956	-0,3956
b_7	-0,8972	-0,8972	-0,8972	-0,8972
b_8	1,1233	1,1233	1,1233	1,1233
b_9	-1,1672	-1,1672	-1,1672	-1,1672
b_{10}	-	3,5102	3,5102	3,5102
b_{11}	-	-3,8393	-3,8393	-3,8393
b_{12}	-	-	0,6208	0,6208
b_{13}	-	-	-0,0661	-0,0661
b_{14}	-	-	6,7282	6,7282
b_{15}	-	-	-3,0646	-3,0646
b_{16}	-	-	-	-2,8785
b_{17}	-	-	-	-0,4177
b_{18}	-	-	-	0,2388
b_{19}	-	-	-	1,283
b_{20}	-	-	-	-3,4948
b_{21}	-	-	-	2,1339
b_{22}	-	-	-	-0,0399
b_{23}	-	-	-	0,3898
b_{24}	-	-	-	7,3853
b_{25}	-	-	-	-5,1613
b_{26}	-	-	-	-4,7851
b_{27}	-	-	-	-1,6193
b_{28}	-	-	-	-19,275

Ao observar a análise de erro, se percebe, como esperado, que quanto maior a quantidade dos componentes principais, menores os erros do modelo.

Capítulo 4

Redes Neurais Artificiais

As RNAs são modelos computacionais inspirados pelos neurônios, possuindo a capacidade de aprender e reter conhecimento através de um conjunto de unidades de processamento que podem ser caracterizadas como neurônios artificiais interconectados por sinapses representadas por vetores ou matrizes de pesos sinápticos (Silva, 2010).

O treinamento de uma RNA depende de um conjunto de dados de entradas e saídas esperadas que são apresentados para a rede, para que os pesos sinápticos sejam recalculados a cada etapa do treinamento da rede, diminuindo o erro geral entre suas respostas e as saídas esperadas. A Figura 23 exibe um diagrama que exemplifica o funcionamento de um neurônio artificial.

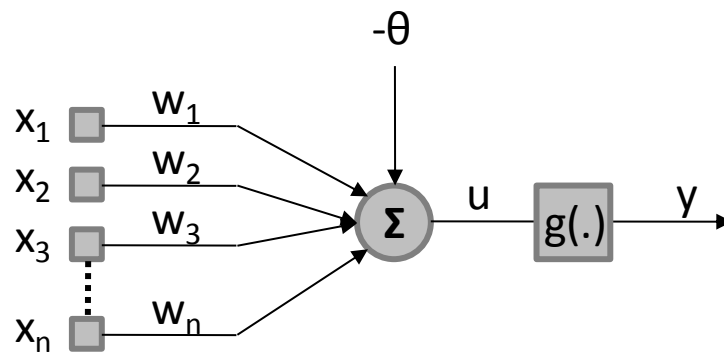


Figura 23 – Neurônio Artificial de uma RNA

As expressões que representam o funcionamento de um neurônio artificial como o exemplificado acima são:

$$u = \sum_{i=1}^n x_i w_i - \theta \quad (13)$$

$$y = g(u) \quad (14)$$

sendo:

- x_i Sinais de entrada;
- w_i Pesos sinápticos;
- θ Limiar de ativação;

- u Potencial de ativação;
 $g(.)$ Função de ativação;
 y Sinal de saída do neurônio;

Os sinais de entrada são os valores a serem avaliados pelo neurônio, através dos pesos sinápticos associados a cada uma das conexões de entrada e do limiar de ativação, este último utilizado para especificar o nível do potencial de ativação que ao passar pela função de ativação poderá ser disparado como saída do neurônio.

As funções de ativação são bastante diversas, sendo que algumas das mais comuns são exemplificadas na Figura 24.

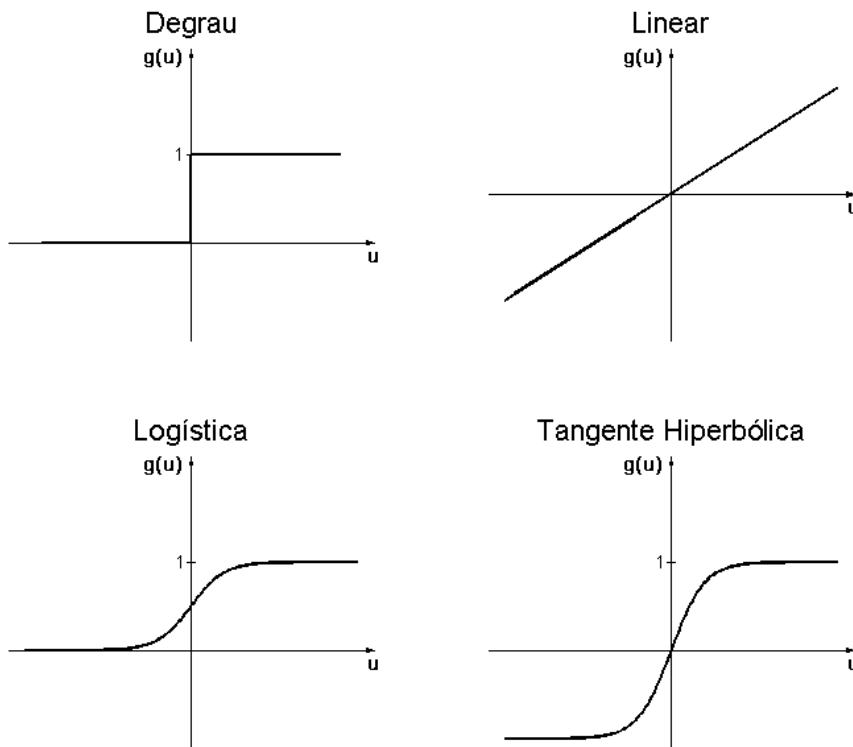


Figura 24 – Exemplos de Funções de Ativação

As equações que representam cada uma das funções de ativação citadas acima seguem:

$$\text{Degrau} \quad g(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u < 0 \end{cases} \quad (15)$$

$$\text{Linear} \quad g(u) = u \quad (16)$$

$$\text{Logística} \quad g(u) = \frac{1}{1 + e^{-\beta u}} \quad (17)$$

$$\text{Tangente Hiperbólica} \quad g(u) = \frac{1 - e^{-\beta u}}{1 + e^{-\beta u}} \quad (18)$$

sendo:

- e Exponencial;
- β Parâmetro de inclinação, tomado como valor unitário nos exemplos.

As RNAs podem ter diversas arquiteturas, desde simples como a rede *Perceptron*, que utiliza apenas um único neurônio artificial, até a complexas estruturas como as Redes *ART* (*Adaptive Resonance Theory*). Cada arquitetura é indicada para a solução de um determinado tipo de problema, sendo que no caso de previsões de séries temporais, a arquitetura mais clássica, que será utilizada neste trabalho, é a Rede *Perceptron* Multicamadas (PMC ou *MLP – Multilayer Perceptron*).

4.1) Rede *Perceptron* Multicamadas

As PMCs são uma das arquiteturas mais versáteis quanto à aplicabilidade, tendo bom desempenho em áreas como as de aproximação de funções, reconhecimento de padrões, otimizações de sistemas e previsão de séries temporais, sendo esta última propriedade a que chamou atenção para sua utilização neste trabalho.

Esta arquitetura consiste em uma primeira camada de neural de entrada, uma ou mais camadas de neurônios artificiais intermediária denominadas camadas escondidas, e uma camada de saída. A Figura 25 exibe uma arquitetura básica de uma PMC com duas camadas neurais escondidas.

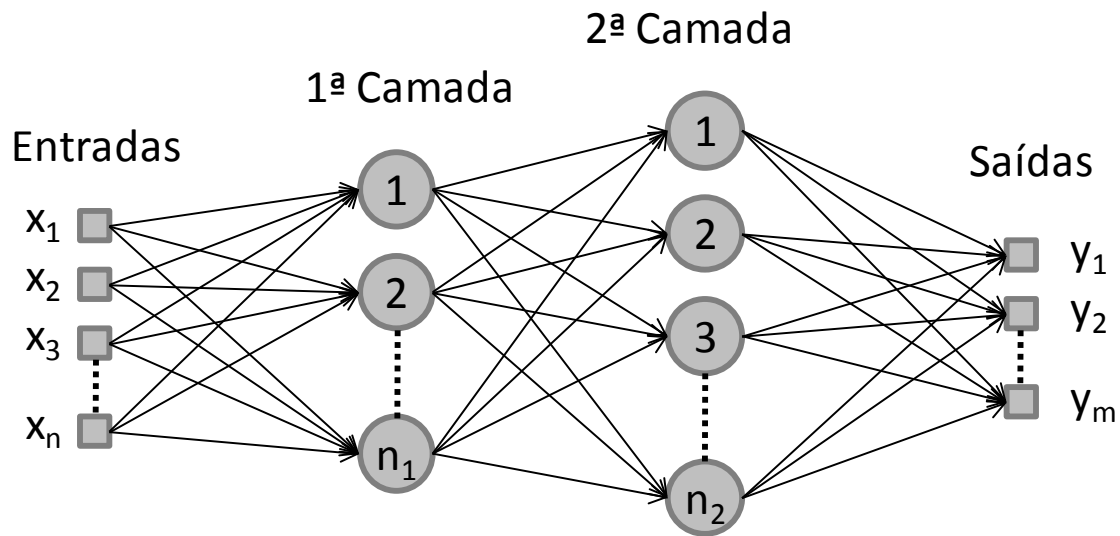


Figura 25 – Arquitetura básica de uma Rede Perceptron Multicamadas

Como já visto, cada interconexão tem um peso sináptico, definido randomicamente antes do treinamento. O objetivo do treinamento da PMC consiste em ajustar estes pesos afim de que o erro entre as respostas geradas pelas entradas e as respostas esperadas apresentadas à rede seja o menor possível. Esta arquitetura utiliza como treinamento o método de *back-propagation*, que consiste em duas etapas alternadas que são aplicadas até um resultado satisfatório.

Durante a primeira etapa, denominada *forward*, as entradas são inseridas na rede, sem qualquer ajuste nos pesos sinápticos, passando por todas as camadas escondidas até resultarem nas saídas. A segunda etapa, denominada *backward*, consiste na comparação das saídas obtidas com as saídas esperadas do sistema, que é então percorrido reversamente, aplicando correções nos pesos sinápticos de cada interconexão.

A aplicação sucessiva de ambas as etapas *forward* e *backward* acabam então por ajustar automaticamente todos os valores sinápticos a cada interação, através da diminuição dos erros das respostas da rede.

Normalmente a arquitetura da PMC é definida de forma empírica, já que seu desempenho depende de inúmeras variáveis. Para o problema da predição do *Spread* foram utilizadas redes de duas camadas escondidas com variações dos números de neurônios artificiais em cada uma, buscando a arquitetura de melhor desempenho geral, e funções de ativação logísticas para as camadas intermediárias e lineares para as camadas de saída.

4.1.1) Spread como entrada da PMC

Primeiramente, afim de se comparar as performances dos modelos $AR(p)$ discutidos no capítulo anterior, foi utilizado como entrada da RNA, o *Spread* de um mês anterior, resultando em uma arquitetura de uma entrada e uma saída, o *Spread* esperado. A partir daí, foram testadas diversas configurações para as duas camadas escondidas, sendo a de melhor desempenho com vinte neurônios em ambas. A Figura 26 exibe dois dos melhores resultados obtidos nesta configuração.

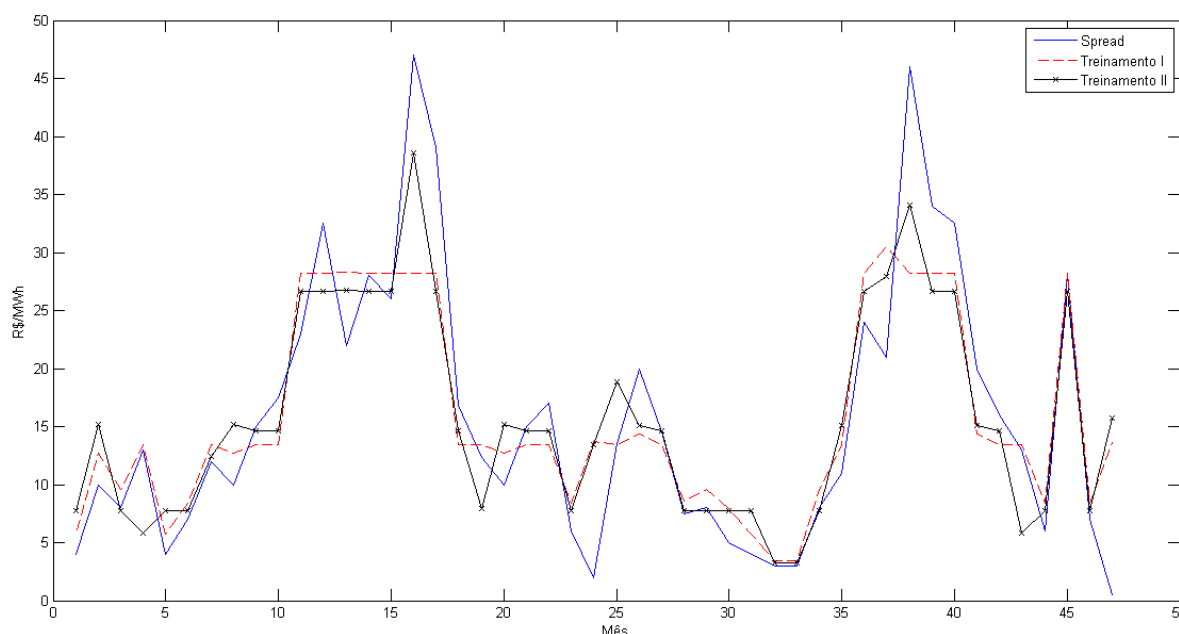


Figura 26 – Resultados das PMC com entrada de $Spread_{t-1}$

A Tabela 8 exibe os erros de alguns dos treinamentos executados na arquitetura implementada, sendo que os Treinamentos II e I correspondem aos resultados da Figura 26.

Tabela 8 - Erros de ajuste da PMC com entrada de $Spread_{t-1}$

	Treinamento I	Treinamento II	Treinamento III	Treinamento IV	Treinamento V
MSE	46,27	41,06	45,11	48,51	41,95
MAE	4,73	4,06	4,69	4,66	4,07
MAPE	75,56%	66,89%	69,43%	73,94%	63,15%

A seguir, foi utilizada uma arquitetura de duas entradas, sendo o *Spread* de um e dois meses atrás, e uma saída, o *Spread* esperado. A mesma configuração de vinte neurônios em cada uma das camadas escondidas foi escolhida como a de melhores resultados. A Figura 27 exibe os melhores resultados obtidos nesta configuração, sendo que o melhor, em linha pontilhada verde, deve ser observado com cautela, pois devido à sua grande exatidão, ele pode ter se “viciado” aos

valores do histórico, apenas se ajustando a estes como uma função, e não realmente prevendo os resultados.

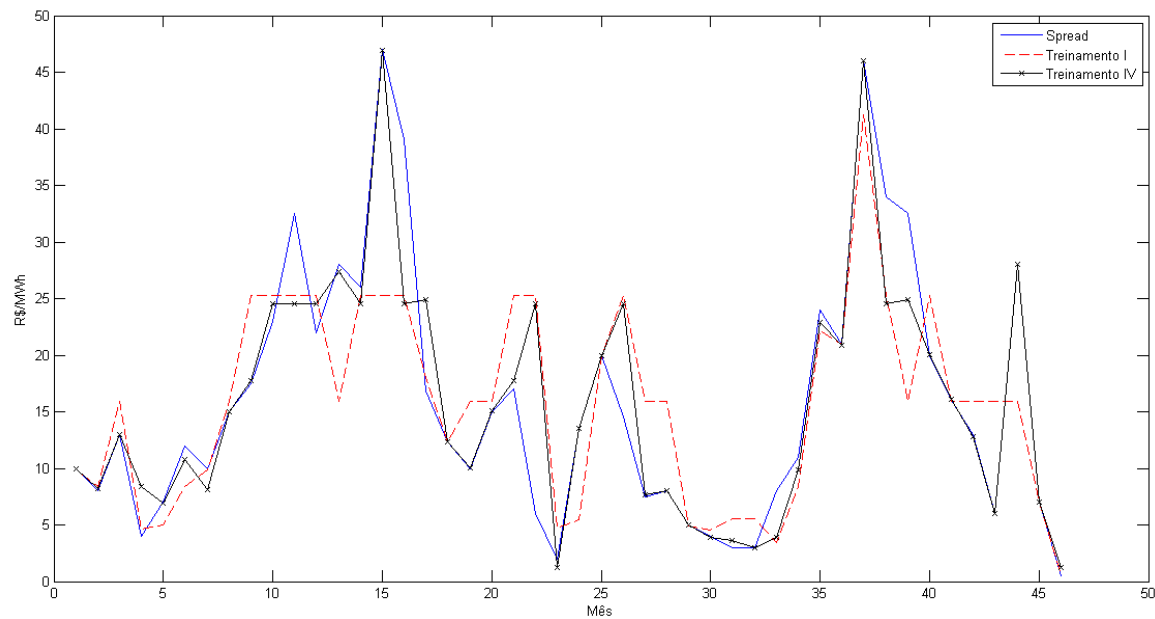


Figura 27 – Melhores resultados das PMC com entradas históricas do *Spread*

A Tabela 9 exibe os erros de alguns dos treinamentos executados na arquitetura implementada, sendo que os Treinamentos I e IV correspondem ao resultados da Figura 27.

Tabela 9 - Erros de ajuste da PMC com entradas históricas do *Spread*

	<i>Treinamento I</i>	<i>Treinamento II</i>	<i>Treinamento III</i>	<i>Treinamento IV</i>	<i>Treinamento V</i>
MSE	43,16	50,95	20,94	17,16	29,38
MAE	4,19	4,82	2,78	1,88	3,38
MAPE	41,67%	79,21%	44,29%	22,06%	37,16%

4.1.2) Variáveis defasadas de um mês como entradas da PMC

Com o intuito de se comparar a performance com o modelo ARX, também foi implementada uma PMC com as sete variáveis apresentadas no Capítulo 2 defasadas em um mês como entrada.

A arquitetura escolhida para esta análise foi a de dez neurônios em cada camada escondida e os resultados gráficos podem ser visualizados na Figura 28.

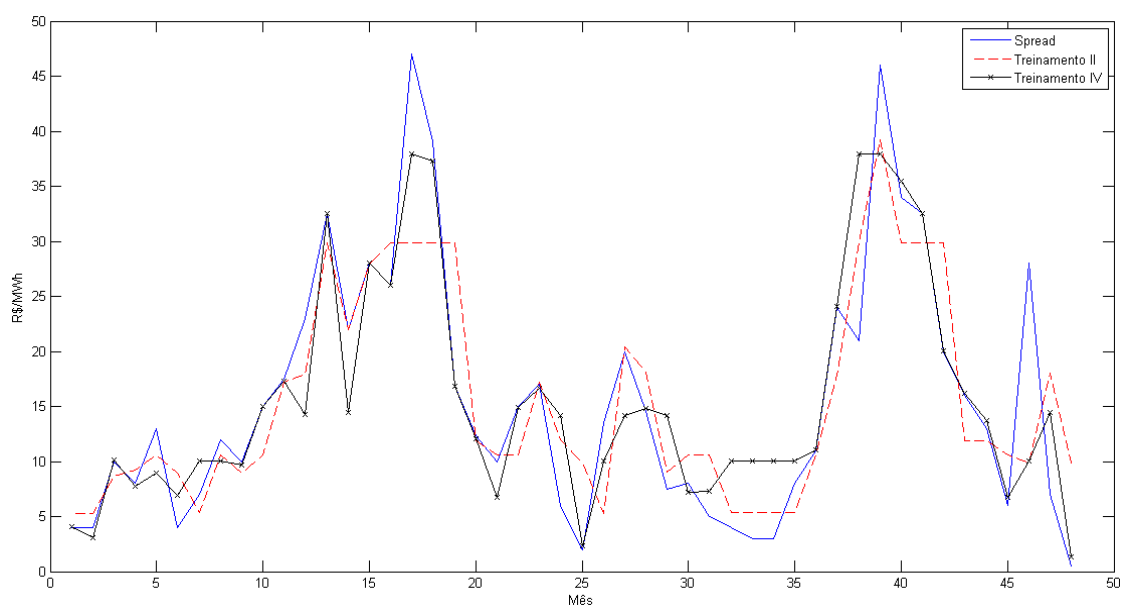


Figura 28 – Resultados da PMC com variáveis do sistema

A Tabela 10 abaixo exhibe os erros para alguns dos treinamentos realizados nesta arquitetura.

Tabela 10 - Erros de ajuste da PMC com variáveis do sistema

	<i>Treinamento I</i>	<i>Treinamento II</i>	<i>Treinamento III</i>	<i>Treinamento IV</i>	<i>Treinamento V</i>
MSE	27,86	37,05	64,20	27,22	43,91
MAE	3,78	4,36	5,88	3,12	4,51
MAPE	56,55%	80,59%	109,36%	37,42%	62,70%

4.1.3) PCA como entradas da PMC

Finalmente, como uma última análise, foram utilizadas os resultados da Análise dos Componentes Principais como entradas para a rede *Perceptron* Multicamadas, utilizando uma arquitetura de dez neurônios em ambas as camadas escondidas. Durante a análise, percebeu-se

que a rede neural tende ao *overfitting*, ou seja, a PMC simplesmente se adapta ao conjunto de treinamento, não generalizando as previsões, que acabam por ter altos erros quando apresentadas a conjuntos de dados que não foram utilizados para treinamento.

O *overfitting* de algumas das redes pode ser observado nos resultados gráficos abaixo, onde foram utilizados como entradas os componentes principais que correspondem às percentagens de 85, 90, 95 e 100 da explicação da variância observados na Figura 29, Figura 30, Figura 31 e Figura 32, respectivamente.

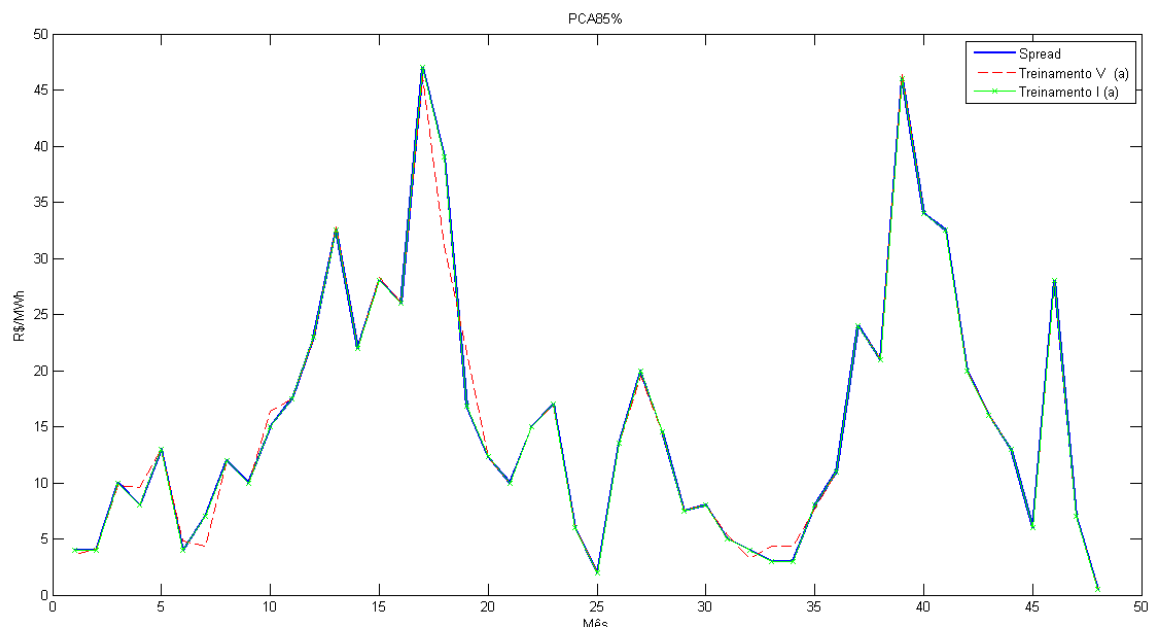


Figura 29 – Resultados da PMC com componentes principais de 85% da variância

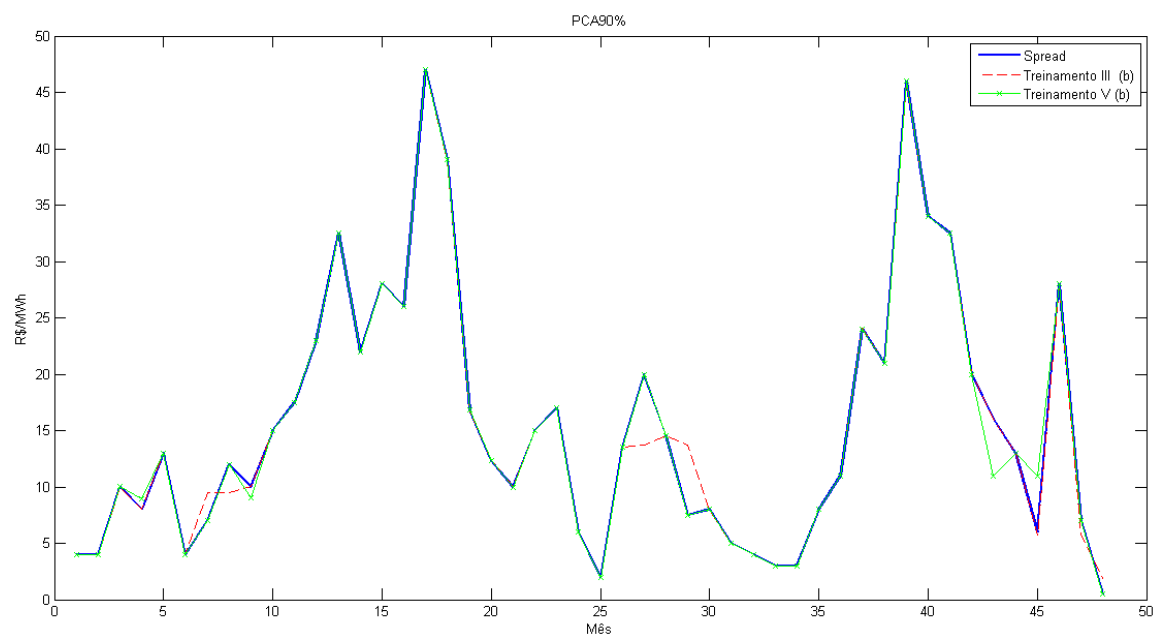


Figura 30 – Resultados da PMC com componentes principais de 90% da variância

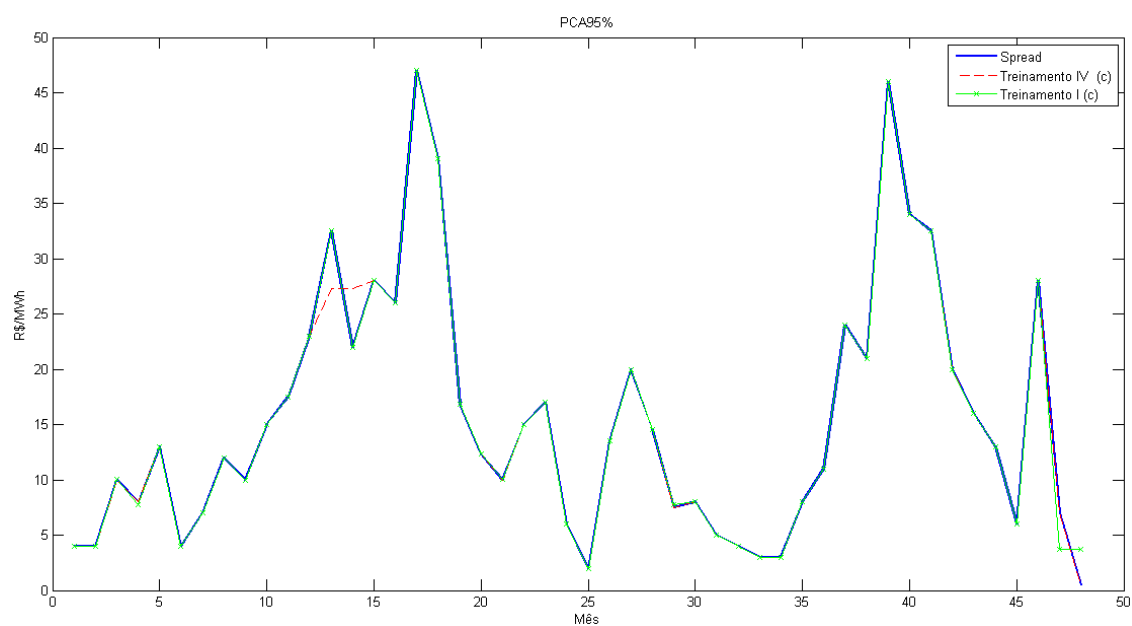


Figura 31 – Resultados da PMC com componentes principais de 95% da variância

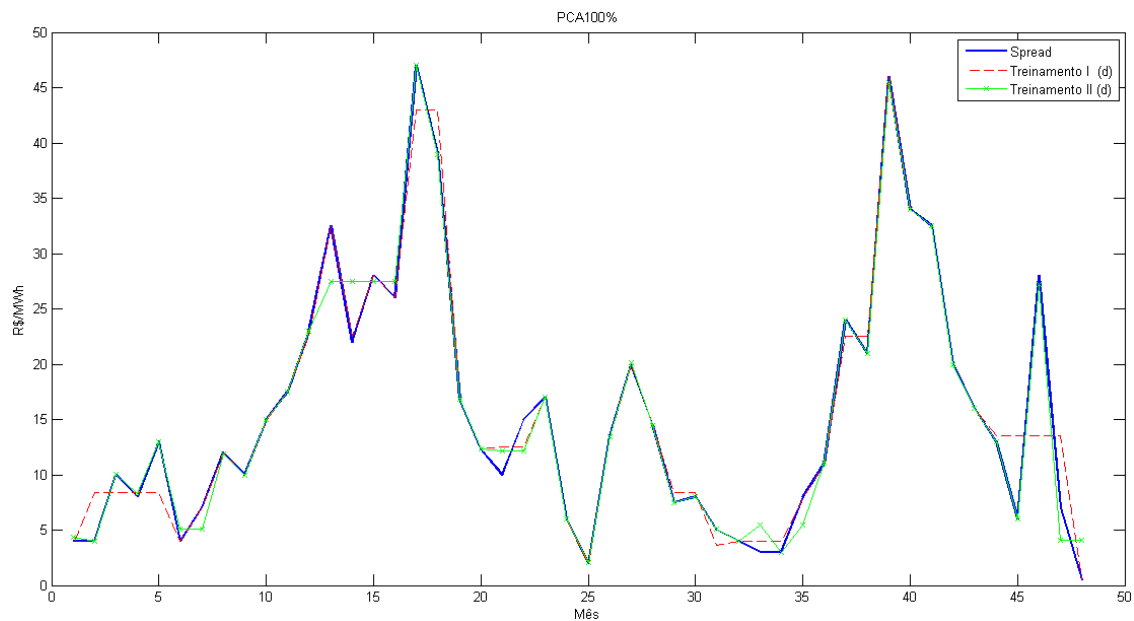


Figura 32 – Resultados da PMC com componentes principais de 100% da variância

A Tabela 11 exibe os erros para alguns dos treinamentos realizados para cada uma das análises feitas. É possível perceber que os erros são muito baixos.

Tabela 11 - Erros de ajuste das PMC com PCA como entradas

<i>Análise</i>	<i>Componentes Principais</i>	<i>% da Variância Explicada</i>	<i>Treinamento</i>	<i>MSE</i>	<i>MAE</i>	<i>MAPE</i>
(a)	9	85%	I	0,00	0,00	0,00%
			II	0,00	0,00	0,00%
			III	4,69	0,50	1,74%
			IV	0,00	0,00	0,00%
			V	2,12	0,57	6,29%
(b)	11	90%	I	0,00	0,00	0,00%
			II	17,05	1,26	58,56%
			III	1,96	0,44	9,92%
			IV	0,32	0,12	0,81%
			V	1,08	0,25	2,87%
(c)	15	95%	I	0,44	0,15	14,64%
			II	64,00	4,62	136,22%
			III	0,00	0,00	0,00%
			IV	1,15	0,22	0,83%
			V	0,00	0,00	0,00%
(d)	28	100%	I	8,47	1,30	13,73%
			II	2,29	0,72	21,80%
			III	0,00	0,00	0,00%
			IV	0,17	0,08	1,56%
			V	15,98	2,24	20,12%

4.1.4) Parada Prematura

Com o objetivo de se evitar que ocorra o *overfitting* durante o treinamento da PMC, foi decidido utilizar o método de parada prematura, ou *early stopping*. Este método é utilizado em conjunto com o *Back-Propagation* e consiste na utilização de conjuntos de dados que não são usados durante o treinamento da rede, mas são avaliados a cada época. Ao identificar que o erro entre a saída esperada e a saída obtida pelo conjunto de validação está maior que o mínimo global em um determinado número de épocas sucessivas, o treinamento é parado, já que a partir daí a rede está tendendo a memorizar os dados de treinamento.

Para o caso deste trabalho, os dados de janeiro a setembro de 2011 foram definidos como dados de validação da previsão e o número máximo de épocas sucessivas cujo erro é maior que o mínimo global foi definido como quatro. A Figura 33 exibe o treinamento de uma das redes através deste critério de parada, onde as linhas tracejadas se encontram no mínimo MSE calculado para os dados de validação.

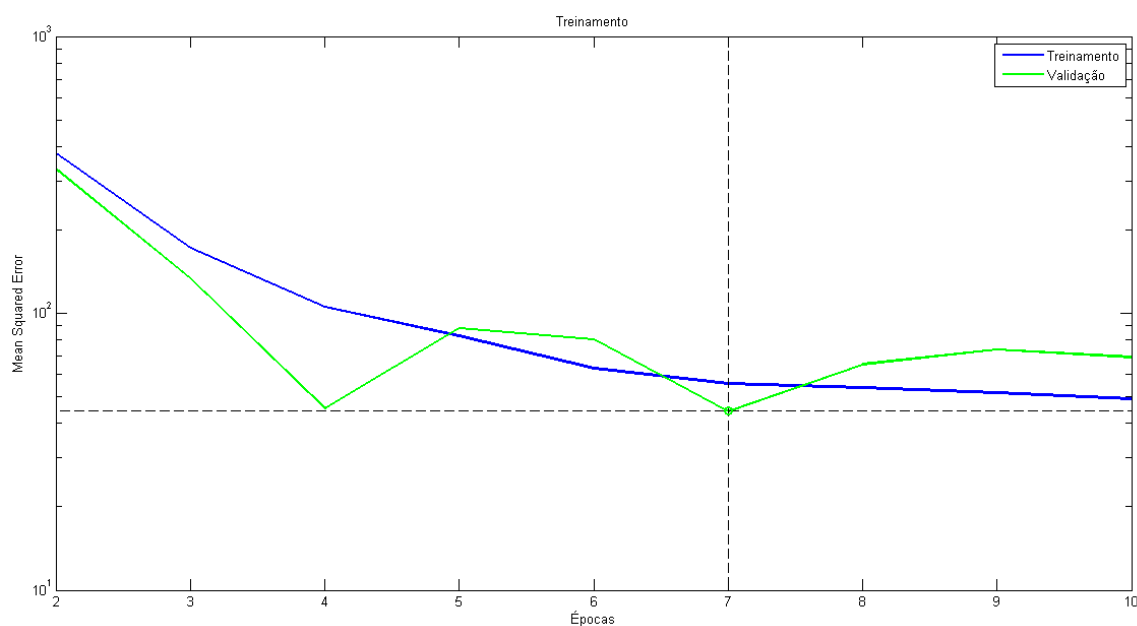


Figura 33 – Treinamento de PMC com critério de parada

Assim, as redes foram novamente treinadas para os quatro conjuntos de PCA, resultando respectivamente nos ajustes da Figura 34, Figura 35, Figura 36 e Figura 37.

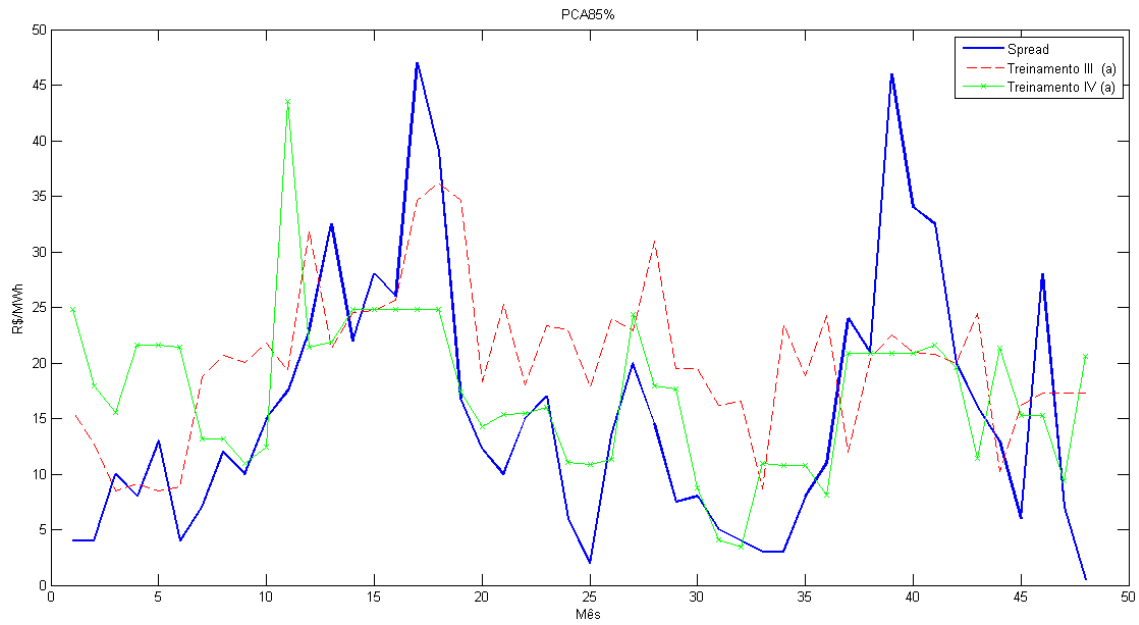


Figura 34 – Ajuste de PMC com PCA 85% e critério de parada

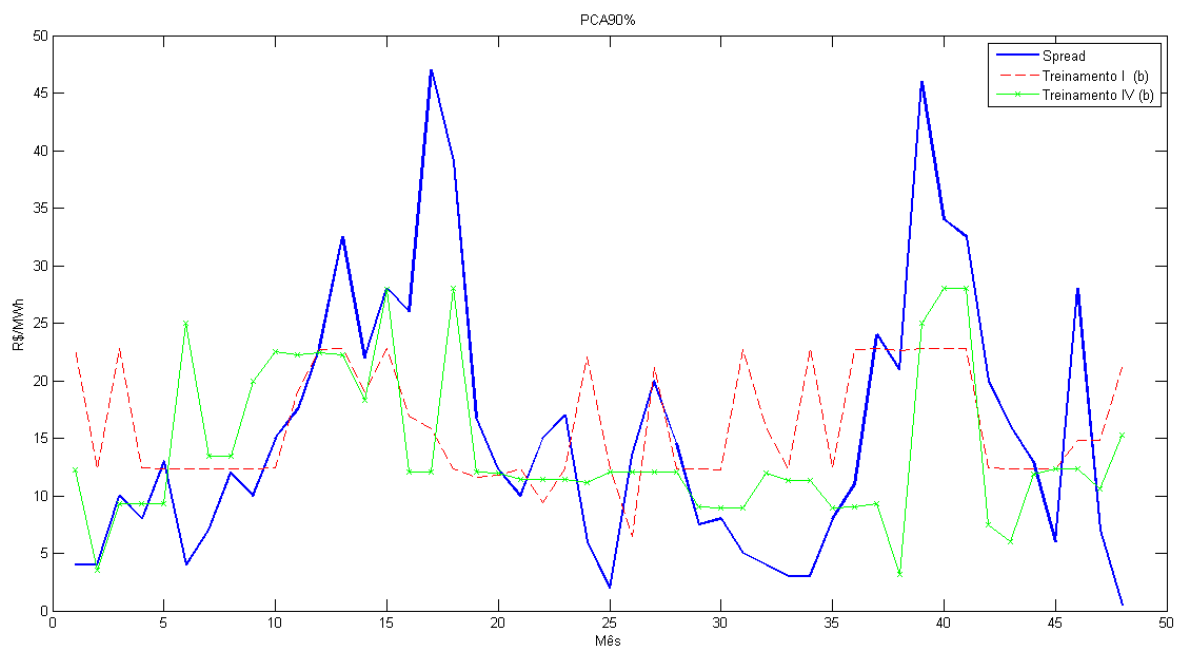


Figura 35 – Ajuste de PMC com PCA 90% e critério de parada

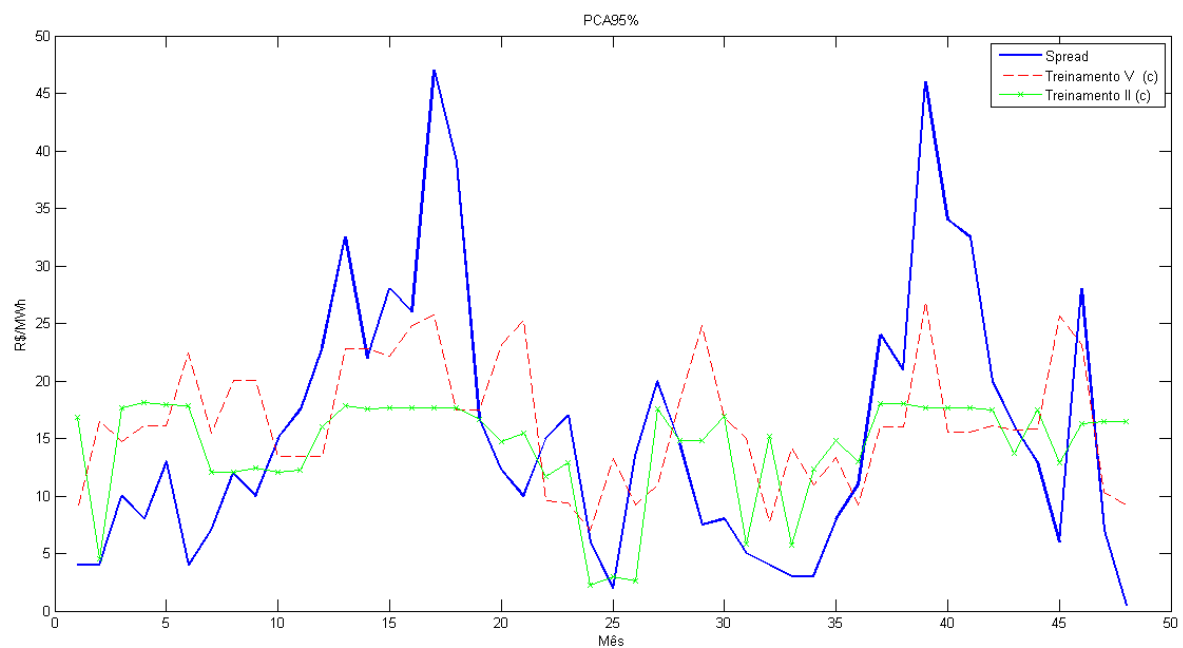


Figura 36 – Ajuste de PMC com PCA 95% e critério de parada

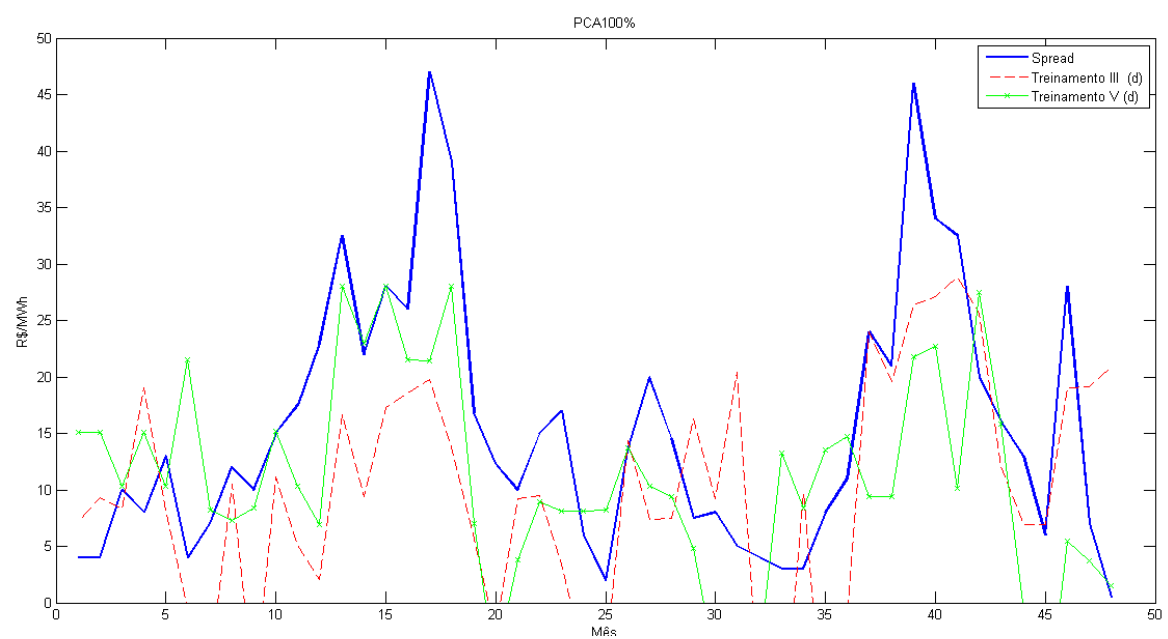


Figura 37 – Ajuste de PMC com PCA 100% e critério de parada

A Tabela 11 foi então refeita para os novos treinamentos com parada prematura, resultando na Tabela 12 onde percebe-se que os erros dos ajustes aumentaram significativamente devido à parada prematura, mas os erros de previsão, que serão exibidos no próximo capítulo, devem ser melhores que os modelos treinados sem o *early stopping*.

Tabela 12 - Erros de ajuste das PMC com PCA e critério de parada

<i>Análise</i>	<i>Componentes Principais</i>	<i>% da Variância Explicada</i>	<i>Treinamento</i>	<i>MSE</i>	<i>MAE</i>	<i>MAPE</i>
(a)	9	85%	I	129,48	9,18	205,78%
			II	48,93	4,81	117,73%
			III	115,87	9,20	183,39%
			IV	101,48	7,28	165,60%
			V	155,03	9,69	158,58%
(b)	11	90%	I	120,62	8,26	185,52%
			II	163,95	9,45	126,67%
			III	96,09	7,13	99,93%
			IV	97,90	7,18	134,69%
			V	84,44	6,47	132,62%
(c)	15	95%	I	161,58	10,17	249,90%
			II	101,62	7,61	133,82%
			III	42,78	5,15	80,41%
			IV	140,22	8,37	102,77%
			V	104,27	8,33	132,05%
(d)	28	100%	I	86,23	6,81	175,41%
			II	224,58	12,03	160,27%
			III	142,05	9,78	182,78%
			IV	81,95	7,49	132,20%
			V	117,65	8,51	92,99%

Capítulo 5

Análise das Previsões

Tendo todos os modelos treinados para a previsão do *Spread*, estes foram então testados com dados verificados entre Janeiro e Setembro de 2011. A Tabela 13 exhibe o resultado dos testes dividindo os modelos de acordo com suas características em cinco blocos distintos.

Tabela 13 – Erros de previsão dos modelos apresentados

<i>Bloco</i>	<i>Modelo</i>	<i>MSE</i>	<i>MAE</i>	<i>MAPE</i>	<i>Entradas</i>	<i>Treinamento</i>	<i>n°</i>
(a)	1 Naive	150,60	9,66	696%	$Spread_{t-1}$	-	-
	2 AR(1)	139,15	8,59	783%	$Spread_{t-1}$	Yule-Walker	-
	3 AR(1)	151,83	9,67	938%	$Spread_{t-1}$	MMQ	-
	4 AR(2)	124,27	7,86	292%	$Spread_{t-1,t-2}$	Yule-Walker	-
(b)	5 ARX	98,41	8,86	89%	Variáveis _{t-1}	MMQ	-
	6 ARX	261,77	13,67	298%	PCA 85%	MMQ	-
	7 ARX	267,79	14,73	567%	PCA 90%	MMQ	-
	8 ARX	597,12	23,48	453%	PCA 95%	MMQ	-
	9 ARX	768,56	23,86	709%	PCA 100%	MMQ	-
(c)	10 PMC	97,58	8,51	90%	$Spread_{t-1}$	<i>Back-Propagation</i>	I
	11 PMC	94,21	8,54	136%	$Spread_{t-1}$	<i>Back-Propagation</i>	II
	12 PMC	259,87	11,94	98%	$Spread_{t-1,t-2}$	<i>Back-Propagation</i>	I
	13 PMC	122,81	8,22	155%	$Spread_{t-1,t-2}$	<i>Back-Propagation</i>	IV
	14 PMC	60,69	5,70	88%	Variáveis _{t-1}	<i>Back-Propagation</i>	II
	15 PMC	203,15	11,78	630%	Variáveis _{t-1}	<i>Back-Propagation</i>	IV
(d)	16 PMC	206,08	11,81	163%	PCA 85%	<i>Back-Propagation</i>	I
	17 PMC	216,26	13,58	100%	PCA 85%	<i>Back-Propagation</i>	V
	18 PMC	234,69	12,54	93%	PCA 90%	<i>Back-Propagation</i>	III
	19 PMC	390,32	16,90	670%	PCA 90%	<i>Back-Propagation</i>	V
	20 PMC	445,48	17,35	117%	PCA 95%	<i>Back-Propagation</i>	I
	21 PMC	435,01	17,16	1.208%	PCA 95%	<i>Back-Propagation</i>	IV
	22 PMC	383,77	14,93	83%	PCA 100%	<i>Back-Propagation</i>	I
	23 PMC	126,62	8,39	215%	PCA 100%	<i>Back-Propagation</i>	II
(e)	24 PMC	94,80	7,39	42%	PCA 85%	<i>Early Stopping</i>	III
	25 PMC	34,44	4,81	46%	PCA 85%	<i>Early Stopping</i>	IV
	26 PMC	29,06	4,52	29%	PCA 90%	<i>Early Stopping</i>	I
	27 PMC	109,70	7,96	50%	PCA 90%	<i>Early Stopping</i>	V
	28 PMC	56,54	6,13	38%	PCA 95%	<i>Early Stopping</i>	II
	29 PMC	102,59	7,85	52%	PCA 95%	<i>Early Stopping</i>	V
	30 PMC	84,41	7,67	90%	PCA 100%	<i>Early Stopping</i>	III
	31 PMC	92,12	7,40	141%	PCA 100%	<i>Early Stopping</i>	V

Percebe-se que apesar da simplicidade dos modelos *Naive* e Autorregressivos do bloco (a), estes obtiveram bons resultados quando comparados a modelos mais complexos como os modelos Autorregressivos Exógenos que utilizam PCA do bloco (b) e as redes neurais do bloco (d). A boa performance destes modelos pode estar ligada à baixa variação do *Spread* durante o

período de teste, já que os modelos deste bloco dependem apenas do seu histórico. Os resultados são exibidos na Figura 38.

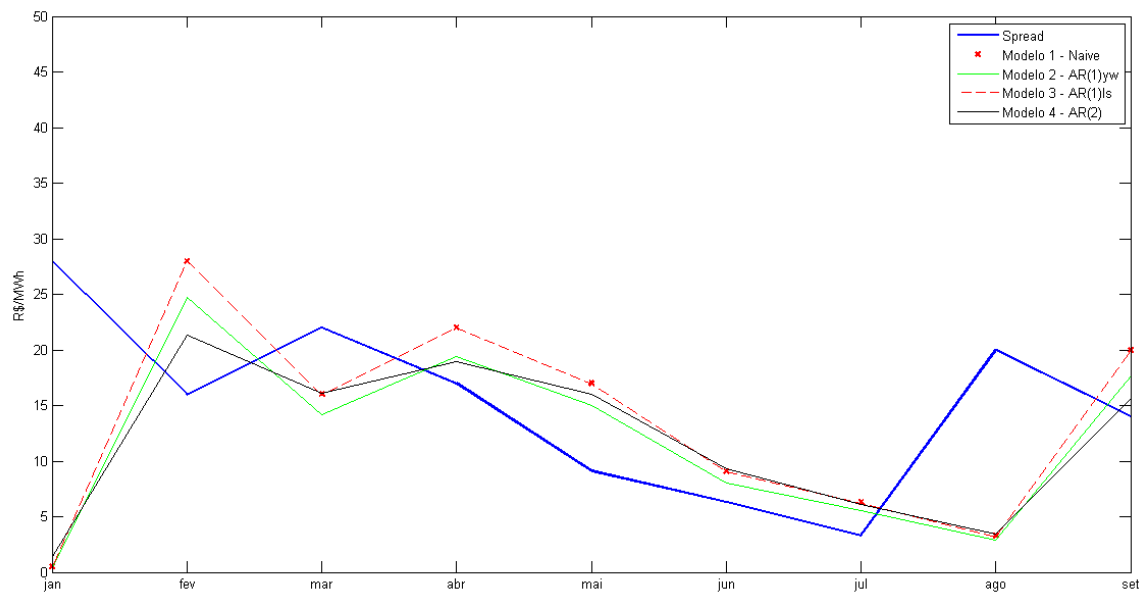


Figura 38 – Previsões do bloco (a)

Dentre os modelos ARX exibidos no bloco (b), o modelo 5, que utiliza as variáveis originais, teve um melhor desempenho que os modelos que utilizaram como entrada os resultados da PCA. Este comportamento pode ser explicado devido à grande quantidade de dados que foram apresentados aos modelos ARX no segundo caso, onde a Análise dos Componentes Principais não foi capaz de reduzir a quantidade de variáveis do problema para um número baixo o bastante, tornando o problema complexo demais para o modelo ARX.

Percebe-se que o desempenho das ARX dos modelos 6 a 9 é inversamente proporcional à quantidade de componentes principais utilizadas como entrada, já que apesar de se aumentar a quantidade de informações para o treinamento, o aumento do nível de complexidade do problema é muito maior, prejudicando o desempenho do modelo. A Figura 39 exibe os resultados dos testes para o bloco (b).

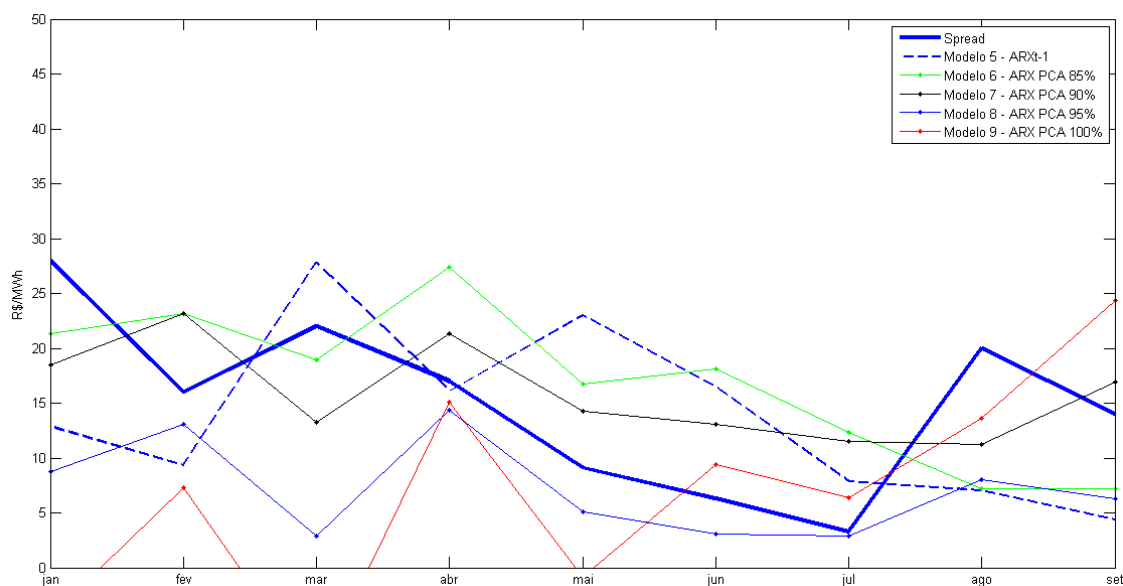


Figura 39 – Previsões do Bloco (b)

No bloco (c) são exibidas as redes PMC com as entradas mais simples e que produzem, em geral, uma previsão melhor que os modelos lineares, mas devido a seu treinamento iniciar de forma aleatória, uma mesma forma de treinamento e um mesmo conjunto de entradas pode resultar em modelos com grandes erros, como o caso do modelo 15, e em modelos com baixos erros, como o caso do modelo 14. A utilização das PMCs com poucas variáveis obteve bons resultados e é possível obter ainda melhores resultados alterando-se empiricamente as formas de treinamento e seus critérios, de forma que a rede se adapte melhor ao problema. A Figura 40 exhibe estes resultados.

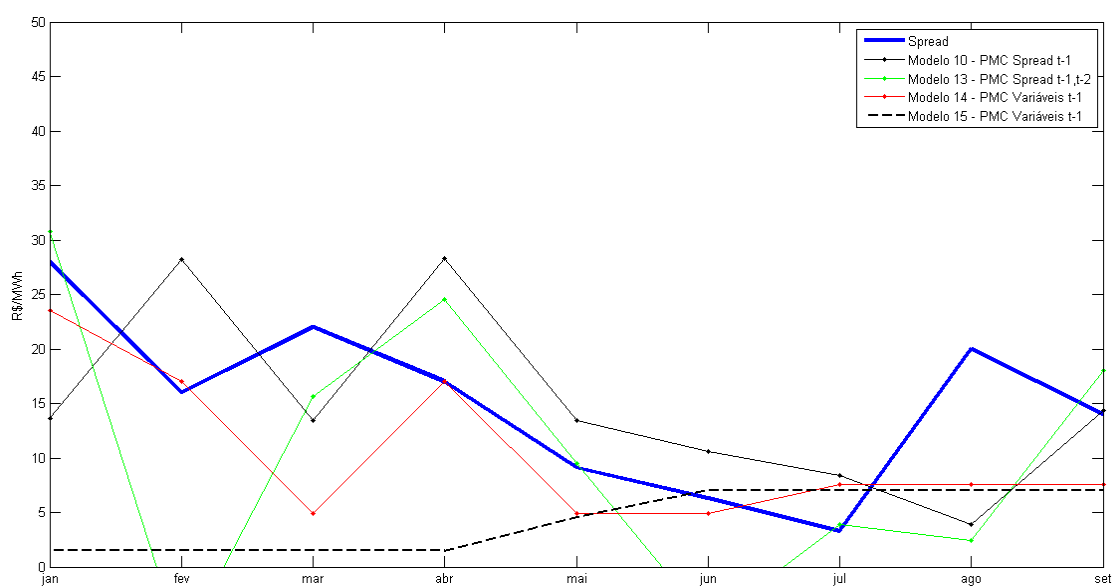


Figura 40 – Previsões do bloco (c)

O bloco (d) exibe as redes PMC sem parada prematura cujas entradas foram os componentes principais. É possível visualizar o resultado do *overfitting*. Apesar dos erros observados durante o treinamento serem praticamente nulos, os erros obtidos na previsão de novos dados foram excessivamente altos. Nos resultados gráficos do bloco (d) na Figura 41 e do bloco (e) na Figura 42 é nítida a diferença entre os treinamentos com e sem a parada prematura.

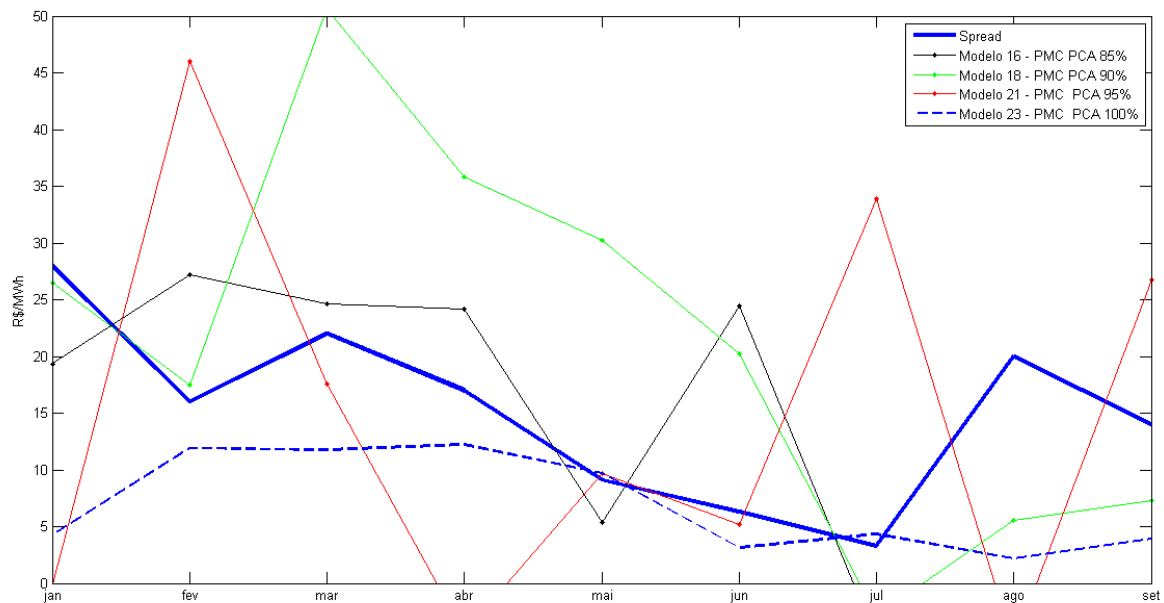
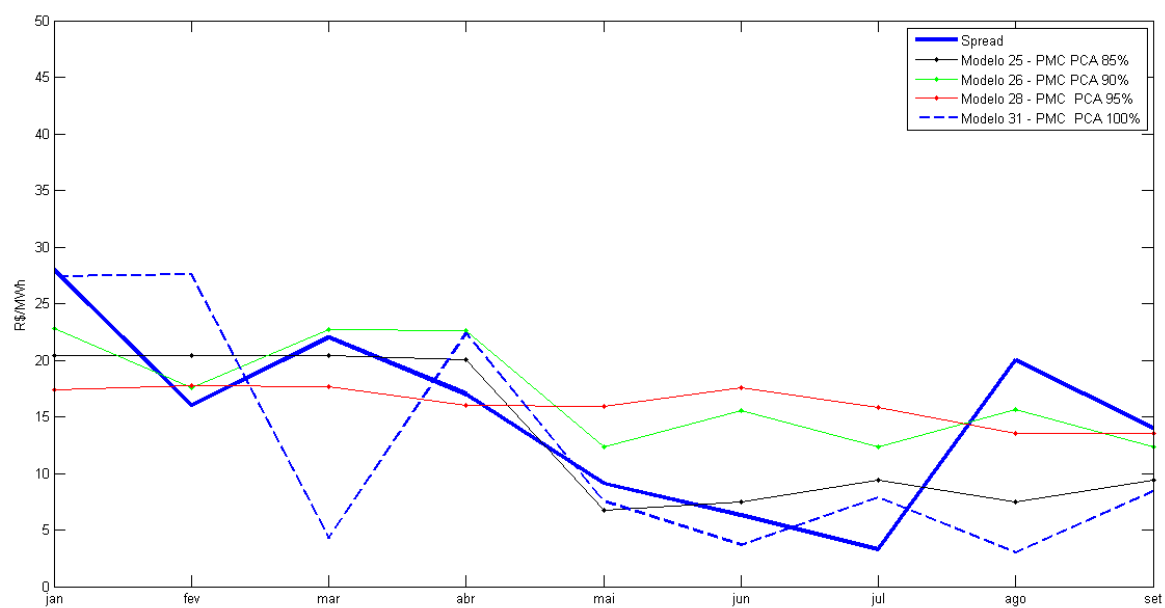


Figura 41 – Previsões do bloco (d)

O bloco (e), com as redes PMC com parada prematura, é o bloco onde se observa um menor erro geral entre todos os estimadores, tendo um grande potencial para a previsão do *Spread*, já que seus resultados podem ser melhorados através da variação de seu treinamento e de suas entradas, ao contrário dos modelos do bloco (a). Seu resultado é exibido na Figura 42.

**Figura 42 – Previsões do bloco (e)**

Capítulo 6

Conclusão e Discussões

Através dos resultados gráficos e numéricos exibidos no Capítulo 5, é possível perceber que em geral, os modelos dos blocos (a) e (e), além do ARX que utiliza como entrada as sete variáveis do sistema elétrico defasadas em um mês foram os modelos que obtiveram os resultados mais satisfatórios durante o teste. Devido à fácil aplicação e bom desempenho, os modelos Autorregressivos de primeira e segunda ordem podem ser aplicados na estimação do *Spread*, mas infelizmente não existe muita possibilidade de melhora no resultado destes modelos.

O mesmo ocorre com o modelo ARX, mas este tem uma capacidade maior de generalização do problema, devido justamente à inserção das variáveis exógenas, que certamente influenciam no *Spread*. Este modelo também tem uma fácil aplicação, sendo necessário apenas a obtenção dos dados históricos do sistema, facilmente adquiridos no site do ONS.

Diferentemente, os modelos que utilizam as redes neurais de arquitetura Perceptron Multicamadas com um maior número de informação, como é o caso dos modelos que utilizam como entrada as variáveis históricas ou os resultados da análise dos componentes principais, têm sim um grande potencial de melhora, sendo necessário variar empiricamente os métodos e configurações de treinamento afim de um melhor ajuste ao problema.

No caso das arquiteturas que utilizam a PCA, é interessante também a variação do conjunto de dados de treinamento, validação e teste para o critério de parada prematura, sendo este o grande fator de generalização no caso destas redes neurais. Outro aspecto que poderia ser avaliado também neste caso é a retroalimentação da rede, ou seja, a utilização dos resultados de previsão atual para dois ou mais meses à frente.

Outra indicação para trabalhos futuros seria utilizar diversas combinações das variáveis aqui apresentadas, além da adição de outras possíveis variáveis do sistema elétrico.

Tendo em vista que o *Spread* é uma variável bastante volátil e que sua relação com as outras variáveis do sistema elétrico nacional é difícil de ser mensurada, a sua previsão dificilmente atingirá erros mínimos, mas o objetivo de sua previsão é a utilização de seus resultados na seleção da melhor estratégia na comercialização de curtíssimo prazo, decidindo entre a venda do lastro positivo de doze meses no caso de uma expectativa de baixa no *Spread*, ou a manutenção deste lastro no caso de expectativa de alta.

A previsão do *Spread* também pode ser associada aos resultados dos modelos NEWAVE e DECOMP, na projeção do patamar em que o preço final da energia elétrica estará nos próximos meses no mercado de curto prazo, informação esta crucial na operação comercial dos agentes do mercado nacional de energia.

BIBLIOGRAFIA

Brockwell, Peter and Davis (2006). Time Series: Theory and Methods, 2nd ed. Springer

CCEE - O Processo de Comercialização. Desenvolvido por Câmara de Comercialização de Energia Elétrica, 2011. Disponível em: <<http://www.ccee.org.br/cceeinterdsm/v/index.jsp?vgnextoid=dbaaa5c1de88a010VgnVCM100000aa01a8c0RCRD#>>. Acesso em: 24 set. 2011.

Jenkins, G. M, and Box, G. E. P. (1994). Time Series Analysis: Forecasting and Control, 3rd edition. Prentice Hall PTR.

ONS – Histórico da Operação. Disponível em: <<http://www.ons.org.br/>>. Acesso em 15 out. 2011.

Pearson, K. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. Philosophical Magazine.

Sacchi, R. (2009). Política de Operação Preditiva Estabilizada via Termo Inercial Utilizando “Analytic Signal”, “Dynamic Modelling” e Sistemas Inteligentes na Previsão de Vazões Afluentes em Sistemas Hidrotérmicos de Potência. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2009.

Silva, I. N. da (2010). Redes Neurais Artificiais para Engenharia e Ciências Aplicadas, Artliber.

Yule, G. Udny (1927). On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers, Philosophical Transactions of the Royal Society of London.

Walker, G. (1931). On Periodicity in Series of Related Terms, Proceedings of the Royal Society of London.

APÊNDICE A

Estimadores de Erro

Estimadores de erro são utilizados para avaliar a qualidade de medidas e previsões, comparando os valores experimentais com os reais e verificando suas diferenças. Normalmente são aplicados quando existem dois ou mais modelos a serem avaliados, onde suas performances são comparadas para cada estimador, a fim de se decidir sobre o melhor modelo.

É grande a diversidade de estimadores e a seleção depende do objetivo da análise a ser feita. Para o caso deste trabalho, foram selecionados três estimadores clássicos: o Erro Quadrático Médio (*Mean Squared Error – MSE*), o Erro Médio Absoluto (*Mean Absolute Error – MAE*) e o Erro Percentual Absoluto Médio (*Mean Absolute Percentual Error*).

Erro Quadrático Médio

É definido pela média do quadrado da diferença entre o valor real e o valor estimado, resultando portanto sempre em um valor positivo. Este estimador elimina o problema encontrado em outros estimadores, como o Erro Médio, onde um valor baixo nem sempre indica uma boa performance, pois caso ocorram valores de erro negativos e positivos, estes acabam se anulando e diminuindo seus efeitos.

O MSE pode ser representado pela equação:

$$MSE = \frac{\sum_{n=1}^N (x_n - \bar{x}_n)^2}{N} \quad (19)$$

sendo, para cada medição n:

MSE	Erro Quadrático Médio;
x_n	Valor real;
$\bar{x}_n$	Valor previsto;
N	Quantidade total de medições.

Apesar de ser um bom estimador, onde menores valores representam melhores resultados, o valor final do MSE não tem uma representação muito clara e pode confundir a avaliação, devido aos seus altos resultados causados pelo termo quadrático.

Erro Médio Absoluto

Enquanto o Erro Quadrático Médio é um bom estimador devido sua habilidade de ignorar falsas sensações de bom desempenho, seus valores não são facilmente interpretados. O Erro Médio Absoluto, porém contempla as duas características, sendo definido como a média do módulo da diferença entre o valor real e o valor estimado, como representado na equação a seguir:

$$MAE = \frac{\sum_{n=1}^N |x_n - \overline{x_n}|}{N} \quad (20)$$

sendo, para cada medição n:

MAE	Erro Médio Absoluto;
x_n	Valor real;
$\overline{x_n}$	Valor previsto;
N	Quantidade total de medições.

Os valores do MAE são mais representativos que os do MSE, por estarem na mesma escala que os dados averiguados, sendo possível o identificar como sendo o erro esperado para o modelo.

Erro Percentual Absoluto Médio

O terceiro estimador de erro utilizado é o Erro Percentual Absoluto médio, estimador que divide a opinião dos especialistas, já que este tem seu valor muito afetado por altos valores da variável real em casos onde a série tem uma grande variância, pois representa a variação percentual do valor previsto e do valor esperado. O MAPE pode ser representado pela

(21)

seguinte equação:
$$MAPE = \frac{\sum_{n=1}^N \left| \frac{x_n - \bar{x}_n}{x_n} \right|}{N} \times 100\%$$

sendo, para cada medição n:

MAPE Erro Percentual Absoluto Médio;

x_n Valor real;

$\bar{x}_n$ Valor previsto;

N Quantidade total de medições.

O MAPE representa a média percentual dos erros em relação ao valor esperado, mas no caso de uma série com grandes picos, seu valor pode acabar distorcendo e dando uma falsa sensação de baixo desempenho do modelo.